Laboratoire d'informatique de Paris Nord

LIPN

Rapport d'activité 2017-2022 et projet 2023-2028

cnrs

UNIVERSITÉ SORBONNE PARIS NORD

Couverture : © Olivier Bodini, « Twitter 7 mars », mars 2023.

Partie I — **Le laboratoire** — **page 5**

9 Informations générales
27 Portfolio
29 Bilan
53 Trajectoire

Partie II — **A^3** — **page 63**

69 Informations générales sur l'équipe A3
83 Introduction du portfolio
85 Auto-évaluation de l'équipe A3
107 Trajectoire de l'équipe A3

Partie III — **AOC** — **page 117**

123 Informations générales sur l'équipe AOC
137 Introduction du portfolio
139 Auto-évaluation de l'équipe AOC
161 Trajectoire de l'équipe AOC

Partie IV — **CALIN** — **page 169**

175 Informations générales sur l'équipe CALIN
189 Introduction du portfolio
193 Auto-évaluation de l'équipe CALIN
217 Trajectoire de l'équipe CALIN

Partie V — **LoVe** — **page 225**

229 Informations générales sur l'équipe LoVe
243 Introduction du portfolio
245 Auto-évaluation de l'équipe LoVe
275 Trajectoire de l'équipe LoVe

Partie VI — **RCLN** — **page 285**

291 Informations générales sur l'équipe RCLN
303 Introduction du portfolio
305 Auto-évaluation de l'équipe RCLN
329 Trajectoire de l'équipe RCLN

Partie VII — **Acronymes** — **page 337**

Partie I
Le laboratoire

Sommaire

Informations générales **9**
1.1 Identification de l'unité . . . 10
1.2 Présentation de l'unité . . . 10
1.3 Les thématiques scientifiques et leurs enjeux . . . 19
1.4 Profil d'activités liées à la recherche . . . 23
1.5 Environnement de recherche . . . 23
1.6 Prise en compte des recommandations du précédent rapport . . . 25

Portfolio **27**

Bilan **29**
3.1 Domaine 1. Profil, ressources et organisation de l'unité . . . 30
3.2 Domaine 2. Attractivité . . . 36
3.3 Domaine 3. Production scientifique . . . 45
3.4 Domaine 4. Inscription des activités de recherche dans la société . . . 49

Trajectoire **53**
4.1 Atouts, Faiblesses, Opportunités et Menaces . . . 54
4.2 Positionnement relativement au projet 2019-2023 . . . 56
4.3 Projet pour la période 2023-2028 . . . 56

Chapitre 1

Informations générales pour le contrat en cours

Sommaire

1.1 Identification de l'unité 10
1.2 Présentation de l'unité 10
1.2.1 Historique, localisation de l'unité. 10
1.2.2 Organisation de l'unité 12
1.2.3 Équipes, plateformes, services communs 14
1.2.4 Effectifs de l'unité et de ses équipes au 31/12/2022 18
1.2.5 Thématiques scientifiques (par équipe le cas échéant) 19
1.3 Les thématiques scientifiques et leurs enjeux 19
1.3.1 Stratégie scientifique 19
1.3.2 Structuration en 5 équipes et interactions 20
1.3.3 Transversalité et interactions 22
1.3.4 Animation scientifique 22
1.4 Profil d'activités liées à la recherche 23
1.5 Environnement de recherche 23
1.6 Prise en compte des recommandations du précédent rapport . . . 25

1.1 Identification de l'unité

Nom de l'unité : Laboratoire d'Informatique de Paris Nord

Acronyme : LIPN

Label et numéro : UMR 7030

Domaine scientifique principal : Informatique

Panels scientifiques par ordre décroissant de pertinence.

Panel 1. ST6 : Sciences et Technologies de l'Information et de la Communication - STIC

Panel 2. ST1 : Mathématiques

- **Équipe de direction**

— Frédérique Bassino, Directrice du LIPN depuis le 01/01/2018.
— Laure Petrucci, Directrice de LIPN du 01/01/2012 au 31/12/2017.

Le LIPN ne compte pas d'équipe de direction à proprement parler, il n'y a pas non plus de direction adjointe. Toutefois l'unité compte, depuis octobre 2022, une responsable administrative, Véronique Criart.

Le conseil de laboratoire est très régulièrement consulté sur les décisions à prendre et les questions de politique scientifique.

Tutelles de l'unité : CNRS, Université Sorbonne Paris Nord

1.2 Présentation de l'unité

1.2.1 Historique, localisation de l'unité.

Le Laboratoire d'Informatique de Paris-Nord (LIPN[1]) de l'Université Sorbonne Paris Nord, créé en 1985, est associé au CNRS depuis janvier 1992 et a le statut d'UMR depuis janvier 2001. Au CNRS, le LIPN est rattaché à l'Institut National des Sciences de l'Information et de leurs Interactions (INS2I).

Les locaux du LIPN sont situés sur le campus de Villetaneuse dans les bâtiments de l'Institut Galilée. Ils sont répartis sur 3 niveaux et sont très insuffisants au regard des besoins (1147 m^2 SU). Toutefois un nouveau bâtiment est en cours de construction sur le campus.

Ce nouveau bâtiment accueillera l'ensemble de la fédération MathSTIC qui regroupe, outre le LIPN, le Laboratoire d'Analyse, Géométrie et Applications (LAGA UMR CNRS 7539) et le Laboratoire de Traitement et Transport de l'Information (L2TI UR 3043), soit un ensemble de 400 chercheurs en mathématiques et en informatique. Cela s'inscrit dans la dynamique de constitution d'un grand pôle de recherche sur cette thématique au nord de Paris.

La première moitié du bâtiment bénéficie d'un financement de 13,18 M€ dans le cadre du CPER 2015-2020, la seconde moitié d'un financement de 18,1 M€ dans le cadre du CPER 2021-2027. Le CNRS a abondé au budget pour un montant de 300k€. Le bâtiment est en cours de construction et devrait être livré entre le printemps 2024 (1ère tranche) et l'automne 2025 (2ème tranche).

1. http://lipn.univ-paris13.fr. (On trouvera en annexe la liste des acronymes et leur signification)

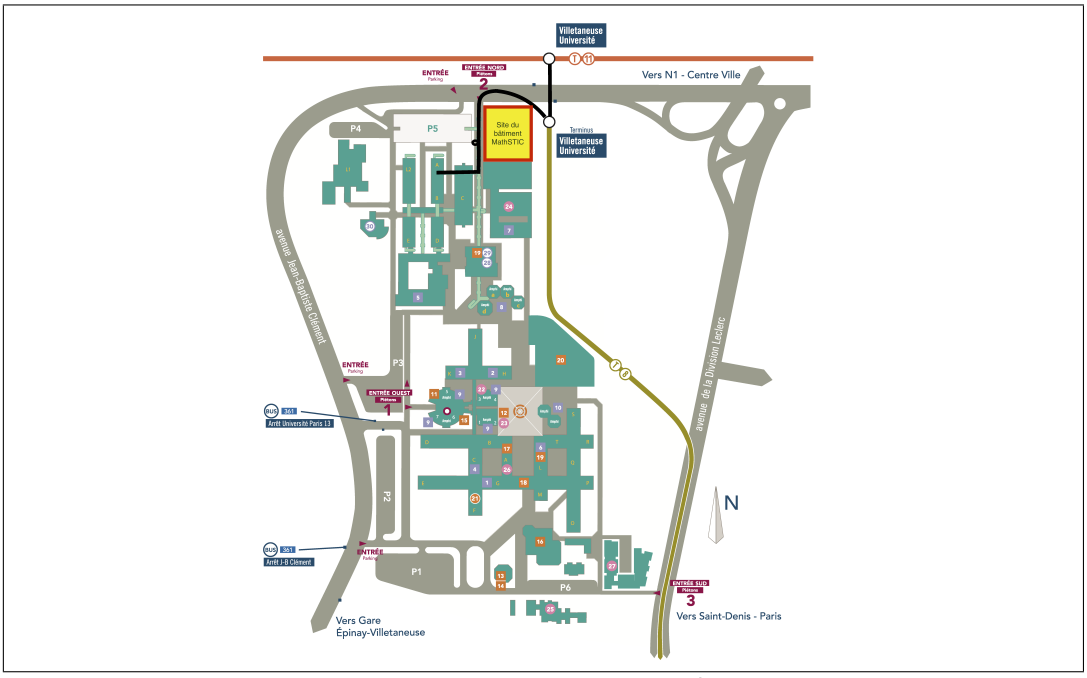


Figure 1.1: Situation futur du bâtiment

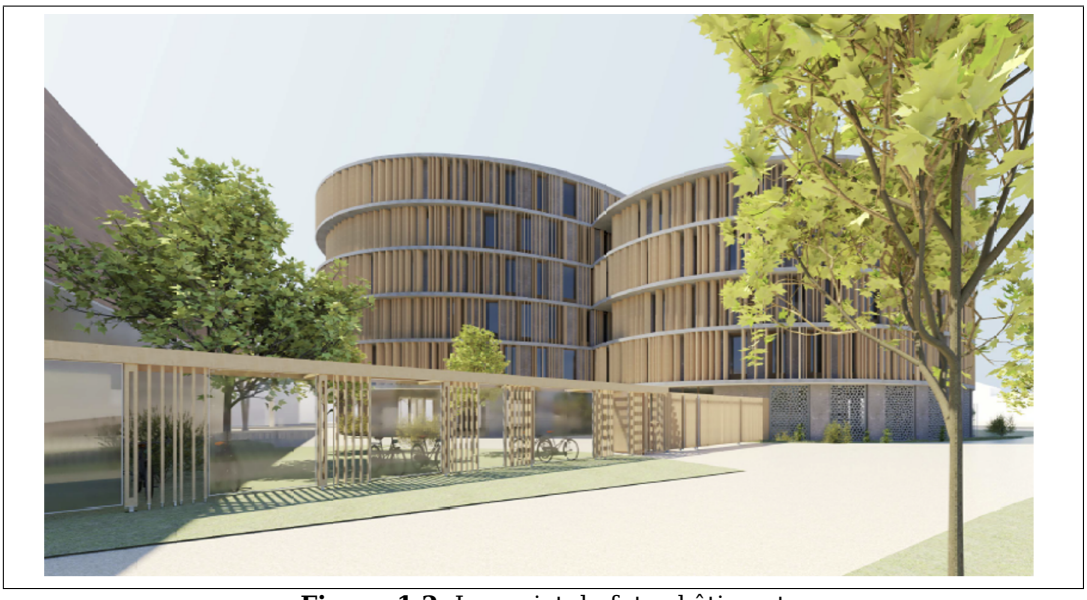


Figure 1.2: Le projet de futur bâtiment

Le LIPN occupera les 4 étages de bureaux de la première tranche (2 151 m^2 SU), dont le rez-de chaussée est dédié aux salles de séminaires et réunions, ainsi qu'à la salle serveurs (soit 20% de la SU de cette tranche).
Le manque cruel de bureaux et d'espaces de réunion devrait ainsi être résolu à la rentrée 2024.

1.2.2 Organisation de l'unité

Le LIPN est organisé autour de 5 équipes scientifiques, complétées par un axe transverse, et 2 équipes support (cf. section 1.2.3). Le conseil de laboratoire y tient un rôle central. Il est assisté dans son travail par plusieurs commissions, dont le rôle est précisé dans ce qui suit, et par un réseau de correspondants.

- **Les équipes de recherche.**

La recherche du LIPN se structure en cinq équipes dont les grandes thématiques sont l'apprentissage automatique (A³), l'optimisation combinatoire et le calcul haute performance (AOC), la conception et l'analyse de modèles combinatoires à l'interface de la physique et de l'algorithmique (CALIN), les fondements du calcul et la vérification formelle (LoVe), le traitement automatique du langage naturel et la représentation des connaissances (RCLN).

La ou le **responsable d'équipe** est nommé.e par la directrice du LIPN après consultation de tous les membres de l'équipe. Il ou elle assure l'animation scientifique et relaie les informations entre la direction du LIPN et les membres de l'équipe. Il ou elle valide du point de vue scientifique les demandes de missions, de stages et les proposition de thèse. Il ou elle met en place l'accompagnement des membres de l'équipe dans leur évolution scientifique et professionnelle.

Aux cinq équipes de recherche s'ajoute un axe transverse qui réunit les membres de l'unité ayant une expertise sur différentes formes de complexité (algébrique, descriptive, d'approximation, de preuve). Il est à noter que, de manière plus générale, les collaborations, la porosité entre équipes se sont très fortement renforcées dans la période (cf. section 1.6)

L'ensemble de ces recherches sont effectuées par 74 chercheurs CNRS, ingénieurs de recherche CNRS et enseignants-chercheurs permanents en poste dans différentes composantes de l'USPN (pour moitié à l'Institut Galilée et pour l'autre moitié à l'IUT de Villetaneuse ou à l'IUT de Saint-Denis). Au-delà c'est en fait environ 150 personnes qui participent aux activités du laboratoire sur le site de Villetaneuse.

- **La gouvernance de l'unité.**

Le LIPN n'ayant pas d'équipe de direction, le **conseil de laboratoire**, qui se réunit au moins une fois par mois, joue un rôle central dans l'unité. C'est le lieu où sont discutées les orientations stratégiques du laboratoire. Son avis, éclairé par les travaux de diverses commissions, prévaut sur les questions suivantes : gestion budgétaire, appels d'offre, classement ou avis sur les demandes d'allocations de recherche, d'inscription en doctorat, choix des profils de poste, propositions pour les demandes de moyens annuelles à l'Université et au CNRS, pour les demandes de professeurs invités ou encore pour les réponses aux appels à projets des tutelles. Ce fonctionnement collégial permet de maintenir une adhésion collective pour des décisions quelquefois difficiles (comme les choix de profils de postes).

Des **assemblées générales** sont convoquées quand cela s'avère pertinent. Il y en a au moins deux par an et généralement davantage. Elles sont toujours suivies d'un moment de convivialité qui permet de poursuivre de manière plus informelle les échanges entre les membres de l'unité.

Le **conseil scientifique** aborde des sujets scientifiques ou organisationnels dans une perspective à moyen terme avec des avis extérieurs éclairés. Il comprend la directrice du LIPN, les responsables d'équipe, trois membres extérieurs (pour la durée du contrat), auxquels s'ajoutent des membres extérieurs apportant une expertise sur les thèmes traités. Il s'est réuni trois fois sur la période.

Le **comité des thèses** assure le suivi des thèses, sans empiéter sur le rôle du directeur de thèse, et l'animation doctorale.

— Il a un rôle d'interface avec l'École Doctorale Galilée,
— Il est en charge de l'examen des demandes d'inscription en thèse et de l'organisation des auditions des candidats.
— Il met en place des entretiens annuels des doctorants.
— Il participe aux comités de suivi.
— Il participe aux évènements doctoraux (organisation d'une réunion annuelle des doctorants, accompagnement du séminaire junior, implication dans les journées de l'Ecole doctorale, *etc.*).
— Il est, de manière générale, l'interlocuteur privilégié des doctorants et encadrants sur les questions et problèmes d'ordre administratif ou personnel.

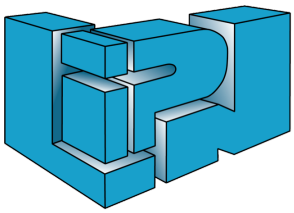
La **commission relations internationales**, qui inclut le correspondant Europe du LIPN, Thomas Seiller (CR CNRS), a pour objectif de favoriser le montage de projets internationaux. Elle joue un rôle informatif sur les appels à projets et incitatif au montage de projets. Cette commission propose également un classement des demandes de professeurs invités (le LIPN bénéficie d'une quinzaine d'invitations de 15 jours à 1 mois financées par l'USPN).

La **cellule communication et web** organise la communication interne et externe du laboratoire. En termes de communication interne, un journal du LIPN a été édité de 2019 à 2022. Son succès dans l'unité incite a réfléchir à un format de communication moins touffu mais plus fréquent.

En termes de communication externe, l'organisation d'un reportage photo, la mise en place d'un nouveau site web, la présence et animation sur les réseaux sociaux sont les principales actions mises en place dans la période.

En 2020, l'unité a organisé, avec la cellule communication de l'INS2I, un *reportage photo*. Un photographe professionnel a passé deux jours à mettre en scène des illustrations des recherches faites au LIPN. Après quelques retards, les images devraient être enfin intégrées dans la base de données CNRS Images et être utilisables pour communiquer sur les activités du laboratoire. Certaines photos ont déjà été utilisées dans des plaquettes du CNRS.

En 2022, le LIPN a mis en ligne un nouveau *site web*, plus ergonomique, plus réactif et mieux adapté aux besoins. Cela a été l'occasion de revoir la charte graphique, et de changer le logo qui datait de la création du laboratoire.



L'ancien logo

Le nouveau logo

Le LIPN a aussi redynamisé sa présence sur les *réseaux sociaux*. Dans un premier temps sur *Twitter*, puis, à la suite du rachat du réseau social par Elon Musk, le LIPN a ouvert, sur ses serveurs, sa propre instance *Mastodon* ouverte à la communauté des chercheurs et chercheuses en informatique en France. Bien évidemment, le LIPN y anime son propre compte. Les nouvelles partagées sur les réseaux sociaux sont typiquement les prix, les événements locaux ou internationaux organisés par des membres du laboratoire, des publications ou invitations prestigieuses, des créations ou récompenses de startup hébergées, ainsi que les présentations des nouveaux arrivés permanents ou non permanents.

La **commission développement logiciel** animée par Jaime Arias (IR CNRS), le responsable de la **`dev-team`** (cf. section 1.2.3), est en charge
— d'améliorer la valorisation de la production logicielle au LIPN,
— de dispenser des formations et partager les bonnes pratiques de développement,
— et de contribuer à l'intensification du développement logiciel au LIPN.
Elle a notamment été à l'initiative de formations régulières sur des outils de base dispensées par des membres du LIPN et essaie de diffuser les bonnes pratiques du développement. Elle compte amplifier sont travail de formation et mieux contribuer à la visibilité des logiciels développés au LIPN (cf. section 4.3.6).

Une **commission pour la parité** femmes-hommes a été créée l'hiver dernier. Animée par la correpondante parité du LIPN, Marie Kerjean (CR CNRS), elle organise des réunions en non-mixité, pour ouvrir un espace de parole sécurisé permettant de remonter d'éventuels soucis au niveau du laboratoire. De nombreuses initiatives se mettent en place (voir Section 4.3.8). Il est à noter qu'une membre du LIPN a coorganisé en 2021 et 2022 les deux premières éditions du *Workshop on Language Technology for Equality, Diversity and Inclusion (LT-EDI)* [ED-9, ED-13].

Une **commission développement durable** vient de se mettre en place. Animée par le correpondant développement durable du LIPN, Étienne André (PU USPN), elle est constituée d'un·e membre par équipe (y compris les équipes technique et administrative). L'objectif est de proposer des actions afin d'améliorer la responsabilité environnementale du laboratoire. Divers projets sont en cours (voir Section 4.3.9).
La commission interagit également avec celle de l'Institut Galilée (deux membres appartiennent aux deux commissions) et relaie au sein du LIPN ses actions (création d'une bibliothèque de livres en lien avec l'écologie et le développement durable, mise en place du tri sélectif au sein du LIPN, *etc*).

En sus des correspondant·e·s parité, Europe et développement durable, le LIPN s'est doté des correspondant·e·s suivant·e·s :
— Partenariat et Valorisation : Hanène Azzag (MCF USPN) – cf. section 3.4.1
— Formation : Jaime Arias (IR CNRS) – cf. sections 1.2.3 et 3.2.2
— Sécurité des systèmes d'information : Marisol Rodriguez Perez (IGE USPN) – cf. section 1.2.3
— Intégrité scientifique : Joseph Le Roux (MCF USPN) – cf. section 3.2.2
— Science ouverte : Étienne André (PU USPN) – cf. section 3.2.2

L'unité compte également deux assistantes de prévention :
— Aude Grezka (IR CNRS).
— Antonia Wilk (Tech. CNRS).

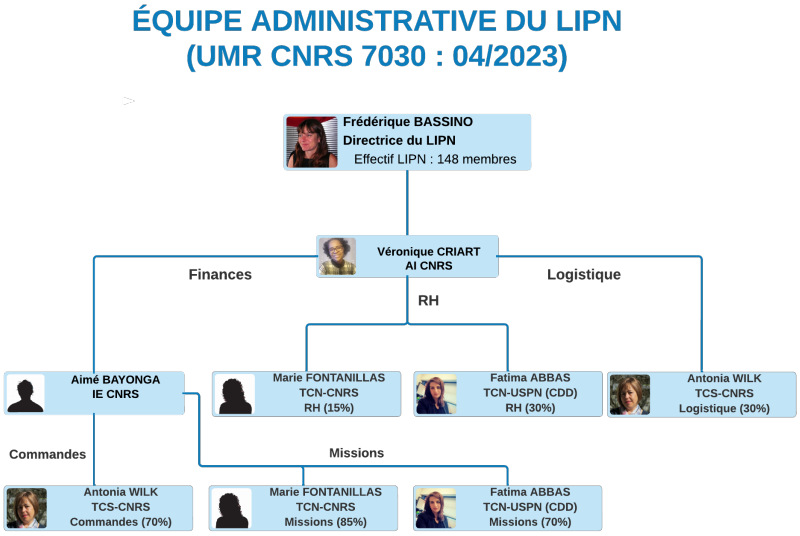
1.2.3 Équipes, plateformes, services communs

Le LIPN compte deux équipes support : l'équipe administrative et l'équipe informatique. Il dispose d'infrastructures informatiques incluant des plateformes et d'une bibliothèque mathématiques-informatique.

• L'équipe administrative

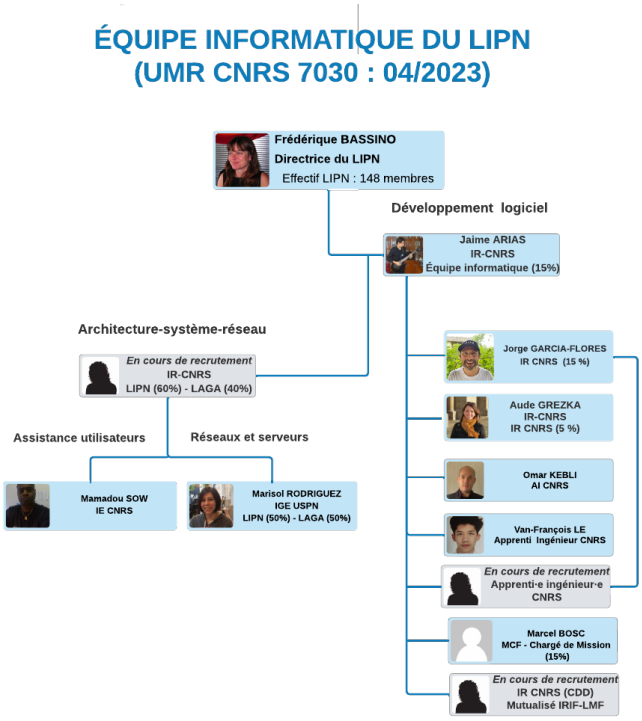
L'équipe administrative a beaucoup évolué depuis la dernière évaluation. Une gestionnaire de l'USPN est partie en mutation en septembre 2019 et une gestionnaire a été recrutée en CDD en 2020 pour la remplacer. Elle prépare des concours pour être titularisée. De plus, l'assistante de direction en charge des questions administratives est partie en retraite en juin 2022. Une responsable administrative de l'unité a été recrutée en octobre 2022 pour piloter l'ensemble du service administratif.

L'équipe est actuellement constituée de quatre personnels CNRS permanents (1 IE, 1 AI, 2 techniciennes qui sont à 80%) et d'une technicienne USPN en CDD. L'organigramme de l'équipe a été clarifié, mettant en évidence les différents pôles d'activités de l'équipe et l'implication de ses membres. Tous les personnels disposent maintenant d'une fiche de poste actualisée précisant l'ensemble de leurs missions. Des ajustements seront faits chaque année pour tenir compte de la montée en compétences des personnels et couvrir aux mieux l'ensemble des besoins de l'unité.



- **L'équipe informatique**

Le LIPN dispose d'une équipe informatique support dont une partie des membres est en charge de l'administration système réseau et l'autre partie du développement logiciel pour le laboratoire.



L'infrastructure. L'infrastructure physique est divisée en trois fonctions différentes : les serveurs de calcul, le système de virtualisation mutualisé et les machines dédiées à différents services. Il est à noter que toute l'infrastructure de virtualisation et de sauvegarde est mutualisée avec le LAGA et partiellement avec le L2TI. La mutualisation du personnel dédié est donc naturelle dans ce contexte.

Des serveurs de calcul à accélérateurs GPU pour entraîner des modèles basés sur de l'apprentissage profond sont accessibles aux membres du LIPN, du L2TI et des laboratoires du LabEx EFL. Il y a actuellement 5 serveurs de calcul GPU, avec un total de 140 cœurs, 1,2 To de RAM et un total de 26 GPUs (8 NVIDIA A40, 11 Quadro RTX 5000, 3 RTX2080 et 4 RTX2080Ti). Cette infrastructure a une cinquantaine d'utilisateur.rice.s. Cela leur permet de réaliser des expériences localement sans trop d'attente, de mettre au point leurs programmes avant de les exécuter sur des infrastructures régionales ou nationales de calcul à haute performance (comme Jean Zay) et de profiter ainsi pleinement des créneaux alloués.

Le système de virtualisation est mutualisé entre le LIPN et le LAGA, et héberge certains services du L2TI, de l'université et des services en collaboration avec d'autres institutions. Il est divisé en 2 parties : la partie la plus récente dans Proxmox dispose de 1,4 To de mémoire vive et de 232 cœurs. La partie VMware a 160 cœurs et 512Gb de RAM. L'équipe RCLN du LIPN dispose également de 2 serveurs de virtualisation (80 cœurs et 640 Go de RAM) pour soutenir différents systèmes tels que CheneTAL, framester, neoveille, morfetik, entre autres.

Il existe aussi des machines dédiées à différents services tels que la visioconférence, la sauvegarde des ordinateurs portables, la sauvegarde des données, les journaux, *etc.*

Le pôle système réseau - Les `admin-sys`. L'équipe est actuellement constituée de deux personnels permanents :

— une administratrice système réseau, IGE USPN titulaire depuis décembre 2020 et mutualisée avec le LAGA,

— un IE CNRS qui assure la gestion du parc informatique du LIPN.

Suite au départ en mutation d'un AI CNRS en 2021, et à des recrutements infructueux, l'administratrice assure seule la gestion de l'infrastructure commune du LAGA et du LIPN. Un poste d'IR mutualisé avec le LAGA est ouvert au recrutement externe. Durant la période, l'équipe a également accueilli un apprenti ingénieur pendant 3 ans.

Les besoins auxquels répond le pôle système réseau du LIPN peuvent être classés en quatre catégories : support au travail collaboratif, services web, besoins de calcul et assistance aux utilisateur·rice·s. Voici une description plus détaillée des services fournis :

1. **Support au travail collaboratif :** l'unité met à disposition de ses membres des services de communication : serveur mail, webmail, messagerie instantanée (Mattermost), de visioconférence (BBB), de stockage de données : Nextcloud, disque réseau partagé et de contrôle de versions : git (GitLab), svn. Ces services souverains sont hébergés sur des machines virtuelles derrière le firewall de securité du LIPN, en vue de faciliter le stockage local des données nécessaires à l'activité scientifique et académique des membres de l'unité. Des services de VPN et de tunnels SSH sont aussi mis à disposition des membres en télétravail. Enfin, l'équipe fournit un service de sauvegarde sécurisée de la totalité de données des postes de travail individuels (Backtop) développé par des membres du pôle système réseau.
2. **Services web** : l'équipe prend en charge la maintenance et la sécurisation des pages web de l'unité, ainsi que la mise à disposition d'un wiki de documentation autogérée, avec une section ouverte en lecture à tous et une partie réservée en lecture-écriture aux membres du laboratoire. La mise en ligne de pages web temporaires pour des conférences ou des projets temporaires, ainsi que des démos de systèmes issus de la recherche du laboratoire est aussi effectuée par l'équipe. Un service de gestion de projet (Openproject) et un service de sondages (Limesurvey) sont mis à la disposition des membres de l'unité.

3. **Services de calcul :** un serveur à 30 accélérateurs GPU est disponible pour la recherche en intelligence artificielle. Il est mutualisé avec le L2TI et le LabEx EFL, son hébergement, sa gestion (notamment de son utilisation concurrente grâce à SLURM) et sa maintenance sont assurée par le LIPN.
4. **Assistance aux utilisateur·rice·s et gestion du parc :** l'équipe prend en charge l'administration et l'exploitation du parc informatique, consistué de serveurs, de postes fixes et d'ordinateurs portables. L'équipe fournit aussi du support aux utilisateur·rice·s dans l'utilisation de ce parc. Un service de prêt d'ordinateurs portables et d'écrans a été mis en place pour faciliter le télétravail pendant la période de confinement, et il a été maintenu après.

Le pôle de développement logiciel - La `dev-team`. La **`dev-team`** est constituée de

— 3 IR CNRS qui, bien que principalement investis dans une activité de soutien à la recherche, consacrent une fraction de leur temps au développement logiciel pour l'unité,
— 1 AI CNRS,
— 1 et bientôt 2 apprentis ingénieurs CNRS,
— 1 MCF du L2TI qui y consacre 15% de son temps,
— 1 IR CNRS CDD mutualisé avec l'IRIF et le LMF est en cours de recrutement,

auxquels s'ajoutent notamment ingénieurs CDD et stagiaires. C'est une structure ouverte dans laquelle les membres du LIPN intéressés par le développement logiciel partagent leurs connaissances, mais aussi leurs problèmes conceptuels et techniques afin de les résoudre ensemble. Hackathon, inclusion d'un développeur malvoyant, mise en place et utilisation d'une plateforme de développement et d'intégration continus, telles sont les principales activités (cf. portfolio 4 du LIPN).

L'équipe développe et maintient le système d'information du LIPN (développement du nouveau site web et refonte de l'intranet de l'unité), ainsi qu'une application pour la direction de la recherche de l'USPN (développement de la plateforme pour la gestion de professeurs invités, utilisée depuis 2019) et contribue à des développements pour la recherche.

L'intranet actuel est un prototype qui n'a pas été maintenu depuis 10 ans et qui présente de nombreuses failles en matière de sécurité et de bonnes pratiques de développement de logiciel. Pour y remédier, la **`dev-team`** en développe une nouvelle version en s'appuyant sur les technologies les plus récentes et une architecture de microservices. Les services de gestion de publications, de réservation de salles, et de gestion de membres ont été mis en production.

La **`dev-team`**, comptant trois IR principalement investis dans une activité de soutien à la recherche, est aussi un espace transverse où apprentis, ingénieurs CDD, stagiaires peuvent découvrir les thèmes de recherche du LIPN. Certains ont ainsi été recrutés pour participer à des projets d'envergure à savoir : Chênetal et Morfetik (RCLN), et CosyVerif (LoVe). Une description détaillée de ces plateformes figure dans les sections respectives de chaque équipe.

Formation des personnels support et personnels d'appui à la recherche. L'unité incite ses personnels à se former et à progresser en terme de polyvalence. Chaque année, les personnels IT CNRS et BIATSS de l'USPN suivent au total près d'une trentaine de formations qui relèvent de l'utilisation d'outils, de l'acquisition de techniques ou de connaissances, ou du management et des RH.

Sur la période, dans l'équipe administrative, une Adjointe Technique a été promue Technicienne et une Technicienne Classe Normale a été promue Technicienne Classe Supérieure.

Les autres promotions concernent des IR CNRS ayant principalement une fonction d'appui à la recherche même s'ils participent à la **`dev-team`**(cf. section 3.2.2).

Accueil d'apprentis ingénieurs CNRS. Le LIPN a accueilli 3 apprentis ingénieurs CNRS en formation à SupGalilée pendant la période et un quatrième est en cours de recrutement pour un début d'apprentissage à la rentrée 2023. Le premier était intégré à l'équipe système réseau, tandis que les deux suivants se sont spécialisés en développement logiciel et ont intégré la **`dev-team`**. Les deux premiers ont terminé leur formation, le troisième finira en 2024.

La bibliothèque recherche mathématiques-informatique. Elle est commune aux membres de la fédération MathSTIC. Elle a actuellement une surface de 350 m^2, 800 mètres linéaires de rayonnages, 45 places de lecture, 5 postes informatiques en libre-service, une imprimante et un scanner. Le fonds de la bibliothèque est constitué d'environ 11 000 ouvrages empruntables par les utilisateurs, 400 ebooks et 150 collections de revues, qui sont consultables sur place ou à distance en version électronique. La bibliothèque sera située au rez-de-chaussée de la deuxième moitié du futur bâtiment MathSTIC.

1.2.4 Effectifs de l'unité et de ses équipes au 31/12/2022

Le tableau ci-dessous indique le nombre de personnels membres du LIPN au 31 décembre 2022 sans tenir compte des temps partiels (*i.e.*, tous les personnels sont pris en compte à part entière).

Équipe	A[3]	AOC	CALIN	LoVe	RCLN	LIPN
DR	0	0	0	2	0	2
CR	0	0	3	3	0	6
PU	2	4	4	5	3	18
MCF	11	7	5	10	10	43
IR	0	0	0	1	2	3
Équipe info.						4
Équipe admin.						5
Post-docs et ATER	3	1	3	2	0	9
Ingénieurs (CDD)	0	0	0	1	0	1
Doctorants	9	9	7	10	6	41
Total	*25*	*22*	*21*	*34*	*21*	*132*

- **Place des femmes dans les effectifs.**

  Le LIPN compte au 31 décembre 2022 :
  — *Parmi les personnels rang A* : 5 femmes PU (soit 27,7% des PU) et 1 femme DR (soit 50% des DR), ce qui fait un total de 6 femmes rang A (soit 33% des personnels rang A). Il s'agit de taux sensiblement au-dessus de la moyenne observée dans les unités de l'INS2I, à savoir 16,7% de femmes parmi les PU et 18,8% parmi les DR.
  — *Parmi les personnels rang B* : 8 femmes MCF (soit 18,5% des MCF) et 1 femme CR (soit $16,6\%$ des CR), ce qui fait un total de 9 femmes rang B (soit 20% des personnels rangs B). Il s'agit de taux cette fois au-dessous de la moyenne observée dans les unités de l'INS2I, à savoir 24,7% de femmes parmi les MCF et 20,5% parmi les CR.
  — L'indice d'avantage masculin est de 0,65 alors qu'il est en moyenne de 1,28 dans les unités de l'INS2I.

  — *Parmi les personnels IT et BIATSS* : on compte autant de femmes que d'hommes. Toutefois il n'y a qu'une seule femme issue de la BAP E et qu'un seul homme issu de la BAP J. Une IR relève de la BAP D, les deux autres IR sont issus de la BAP E.

— *Parmi les doctorant.e.s* : on compte 12 femmes, elles représentent 29% des doctorantes, ce qui est proche de la moyenne des unités de l'INS2I, à savoir 25%.

La dynamique de recrutements de femmes s'est améliorée sur la période. Parmi les 4 PU, recrutés on compte 1 femme ; parmi les 2 MCF recrutés, on compte également 1 femme ; parmi les 2 CR recrutés on compte 1 femme ; de plus les 2 recrutements de personnel support permanents (BAP E et BAP J) sont des femmes. Un effort particulier est fait pour attirer des candidates féminines sur les différents concours (notamment CR et MCF). Sur les 3 recrutements MCF qui sont en cours en 2023, une femme devrait être recrutée.

Il est à noter qu'en termes de responsabilités le LIPN a une directrice, et que le comité d'experts et une des cinq équipes du LIPN sont dirigés par des femmes.

1.2.5 Thématiques scientifiques (par équipe le cas échéant)

Les grandes thématiques scientifiques par équipe et pour l'axe transverse sont les suivantes :

— A³ : Apprentissage (non-supervisé, méta-apprentissage)
— AOC : Optimisation Combinatoire, Programmation mathématique, Calcul à hautes performances
— CALIN : Combinatoire (énumérative, analytique, géométrique et algébrique)
— LoVe : Logique, Sémantique des programmes, Vérification formelle
— RCLN : Traitement Automatique des Langues, Graphes de connaissances
— Axe Complexités : Complexité (algébrique, descriptive, d'approximation, de preuve)

1.3 Les thématiques scientifiques et leurs enjeux

Les recherches menées au LIPN couvrent un spectre scientifique assez large – du TAL à la combinatoire, en passant par l'apprentissage, l'optimisation combinatoire, la vérification et la logique –, elles allient contributions théoriques et applications avec une ouverture interdisciplinaire délibérée.

1.3.1 Stratégie scientifique du LIPN

La stratégie scientifique du LIPN est articulée autour de deux idées fondamentales :

— **Appréhender les grands problèmes scientifiques depuis la recherche fondamentale jusqu'à la valorisation.** La recherche en sciences de l'information est en effet traversée par une double influence : d'une part l'étude de grandes questions fondamentales qui ont des répercussions économiques et sociétales tout en contribuant à l'évolution même de la discipline, et d'autre part la recherche de solutions concrètes, éventuellement logicielles, à des problèmes pratiques. Ainsi P vs NP est l'un des 7 problèmes du millénaire selon le Clay Mathematics Institute ; le développement de l'informatique quantique remet en question la nature du calcul telle que conçue par Turing, et elle est déjà en partie utilisée dans certaines applications. La croissance du monde numérique alimente également la recherche en informatique, de ses applications jusqu'à ses travaux fondamentaux.
— **Se développer en favorisant collaborations et interdisciplinarité avec l'environnement proche ou plus éloigné.** Nombre d'organisations et de sociétés scientifiques insistent sur la richesse et la nécessité des interactions entre l'informatique et les autres disciplines scientifiques, tant du point de vue de l'échange d'outils et de méthodes que dans celui de l'échange de concepts fondamentaux. De multiples avancées scientifiques récentes n'auraient en effet pas vu le jour

sans cette interdisciplinarité et de nombreux domaines en plein essor ont un besoin crucial de ces collaborations entre différentes disciplines scientifiques. Ainsi la théorie géométrique de la complexité, vaste programme lancé par Mulmuley et Sohoni, s'attaque à une reformulation algébrique de P vs NP due à Valiant à l'aide des outils de la géométrie algébrique et de la théorie des représentations. Et l'apprentissage quantique est un domaine de recherche émergent croisant apprentissage et physique quantique. Notre objectif est de contribuer au développement d'échanges fructueux dans lesquels l'informatique, en tant que science à part entière, peut apporter ses méthodologies et ses paradigmes et en retour s'inspirer de ceux issus d'autres disciplines telles que les mathématiques, la physique, les sciences humaines et sociales ou encore la biologie.

1.3.2 Structuration en 5 équipes

Le LIPN soutient les orientations originales de ses équipes en se fixant comme objectifs d'accroître sa visibilité sur la scène internationale ainsi que son attractivité, en étant leader sur des thématiques de pointe. Ces thématiques sont explicitées dans les projets des équipes.

- **L'équipe A³** s'intéresse aux problèmes d'apprentissage automatique et couvre un large spectre de questions, allant de l'apprentissage supervisé et non supervisé à l'apprentissage par renforcement.
  Les résultats marquants qu'elle a obtenus sur la période incluent (cf. portfolios A³ 1 à 5) :
  1. Des avancées en adaptation du domaine en lien avec la théorie du transport optimal ;
  2. Des résultats concernant les données relatives et l'apprentissage de représentation de concepts ;
  3. Un livre sur la théorie de l'adaptation du domaine ;
  4. La création d'une startup HelphIA centrée sur l'exploitation et la visualisation de données massives ;
  5. L'apprentissage de modèles structurés dans le cadre du projet CHIST-ERA Adalab dont le but était de rendre la découverte scientifique plus rapide et reproductible en s'appuyant sur des robots scientifiques.

- **L'équipe AOC** est visible pour ses contribution sur approches non-linéaires et polyédriques pour améliorer la résolution de problèmes d'optimisation de grande taille, que la résolution soit exacte ou approchée ; des questions relatives aux environnements distribués sont également prises en compte.
  Les résultat marquants qu'elle a obtenus sur la période incluent (cf. portfolios AOC 1 à 5) :
  1. Box-total dual integrality, box-integrality, and equimodular matrices, publié dans *Mathematical Programming* ;
  2. An exact algorithm for robust influence maximization, publié dans *Mathematical Programming* ;
  3. Méthodes de Décomposition Appliquées à des problèmes Non Linéaire, publié dans *Mathematical Programming* ;
  4. The Anchor-Robust Project Scheduling Problem, publié dans *Operations Research* ;
  5. A Methodology to Scale Containerized HPC Infrastructures in the Cloud, présenté à la conférence *Euro-Par*.

- **L'équipe CALIN** qui regroupe des chercheurs travaillant dans différents domaines de la combinatoire (analytique, bijective, géométrique et algébrique) est très visible au niveau international sur les plans de l'analyse d'algorithmes et de la physique combinatoire avec de façon transversale un apport important de la géométrie discrète et algorithmique.
  Les résultat marquants qu'elle a obtenus sur la période incluent (cf. portfolios CALIN 1 à 5) :
  1. L'universalité du permuton séparable brownien, l'objet limite de permutations tirées uniformément au hasard dans de nombreuses classes de permutations, résultat publié dans *Journal of the European Mathematical Society, Advances in Mathematics et Annals of Probability* ;
  2. Une interprétation combinatoire des coefficients de Kronecker, résolvant un problème posé en 1938, résultat publié dans *Algebraic Combinatorics* ;
  3. Cartes et hiérarchie BKP, un analogue non-orienté de la hiérarchie KP, résultat publié dans *International Mathematics Research Notice* ;
  4. Échantillonnage aléatoire : aspects probabilistes, combinatoires et géométriques, livre publié par l'American Mathematical Society
  5. Une expression de la complexité de l'optimisation linéaire via les zonotopes, résultat publié dans *Mathematika* et *Proceedings of the American Mathematical Society*.

- **L'équipe LoVe** (ex-LCR) est particulièrement reconnue au niveau international pour son expertise en applications de la logique linéaire aux langages de programmation et en complexité implicite, ainsi que pour son expertise en model checking et spécialement en vérification d'automates temporisés paramétrés.
  Les résultat marquants qu'elle a obtenus sur la période inclus (cf. portfolios LoVe 1 à 5) :
  1. Etude sémantique de la logique TATL (Timed Alternating-Time Temporal Logic), publiée sur le *Journal of Artificial Intelligence Research* ;
  2. Distingushed paper à *POPL 2020* sur l'expansion de Taylor en lambda-calcul ;
  3. Model checking appliqué à la vérification de processus de blockchain (*ICSOC 2021*) ;
  4. Première analyse formelle de la correction de la différentiation automatique dans un langage fonctionnel expressif (*POPL 2021*) ;
  5. Publication d'un livre "mise-à-jour" de la référence fondamentale pour le lambda-calcul (co-signé avec Henk Barendregt).

- **L'équipe RCLN** est très visible pour ses travaux en analyse syntaxique et sémantique qui s'appuient sur de l'apprentissage profond et de l'optimisation combinatoire, ainsi que pour son expertise en construction ou enrichissement de ressources langagières et de graphes de connaissances à partir de corpus ou de données structurées.
  Les résultat marquants qu'elle a obtenus sur la période inclus (cf. portfolios RCLN 1 à 5) :
  1. L'analyse syntaxique en dépendances d'ordre supérieur par descente de gradient et algorithme génétique ;
  2. Le traitement automatique de la langue arabe ;
  3. La génération de graphes de connaissances à partir de la littérature scientifique ;
  4. Une campagne d'évaluation en extraction et classification de relations sémantiques dans les papiers scientifiques ;
  5. Morfetik, une ressource morphologique.

1.3.3 Recherches transverses et recherches inter-équipes

Davantage décloisonner les activités entre les différentes équipes du LIPN était un des objectifs du LIPN sur la période, répondant en cela par avance à une recommandation faite à l'issue de la dernière évaluation HCERES. Les résultats ont dépassé les objectifs envisagés : un axe transverse a été créé, une dynamique autour de la science des données est en train d'émerger et les nombres de co-encadrements et de contrats inter-équipes ont progressé et concernent toutes le équipes.

L'axe transverse Complexités. C'est avec la conviction de l'importance de combiner différentes approches pour progresser sur les grandes questions de complexité que le LIPN a créé en 2022 un axe transverse Complexités (cf. portfolio 1 du LIPN). Cet axe, qui donne de la visibilité à cette thématique de recherche en pleine expansion au sein de l'unité, est destiné à inciter les experts de différentes approches, membres d'équipes différentes, à se réunir et à travailler ensemble. En effet, des expertises de longue date sur des aspects variés de la complexité algorithmique sont déjà présentes notamment dans les équipes LoVe, AOC et CALIN. Des résultats remarquables en lien avec ces questions ont été obtenus dans la période (cf. par exemple les portfolios 2 et 5 de l'équipe CALIN ainsi que [CI-61, CI-80, CI-108] ou [RI-50, CI-26]).

Le LIPN est également très fortement investi pour fédérer et promouvoir la recherche en complexité aux niveaux national (Écoles Caléïdoscope) et international (École et conférence GCT 2022).

La science des données au LIPN. La science des données est présente dans les activités de toutes les équipes de recherche du LIPN (cf. portfolio 2 du LIPN). Si cela est attendu dans les équipes d'apprentissage (A³) et de TAL et représentation des connaissances (RCLN), voire dans l'équipe d'optimisation combinatoire (AOC), c'est nettement plus surprenant en ce qui concerne les équipes de combinatoire (CALIN) et de logique et vérification (LoVe).

On peut noter 3 co-encadrements inter-équipes sur cette thématique pendant la période : entre les équipes A³ et AOC sur du clustering pour l'allocation de ressources, entre les équipes A³ et CALIN sur de l'apprentissage de données tensorielles et entre AOC et RCLN sur l'utilisation de l'apprentissage pour accélérer des solveurs en optimisation. Les deux dernières thèses sont en cours. Des contrats (les projets PIA Square Predict, FUI PCUv2, ANR PROText, CIFRE Alterway, *etc.*) et des publications font également partie des activités inter-équipes sur le sujet.

Autre collaborations inter-équipes. À ces deux grandes thématiques de recherche inter-équipes s'ajoutent des collaborations plus informelles. Ainsi l'équipe LoVe, et en particulier son axe logique, et l'équipe CALIN ont deux projets ANR en commun : un sur la combinatoire des λ-termes (ANR LambdaComb) et l'autre sur les aspects de théorie des représentations liés à la complexité (ANR DYSCO).

La période a également vu une collaboration fructueuse entre l'axe vérification de LoVe et l'équipe AOC (publications cosignées, 2 co-encadrements de thèse), qui n'est cependant plus d'actualité car cela concernait une MCF de l'équipe AOC ayant pris un poste de professeure à l'Université du Québec à Montréal en 2021.

On dénombre au total 29 publications inter-équipes sur la période.

1.3.4 Animation scientifique

En sus des séminaires et groupes de travail organisés par chaque équipe, d'un groupe de travail avec le LAGA et des journées thématiques de la fédération MathSTIC, l'animation scientifique se décline également au niveau du laboratoire.

Elle prenait notamment la forme des *mercredis du* LIPN jusqu'à la pandémie. Ces séminaires destinés à toute l'unité étaient l'occasion d'exposés pour une large audience dispensés par les professeurs invités que le LIPN accueille chaque année.

A contrario, c'est depuis la fin du confinement et des mesures très restrictives que le *séminaire junior* fonctionne avec une très bonne fréquence (toutes les 2 à 3 semaines) avec une importante audience. Il se termine par un moment de convivialité entre doctorants et postdoctorants.

Les deux autres principales animations de la vie scientifiques du LIPN sont les séminaires de l'axe Complexités et les journées autour de la *Science des données*, qui ont vocation à prendre plus d'ampleur dans la période à venir.

On peut également mentionner les journées scientifiques organisées dans le cadre des réunions du Conseil Scientifiques ou de la venue du directeur de l'INS2I.

1.4 Profil d'activités liées à la recherche

Activités (répartir 100 points sur ces 7 items)	
Administration et animation de la recherche : pilotage de la recherche (VP, direction d'institut, DAS, par exemple), participation à des instances d'évaluation (CNU, CoNRS, CSS, Hcéres, par exemple), responsabilité de dispositifs Idex ou Isite, direction de projets (ANR, Horizon Europe, ERC, CPER, PIA, France 2030, par exemple), responsabilités éditoriales dans des revues ou collections nationales et internationales.	25
Aide aux politiques publiques et expertise technique : pouvoirs publics aux niveaux européen, national et régional, entreprises, instances internationales comme FAO, OMS, etc.	2
Contribution à l'adossement d'enseignements innovants à la recherche : EUR, SFRI, etc.	10
Dissémination de la recherche : partage de connaissances avec le grand public, médiation scientifique, interface sciences et société	5
Recherche et encadrement de la recherche.	43
Valorisation, transfert, innovation.	15
Autres activités. (à préciser en une ligne maximum).	0

1.5 Environnement de recherche

- **Fédération MathSTIC**

La fédération MathSTIC contribue à intensifier les collaborations entre le LIPN, le LAGA et le L2TI et favorise l'émergence de recherches tirant parti des compétences présentes dans les trois laboratoires. Ses activités se structurent autour de 4 axes de recherche : i) Optimisation et apprentissage appliqués aux contenus numériques, ii) Calcul haute-performance, systèmes distribués, iii) Physique mathématique, physique statistique, combinatoire et iv) Catégories : entre calcul et topologie. Les membres du LIPN sont impliqués dans chacun d'eux ainsi que dans l'animation et la direction scientifique de la fédération.

Sur la période, 4 thèses ont été co-encadrées avec le LAGA (cela concerne les équipes A³, AOC et CALIN) et deux l'ont été avec le L2TI (cela concerne l'équipe A³). On dénombre également 4 postdocs co-encadrés avec le LAGA (cela concerne les équipes AOC, CALIN et LoVe) et un avec le L2TI (cela concerne l'équipe AOC).

De plus, un LabCom autour du traitement d'images aériennes, co-porté par le LIPN et le L2TI en partenariat avec l'entreprise COSE, vient de démarrer en 2022.

D'une manière générale, les activités inter-laboratoires sont en plein dynamisme et l'emménagement à partir de 2024 dans de nouveaux locaux (cf. section 1.2.1) communs aux trois laboratoires de la fédération devrait renforcer encore ces collaborations.

L'EUR

L'École Universitaire de Recherche de Paris Nord en Mathématiques et Informatique - EUR M&CS est une formation internationale de haut niveau sur 5 ans : master et doctorat. Elle a notamment pour objectif de mettre en valeur les interactions et la complémentarité des mathématiques et de l'informatique afin de favoriser enseignements et recherches interdisciplinaires.

Il s'agit d'un consortium USPN et CNRS qui réunit trois laboratoires (le LAGA, le LIPN et le L2TI) et de la fédération MathSTIC, les masters d'informatique et de mathématiques et l'ED Galilée. Financée dans le cadre du PIA 3, l'EUR bénéficie d'un budget de 4,45 M€ sur la période 2020-28. Cela finance principalement des bourses de master, des contrats doctoraux auxquels s'ajoutent financements de mobilités, de matériel, d'activités avariées, d'intervenants.

Après un démarrage difficile du à la pandémie, l'EUR se développe rapidement depuis 2022 en attirant des candidatures d'une quarantaine de pays. La formation compte actuellement 25 étudiant.e.s en master. Cela constitue un vivier local de doctorant.e.s de qualité ayant si ce n'est une double culture scientifique mathématiques-informatique, au moins un fond de connaissances mathématiques.

LabEx

L'équipe RCLN copilote le LabEx EFL, dans lequel elle est fortement impliquée, et dirige deux de ses cinq axes : « Analyse sémantique computationnelle » et les formations. La pluridisciplinarité est un atout majeur pour l'évolution et le renouvellement des sujets de recherche à la croisée entre linguistique et informatique. L'implication dans ce laboratoire d'excellence est une opportunité pour la réalisation de projets d'envergure et l'ouverture d'un Master interdisciplinaire. Le renouvellement par l'Université Paris Cité de ce LabEx qui se termine fin 2024, et le maintien des conditions de portage des projets (ouvert aux autres universités partenaires hors IdEx) sont des enjeux essentiels pour le LIPN.

Le DIM RFSI

Le Réseau Francilien en Sciences Informatiques (RFSI) a fédéré de 2017 à 2022 l'ensemble des acteurs franciliens en sciences informatiques, en s'appuyant sur un réseau de 27 laboratoires (soit environ 1700 chercheurs) répartis dans toute l'Île-de-France. Ce DIM a financé des projets collaboratifs et co-financé des bourses de thèse et de postdoc.

Le LIPN a non seulement participé et mais également dirigé, sur la période, ce Domaine d'Intérêt Majeur de la région Île-de-France en Sciences Informatiques. Cela lui a permis de renforcer, en tant que structure, son insertion dans l'environnement francilien et de créer un vrai réseau de solidarité avec les autres laboratoires franciliens.

Un partenariat d'envergure s'est également créé autour du besoin en calcul intermédiaire entre les moyens usuels et ceux fournis par le supercalculateur Jean Zay du CNRS à l'IDRIS. Ainsi avec le soutien de la région, le LIP6, LIPN et l'IRIF ont soutenu le projet ConvergenceParis de grilles de calcul CPU et GPU à disposition de leurs membres. L'infrastructure a été installée et est maintenue par l'équipe soutien en informatique du LIP6, une équipe qui a une expérience de ce type de projet d'envergure.

- **Campus Condorcet**

Le Campus Condorcet est un établissement public national associant 10 partenaires fondateurs, dont l'USPN[2], et qui a vocation à devenir un campus de recherche et de formation en sciences humaines et sociales d'envergure européenne et internationale.

Il est situé à la frontière de Paris et d'Aubervilliers en liaison directe par tramway avec le campus de Villetaneuse. Les premiers bâtiments ont été livrés en 2019. Le Campus Condorcet a notamment pour mission de collecter, enrichir, valoriser, mettre à disposition et conserver des ressources documentaires. Il devrait devenir un lieu privilégié pour conduire et développer les recherches collaboratives et pluridisciplinaires associant informatique et sciences humaines et sociales.

Pour le moment, les collaborations du LIPN avec le Campus Condorcet restent encore limitées. Le seul fait marquant est un projet ANR et un co-encadrement de thèse avec le laboratoire Géographie-Cité. Sur les thématiques de recherches en lien avec les SHS, le LIPN a la volonté de profiter de l'opportunité que représente le regroupement de tant de laboratoires de SHS sur le Campus Condorcet.

1.6 Prise en compte des recommandations du précédent rapport

Recommandations concernant les produits et activités de la recherche

— *La visibilité des activités de recherche est à améliorer, en développant, par exemple, la valorisation industrielle et la diffusion de l'information scientifique et technique.*
Sur la période, la LIPN a fait un important effort de valorisation de ses recherches via le développement de logiciel (cf. portfolio 4 du LIPN), la création d'une start-up, le démarrage d'un LabCom, le développement d'une preuve de concept d'un détecteur de malware (cf. portfolio 5 du LIPN), sans compter les mobilités en entreprise de membres de l'unité.
Les actions de diffusion de la culture scientifique et technique sont abordées dans la section 3.4.3. Le LIPN s'est notamment fortement investi dans la promotion de l'informatique dans le secondaire (cf. portfolio 6 du LIPN).

— *Il faut réfléchir au sein du laboratoire à une stratégie de diversification des sources de financement*
Les financements du LIPN se sont largement diversifiés et incluent des financements PIA, ANR, contrats CIFRE, et industriels (cf. section 3.1.2). Les projets ANR représentaient 34,8% du financement annuel moyen sur la période précédente : ils ne représentent plus que 19,1% du financement annuel moyen sur la période. Le nombre de contrats CIFRE a progressé de 70% par rapport à la période précédente et le montant de l'environnement est sensiblement plus élevé. L'évolution la plus notable, en termes de financements, est l'obtention d'important fonds issus des PIA : démarrage d'un LabCom, réussité à un appel AMI Cloud BPI, obtention de financement du LabEx EFL, et de manière plus ponctuelle du LabEx AMIES. Cela représente sur la période 1642 k€ des 5834 k€ (soit 28% du total) de crédits obtenus par l'unité dans des appels à projets.

— *La durée des thèses doit être surveillée et réduite progressivement. La Direction doit saisir l'opportunité d'engager une activité de formation destinée à accroître le flux doctoral en direction de la recherche pour toutes les équipes.*

2. Le Centre National de la Recherche Scientifique (CNRS), l'École des Hautes Études en Sciences Sociales (EHESS), l'École Nationale des Chartes (ENC), l'École Pratique des Hautes Études (EPHE), la Fondation Maison des Sciences de l'Homme (FMSH), l'Institut National d'Études Démographiques (INED) et les Universités Paris 1 Panthéon-Sorbonne, Sorbonne Nouvelle Paris 3, Paris 8 Vincennes - Saint-Denis et l'USPN.

La durée moyenne des thèses a sensiblement diminué sur la période passant de 3,7 ans à 3,3 ans et ce malgré la crise sanitaire (cf. section 3.2.2).
L'École Universitaire de Recherche de Paris Nord en Mathématiques et Informatique - EUR M&CS bénéficie d'un financement PIA à hauteur de 4,45 M€ sur la période 2020-2028. Ses étudiants constituent un vivier local qui devrait permettre d'accroître le flux de doctorants au LIPN (cf. section 1.5).

Recommandations concernant l'organisation et la vie de l'unité

— *La définition d'un organigramme du laboratoire est nécessaire pour situer les différentes instances et structures du laboratoire. Il faut clarifier les rôles et les compétences des différents comités mis en place ou qui vont être mis en place.*
Les diverses commissions ont vocation à élaborer des propositions en amont du conseil de laboratoire sur des tâches spécifiques. Ces dernières font ensuite l'objet de discussions et décisions du conseil de laboratoire. Quand cela est adapté, les commissions assurent ensuite le suivi de la mise en œuvre des décisions et rendent compte à la direction de l'unité et au conseil de laboratoire au besoin.

— *Les activités du service administratif et financier doivent être structurées et identifiées ; le service doit disposer d'un responsable participant au comité de direction. Il en est de même du service informatique qui doit être structuré en tenant compte de l'existence de postes mutualisés et en identifiant les missions transverses. Les responsabilités doivent être explicitement établies.*
Le service administratif a été restructuré depuis la dernière évaluation, un organigramme plus clair a été mis en place, il y a maintenant une responsable et tous les personnels disposent de fiches de poste mises à jour.
Le/la responsable du service informatique (pôle système réseau) est en cours de recrutement, ce pôle est pour le moment sous-dimensionné vu les enjeux. La **dev-team**travaille de manière autonome mais parfaitement concertée avec le pôle système réseau et a son propre responsable. Pour plus de détails on se reportera à la section 1.2.3.

— *La création d'un comité de direction ne doit pas priver le conseil de laboratoire de son rôle d'instance où sont débattues les orientations stratégiques du laboratoire.*
Il n'y a pas eu de mise en place de comité de direction, le rôle du conseil de laboratoire a été maintenu voire accru avec des réunions très régulières (au moins une fois par mois) et aussi fréquentes que nécessaire.

Recommandations concernant les perspectives scientifiques à cinq ans et la faisabilité du projet

— *Des actions nouvelles d'attractivité doivent être imaginées et mises en œuvre.*
Le LIPN s'est fortement impliqué dans l'organisation d'évènements internationaux à forte visibilité (écoles, conférences, *etc.*). Il s'est également lancé dans une démarche volontariste pour attirer des très bon.ne.s candidat.e.s surles postes ouverts au concours, ainsique des candidat.e.s CR CNRS.

— *Les interactions entre équipes et axes sont à inciter par des mesures de fléchage de moyens et par une priorité explicite au niveau du laboratoire.*
Décloisonner les activités de recherche des équipes était un des points de la stratégie du laboratoire et les résultats ont été au delà des ambitions initialement formulées. Un axe transverse Complexités a été créé (cf. portfolio 1 du LIPN), la *Science des données* a émergé comme thématique de recherche partagée par l'ensemble des équipes (cf. portfolio 1 du LIPN) et les recherches inter-équipes ont de manière générale beaucoup progressé (cf. section 1.3.3).

Chapitre 2
Introduction du portfolio

Le LIPN a souhaité dans la partie portfolio de l'unité présenter une sélection représentative de ses activités transverses, qui n'apparaissent pas ou seulement en filigrane dans les portfolios des équipes. Les productions scientifiques *stricto sensu* sont exposées dans les portfolios des équipes thématiques.

1. **La recherche en complexité au LIPN et au delà.** Le LIPN a créé en 2022 un axe transverse « Complexités » pour donner de la visibilité à cette thématique de recherche en pleine expansion au sein de l'unité. En effet, le rôle de membres du LIPN pour fédérer et promouvoir la recherche en complexité tant au niveau local, qu'au niveau national ou international, et le dynamisme scientifique que cela engendre sont à l'image de la recherche que l'unité veut développer, une recherche ambitieuse sur ses objectifs et fédérative à large échelle.
2. **La science des données au LIPN.** La science des données est présente dans les activités de toutes les équipes de recherche du LIPN. Si cela est attendu dans les équipes d'apprentissage (A³) et de TAL et représentation des connaissances (RCLN), voire dans l'équipe d'optimisation combinatoire (AOC), c'est nettement plus surprenant en ce qui concerne les équipes de combinatoire (CALIN) et de logique et vérification (LoVe). Mettre en évidence ce trait d'union qui apparaît en filigrane dans les rapports des équipes et quelques résultats marquants associés donne un éclairage sur ces recherches transverses du LIPN.

3. **Un parti pris d'interdisciplinarité.** Convaincu de l'apport à la recherche scientifique de l'interdisciplinarité tant du point de vue de l'échange de concepts que du transfert de méthodes, le LIPN a une démarche volontariste d'ouverture aux autres sciences (mathématiques, physique, linguistique, philosophie ou encore chimie, médecine). Cela se traduit dans les recrutements, les projets, les encadrements et la production scientifique.
4. **Dynamisation du développement logiciel** : Pour dynamiser et amplifier le développement logiciel dans l'unité, les membres du LIPN impliqués dans le développement de logiciels (5 permanents auxquels s'ajoutent apprentis, ingénieurs CDD et stagiaires) se retrouvent au sein d'un groupe de travail, la **`dev-team`**. C'est une structure ouverte dans laquelle ils partagent leurs connaissances, mais aussi leurs problèmes conceptuels et techniques afin de les résoudre ensemble. Cohésion, inclusion, efficacité et bonne ambiance sont au rendez-vous.
5. **La valorisation au LIPN : Start-up, LabCom et POC** : Au delà de sa participation aux activités des pôles de compétitivité Cap Digital et Systematic, ainsi qu'à celles organisées par la cellule de valorisation du CNRS, le LIPN tient à valoriser ses recherches sous toutes les formes possibles et notamment via le transfert aux entreprises. Création de start-up, soutien aux start-up, développement de partenariat via des LabCom, preuve de concepts font partie des formes de soutien à l'innovation que développe le LIPN.
6. **Promotion de l'enseignement de l'informatique dans le secondaire** : Le LIPN considère comme essentiel que se développe enfin l'enseignement de l'informatique dans le secondaire. C'est la raison pour laquelle l'unité a soutenu les actions de formation des enseignants, s'est engagée dans les jurys de concours de recrutement de l'Éducation Nationale et a participé à la promotion de l'enseignement de l'informatique. Enfin, un chercheur du laboratoire est porteur d'un vaste projet de développement de salles informatiques virtuelles à destination d'un large public.

Chapitre 3
Auto-évaluation du bilan

Sommaire

3.1 Domaine 1. Profil, ressources et organisation de l'unité 30

3.1.1 Référence 1. L'unité s'est assigné des objectifs scientifiques pertinents. 30

3.1.2 Référence 2. L'unité dispose de ressources adaptées à son profil d'activités et à son environnement de recherche et les mobilise. . . 32

3.1.3 Référence 3. Les pratiques de l'unité sont conformes aux règles et aux directives définies par ses tutelles en matière de gestion des ressources humaines, de sécurité, d'environnement, de protocoles éthiques et de protection des données ainsi que du patrimoine scientifique. 34

3.1.4 Synthèse de l'auto-évaluation 36

3.2 Domaine 2. Attractivité 36

3.2.1 Référence 1. L'unité est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche. 36

3.2.2 Référence 2. L'unité est attractive par la qualité de sa politique d'accompagnement des personnels. 39

3.2.3 Référence 3. L'unité est attractive par la reconnaissance de ses succès à des appels à projets compétitifs. 43

3.2.4 Référence 4. L'unité est attractive par la qualité de ses équipements et de ses compétences techniques. 43
3.2.5 Synthèse de l'auto-évaluation 44
3.3 Domaine 3. Production scientifique 45
3.3.1 Référence 1. La production scientifique de l'unité satisfait à des critères de qualité. 45
3.3.2 Référence 2. La production scientifique de l'unité est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels. 46
3.3.3 Référence 3. La production scientifique de l'unité respecte les principes de l'intégrité scientifique, de l'éthique et de la science ouverte. Elle est conforme aux directives applicables dans ce domaine. 49
3.3.4 Synthèse de l'auto-évaluation 49
3.4 Domaine 4. Inscription des activités de recherche dans la société . 49
3.4.1 Référence 1. L'unité se distingue par la qualité et la quantité de ses interactions avec le monde non-académique. 49
3.4.2 Référence 2. L'unité développe des produits à destination du monde culturel, économique et social. 50
3.4.3 Référence 3. L'unité partage ses connaissances avec le grand public et intervient dans des débats de société. 50
3.4.4 Synthèse de l'auto-évaluation 52

3.1 Domaine 1. Profil, ressources et organisation de l'unité

3.1.1 Référence 1. L'unité s'est assigné des objectifs scientifiques pertinents.

Les principaux éléments concernant la stratégie scientifique et les thématiques de recherche du LIPN se situent dans les Sections 1.2.2 et 1.3. Ils sont succinctement rappelés ici. Est ensuite détaillée l'évolution des ressources humaines dont dispose l'unité pour mettre en œuvre ce projet.

Stratégie et objectifs scientifiques. La stratégie scientifique du LIPN est articulée autour de deux idées fondamentales :

— Appréhender les grands problèmes scientifiques depuis la recherche fondamentale jusqu'à la valorisation des résultats obtenus.

— Se développer en favorisant collaborations (IRIF, DIM RFSI dont les deux derniers responsables étaient membres du LIPN) et interdisciplinarité avec l'environnement proche (avec le LAGA et le L2TI au sein de la Fédération CNRS MathSTIC à l'USPN, LabEx EFL au niveau d'ASPC) ou plus éloigné (laboratoires de philosophie à Lyon et à Paris) (voir portfolio 3 du LIPN).

A cela s'ajoutent l'insertion dans le tissu national (implications dans les GDR RO, IM et TAL) et international (GDRI AleanetWork et bientôt l'IRN CLoVe avec le Danemark).

En prenant délibérément le parti de l'interdisciplinarité, le LIPN s'inscrit pleinement dans la stratégie de l'USPN qui a comme objectif de promouvoir l'interdisciplinarité et l'ouverture sur la société. En faisant émerger des activités transverses en science des données, le LIPN s'inscrit dans la politique que l'USPN met en place dans le cadre de l'alliance européenne Uninovis centrée autour de la science des données.

Ressources humaines. Le tableau ci-dessous indique les effectifs du LIPN au 1^{er} janvier de chaque année sans tenir compte des temps partiels (*i.e.*, tous les personnels sont pris en compte à part entière, y compris les apprentis).

Catégorie	2017	2018	2019	2020	2021	2022
DR	1	0	1	1	2	2
CR	6	7	7	7	7	7
PU	18	16	15	15	16	17
MCF	50	49	47	46	45	45
IR	2	3	3	3	2	3
Équipe info.	5	5	5	5	5	6
Équipe admin.	5	5	4	5	5	5
Post-docs et ATER	6	4	9	10	5	2
Ingénieurs (CDD)	1	3	0	0	2	1
Doctorants	35	34	38	35	41	39
Total	*129*	*126*	*129*	*127*	*130*	*127*

Sur la période le nombre d'enseignants-chercheurs membres du LIPN s'est érodé.

L'unité a privilégié sur la période le recrutement de PU (4 sur la période) ce qui suffit à peine à maintenir constant leur nombre. Le manque de rangs A (à peine 25% des (enseignants)-chercheurs) pour animer la recherche, encadrer des doctorants et participer à la gestion de l'unité était déjà signalé comme un risque lors du dernier rapport HCERES, cela reste d'actualité.

Parallèlement, le LIPN a connu des départs de MCF sur la période (promotions, départs en retraite, détachements, disponibilités) et le phénomène s'accentue très fortement en 2022-23. Sur la période, seuls 2 MCF ont été recrutés, ce qui conduit à une chute du nombre de MCF. Cela fragilise fortement certaines thématiques, notamment autour de l'IA. Il est à noter que 3 recrutements sont en cours (mais de nombreux départs s'annoncent également).

Avec la multiplication des opportunités de mobilité (au niveau national et international, dans le public comme dans le secteur privé), les augmentations de charge de travail (administration, enseignement), la question du maintien du potentiel de recherche devient réellement critique.

Implication des personnels. La politique scientifique de l'unité est décidée par le conseil de laboratoire qui se réunit au moins une fois par mois, et autant que nécessaire. Avec la préparation du rapport HCERES, il s'est réuni en moyenne deux fois par mois au premier semestre 2023. Les diverses commissions associant des membres de toutes les équipes travaillent en amont du conseil de laboratoire à des propositions sur les sujets dont elles ont la charge (cf. section 1.2.2). Ainsi une bonne partie des membres du laboratoire contribue directement à l'élaboration de la politique scientifique du LIPN. À cela s'ajoutent les assemblées générales qui permettent d'engager des discussions avec tout le personnel de l'unité sur les sujets importants.

Impact économique et sociétal L'unité a beaucoup progressé en terme de développement logiciel (cf. portfolio 4 du LIPN) et dans ses activités de transfert (cf. portfolio 5 du LIPN). Sur la période, 20% des thèses sont financées par des contrats CIFRE et les relations industrielles de l'unité sont en plein essor.

Le LIPN s'est également fortement impliqué dans la promotion de l'enseignement de l'informatique dans l'enseignement secondaire. La France est en effet extrêmement en retard de ce point de vue et réussir à former des enseignants en informatique et à attirer des étudiant.e.s vers l'informatique est un véritable enjeu de société.

3.1.2 Référence 2. L'unité dispose de ressources adaptées à son profil d'activités et à son environnement de recherche et les mobilise.

Budget. Le budget du laboratoire est explicité dans la Figure 3.1. Celle-ci n'intègre pas les salaires des permanents et contractuels (ATER, BIATSS, IT) payés par l'Université Sorbonne Paris Nord ou par le CNRS sur subvention d'état, ni ceux des thèses CIFRE, ni les dépenses liées à l'infrastructure qui sont couvertes par l'Université (locaux, téléphone, électricité, chauffage). Elle n'intègre pas non plus les recettes et subventions pour l'organisation de colloques.

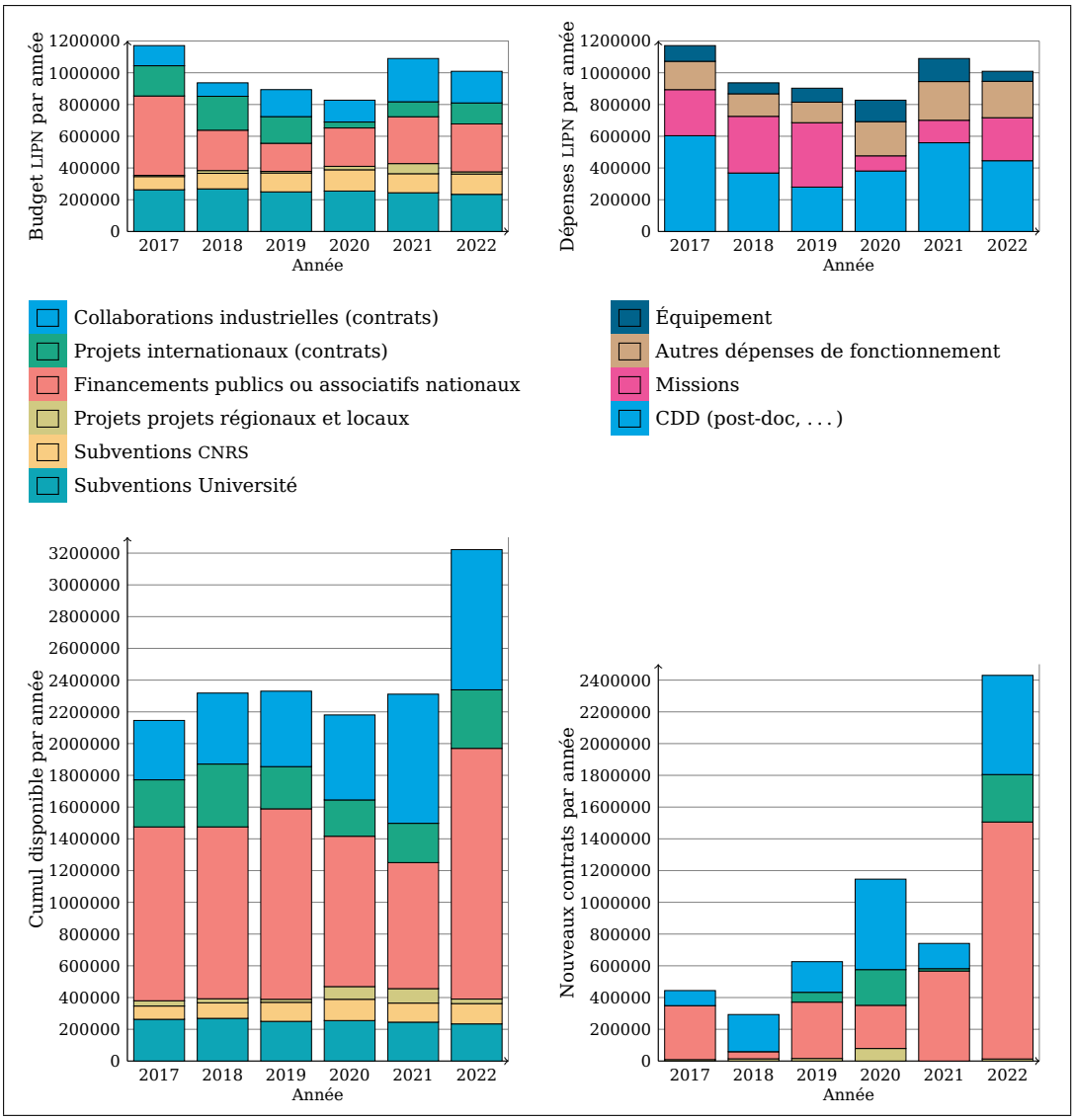


Figure 3.1: Données financières du LIPN sur 2017-2022

Le « budget par année » montre les origines des crédits consommés, alors que les « dépenses par année » indiquent les postes de dépenses. Ces dépenses (un peu moins de 0,6 M€ par an hors salaires) sont stables (si l'on fait exception de l'année 2020). Elles concernent principalement les missions des chercheurs (visites de travail, conférences), l'équipement, le mobilier, les gratifications de stages[1]. Elles intègrent également les

1. La contribution (50% pour les M2) de l'USPN au financement de ces gratifications est comptabilisée.

frais de gestion des contrats.

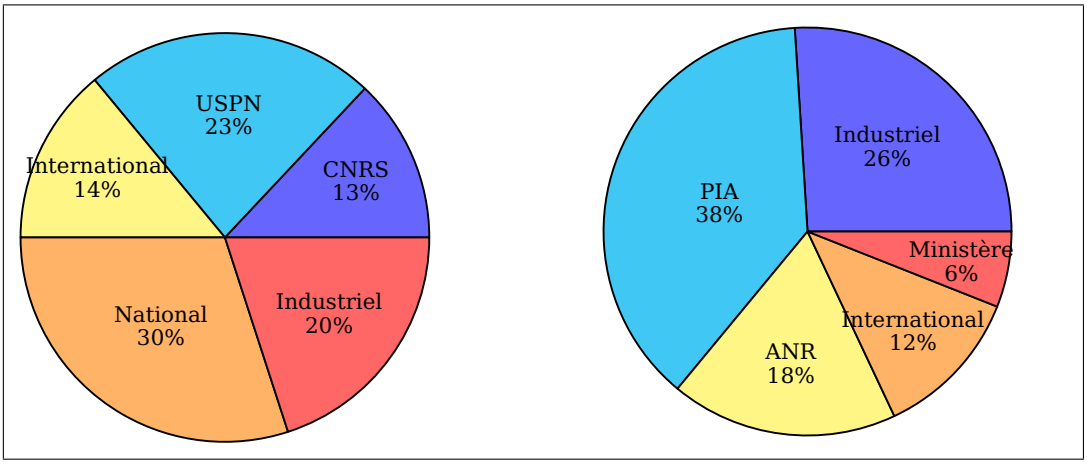


Figure 3.2: La structure du budget 2022 et celle des nouvelles RP 2022

Ressources financières. De nombreux projets étant pluri-annuels, sur 2 à 5 ans, le montant total disponible par année est plus important, comme indiqué dans le « cumul disponible ». Enfin, « nouveaux contrats » fait la synthèse du montant total des contrats obtenus dans l'année [2].

L'année 2022 a été faste, avec le démarrage d'une grande variété de contrats : 2 contrats ANR, 3 financements de thèses (Ministère de l'armée, Université franco-italienne, une entreprise), 5 financements de postdocs (LabEx EFL et AMIES, Cofund - UE, 2 contrats CIFRE (BPCE, Orange), 1 participation au PEPR Cybersécurité, le démarrage du LabCom IRISER, le financement d'un IE via France relance et 3 contrats industriels. Comme le montre la figure 3.2, 38% du montant des crédits obtenus correspondent à des projets financés par des PIA, 26% correspondent à des projets industriels et 18% correspondent à des priojets ANR. Cela marque une profonde évolution de la structure des financements que le LIPN en comparaison de la période précédente où les projets ANR représentaient 34,8% du financement annuel moyen sur la période précédente : ils ne représentant plus que 19,1% du financement annuel moyen sur la période,

A contrario, l'année 2018 a vu peu de contrats démarrer. Il s'agit probablement d'un concours de circonstances du à de nombreux contrats en cours, au basculement du début des contrats ANR à l'année suivante et à la rédaction au cours de l'année 2017 (dépôt en septembre) du précédent rapport HCERES.

Il est donc vraiment indispensable pour le laboratoire de continuer à avoir un financement récurrent conséquent afin de prévenir les fluctuations fortes engendrées par les contrats financièrement importants mais non lissés sur les différentes années.

> Le LIPN a préservé une politique de mutualisation de l'ensemble des ressources à l'échelle du laboratoire (absence de budget d'équipe, mutualisation d'une fraction des ressources sur contrat). Cela facilite les conditions de travail des membres du laboratoire, en leur donnant accès à des équipements de qualité et aux ressources financières nécessaires pour leur poste de travail et pour leurs missions en particulier, et ce quelle que soit la situation des contrats de l'équipe concernée. Cette gestion financière est appréciée des membres du laboratoire qui n'hésitent pas pour autant à monter et soumettre de nouveaux projets.

2. Pour les contrats obtenus en fin d'année, les crédits ont souvent été notifiés à l'exercice suivant.

Matériel informatique et de visioconférence. Dans la période, l'unité a ainsi beaucoup investi en matériel (ordinateurs portables, casques, *etc.*) pour permettre à l'ensemble de son personnel (y compris les non permanents et les stagiaires) de travailler dans les meilleures conditions à distance quand cela a été nécessaire et de télétravailler quand cela est souhaité et possible.

Le LIPN a doté sa principale salle de séminaire d'un équipement performant de visioconférence suite à l'intensification des usages mixtes (participation en présentiel et en distanciel) pendant la pandémie. Un équipement mobile de visioconférence vient compléter les investissements faits pour permettre le travail à distance dans cette période où limiter les déplacements devient une question prégnante. Les équipements ont été choisis de manière à pouvoir être réinstallé dans les futurs locaux du LIPN.

Pour assurer la qualité du code développé au LIPN, en particulier par la **`dev-team`**, le LIPN a investi dans un serveur d'intégration continue (compilation, documentation et tests) et de publication de logiciels.

Pour valoriser les recherches en TAL menées au LIPN, un prototype de la plateforme ChêneTAL destinée à la diffusion des logiciels développés par l'équipe RCLN en direction de la communauté TAL et de celle des chercheurs SHS intéressés par le TAL ou la linguistique de corpus est disponible depuis l'automne 2020. Dans ce contexte, le LIPN a décidé d'acheter un nouveau serveur exclusivement réservé à cette plateforme. Un apprenti ingénieur qui rejoindra le LIPN en septembre 2023 participera à l'intégration des modules à cette plateforme en plein déploiement.

Locaux. Le LIPN manque cruellement de bureaux et d'espaces de réunion. Pour améliorer la situation, des bureaux sont régulièrement réaménagés avec du mobilier plus fonctionnel et plus modulable qui pourra être réutilisé dans les futurs locaux du LIPN. De plus, un MCF, en charge de l'attribution des bureaux, veille à optimiser l'occupation des bureaux de manière à ce que tous les stagiaires de master puissent être accueillis au LIPN. Le déménagement de l'unité en 2024 devrait résoudre ce problème de manque de place.

3.1.3 Référence 3. Les pratiques de l'unité sont conformes aux règles et aux directives définies par ses tutelles en matière de gestion des ressources humaines, de sécurité, d'environnement, de protocoles éthiques et de protection des données ainsi que du patrimoine scientifique.

Politique de ressources humaines. Les (enseignants)-chercheurs sont accompagnés dans leur projet de carrière de diverses manières. L'USPN a mis en place des mesures pour préserver leur temps de recherche. Ainsi les néo MCF bénéficient d'une décharge de 64h de service lors leur première année à l'université et de 32h la seconde année. Les enseignants-chercheurs porteurs locaux ou coordinateurs de projets de recherche suffisamment dotés peuvent bénéficier de 48 à 64h de décharge par an sur la durée de leurs projets.

Par ailleurs, l'unité incite MCF et PU ayant un projet de recherche ambitieux à demander des délégations CNRS, des CRCT, voire à candidater à l'IUF pour bénéficier du temps nécessaire à sa réalisation. Les MCF ayant acquis expérience et maturité en recherche sont incités à soutenir leur HDR. Les dossiers de promotions de corps ou de grade sont relus à la demande des intéressé.e.s. De nombreuses promotions de grade (PU et MCF) ont été obtenues sur la période.

La formation et les carrières des personnels d'appui à la recherche sont abordées dans la section 3.2.2.

Le LIPN qui est régulièrement dirigé par une femme (3 directrices et 3 directeurs depuis sa création) est particulièrement sensible à la question de la place des femmes dans l'unité. Si la proportion de femmes rangs A est bonne, il n'en est pas de même en

ce qui concerne les rangs B (cf. section 1.2.4). Un réel effort est fait pour attirer des candidatures féminines sur tous les types de postes mis au recrutement. Une commission pour la parité femmes-hommmes (cf. sections 1.2.2 et 4.3.8) qui souhaite traiter plus généralement des questions d'inclusion commence également à mettre en place des actions.

Protection du patrimoine scientifique et des systèmes informatiques. La stratégie de protection de l'infrastructure informatique du LIPN et du patrimoine logiciel et scientifique que l'unité héberge se décline en deux parties : mesures préventives et mesures correctives. Pour prévenir les possibles attaques et réduire les brèches de sécurité, on sécurise les accès aux services souverains du LIPN (serveurs de calcul, accélérateurs GPUs, serveurs web, courriel, messagerie instantanée, contrôle de versions logicielles, cloud de stockage, service de visioconférences, gestion de projets, etc.) derrière un Firewall et on surveille les anomalies dans la connexion en direction de ces machines virtuelles (proxmox, vmware) avec Fail2Ban. On veille également à mettre à jour les systèmes d'exploitation de ces serveurs avec les dernières actualisations de sécurité. L'augmentation du personnel en télétravail requiert la mise en place de protocoles de connexion aux services internes au LIPN par des tunnels ssh ou des VPN pour préserver la confidentialité des données échangées.

Le parc informatique est surveillé par des systèmes tels que Zabbix et LibreNMS qui permettent de détecter tout comportement anormal et de le corriger. Cette stratégie préventive est entièrement assurée par l'équipe informatique du LIPN, qui maintient aussi les logiciels maison, développés en respectant les meilleures pratiques en termes de codage pour éviter les brèches par injections SQL ou Javascript, qui sont mis à disposition des collaborateurs externes (démos, systèmes collaboratifs, applications web). Quant aux mesures correctives, dans le scénario d'une éventuelle attaque, des sauvegardes quotidiennes des données et des ressources sont réalisées, et les logs détaillés de l'activité du réseau sont conservés indépendemmant pour pouvoir être réactifs et avoir toutes les données en cas d'intrusion éventuelle. Suite à une intrusion malveillante en 2016 sur l'ancien intranet en PHP, qui a donné lieu à une demande de rançon (non payée), l'équipe informatique a entrepris la refonte d'un nouvel intranet de gestion du laboratoire (publications, thèses, stages, salles de réunions) en Javascript (Mongo, Vue, Nodejs, Expressjs) qui met en place les meilleures pratiques de sécurité pour éviter toute nouvelle intrusion.

Développement durable. Le LIPN s'est récemment doté d'une commission développement durable. Les activités dans ce domaine sont décrites dans les sections 1.2.2 et 4.3.9.

Le LIPN participe à la mise en place de l'opération *Extended Stay Support Scheme* qui donne la possibilité aux conférences d'offrir à leurs participants d'étendre leur séjour dans un laboratoire de recherche du territoire, ou à proximité (par exemple dans le cadre de la conférence *ETAPS 2023*).

L'unité a également organisé *The APSA Challenge for Science and Technology - Ethiopia 2018 : Technologies et innovations pour le développement durable* destiné à promouvoir la recherche et les innovations en science et technologie et l'entrepreneuriat pour le développement durable en Éthiopie et, plus généralement, en Afrique.

Enfin des travaux de recherche en lien avec l'éco-responsabilité ont démarré récemment au LIPN (cf. section 4.3.9).

Plan de continuité d'activité et anticipation des situations d'urgence. Le plan de continuité d'activité du LIPN repose sur la robustesse et la résilience de ses services informatiques, de leur capacité à s'adapter à toute situation comme durant celle de la pandémie. Les **`admin-sys`** du LIPN sont en charge de cette tâche. L'équipement individuel (ordinateur portable, écran, casque) des membres (permanents et non-permanents) du LIPN qu'ils soient au travail ou à domicile fait également partie de ce plan,

Il n'y a pas d'anticipation spécifique de situations d'urgence en dehors de la question des travailleur.se.s isolé.e.s. L'USPN a mis en place pour eux la distribution systématique d'un appareil adapté permettant de signaler un problème.

3.1.4 Synthèse de l'auto-évaluation

- **Forces et Atouts**
  - *Stratégie scientifique en cohérence avec l'environnement* : la stratégie voulue et mise en œuvre par LIPN est pertinente compte tenu de l'environnement et s'inscrit naturellement dans la politique scientifique des tutelles.
  - *Mutualisation du budget* : la mutualisation des ressources financières à l'échelle du laboratoire facilite le fonctionnement de l'unité et est source de cohésion.
  - *CR et IR CNRS* : l'arrivée au LIPN de deux CR et deux IR sur la période a donné un coup de fouet aux activités des équipes concernées (LoVe et RCLN).

- **Faiblesses**
  - *Absence de projet européen* : l'absence de projet européen H2020 - Horizon Europe ou ERC limite les marges de manœuvre que peut dégager sur ressources propres le LIPN.
  - *Un affaiblissement des ressources humaines* : le nombre de PU est stable sur la période mais beaucoup d'entre eux ont ou ont eu de lourdes responsabilités scientifiques ; le nombre de MCF est en baisse ce qui fragilise certaines thématiques de recherche de l'unité.

3.2 Domaine 2. Attractivité

3.2.1 Référence 1. L'unité est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche.

Le LIPN accueille des chercheurs et chercheuses bien intégré.e.s et reconnu.e.s dans leurs communautés scientifiques, comme en témoignent notamment l'organisation et le pilotage de conférences, ainsi que de nombreuses responsabilités dans des instances de pilotage ou évaluation de la recherche.

Dans la suite, sont présentés les éléments marquants qui montrent le rayonnement du LIPN, en termes d'organisation de manifestations scientifiques, de participation à des instances de pilotage ou d'évaluation de la recherche et de prix et distinctions.

Invitations de membres de l'unité. Les membres du LIPN sont fréquemment invités dans des institutions académiques (Universités en Afrique, en Amérique du Nord et du Sud, en Asie et en Europe), ainsi que dans des centres de recherche internationaux (Dagstuhl, Fields Institute, ICERM, Oberwolfach, Banff International Research Station, *etc.*) ou des manifestations internationales (cf. auto-évaluations des équipes). Et si les exposés invités dans des worskhops internationaux ou conférences sur des thématiques ciblées sont nombreux, les exposés invités dans des conférences de référence sont en nombre beaucoup plus limités.

Organisation de manifestations scientifiques. Le LIPN s'investit beaucoup dans l'organisation de manifestations scientifiques internationales, parfois de taille relativement importante. Cet aspect est détaillé dans le portfolio 7 de l'équipe LoVe notamment. Les conférences les plus importantes organisées par des membres du LIPN sur la période sont indiquées ci-dessous.

— *European joint conferences on Theory Ands Practice of Software* (ETAPS 2023) et *29th International SPIN Symposium on Model Checking of Software* (SPIN 2023) ;
— *International Conference on Microservices 2022* ;
— *42nd international conference on Application and Theory of Petri Nets and Concurrency* (PetriNets 2021) ;
— *10th International Joint Conference on Automatic Reasoning* (IJCAR 2020) et *5th International Conference on Formal Structures for Computation and Deduction* (FSCD 2020) ;
— *41st international conference on Application and Theory of Petri Nets and Concurrency* (PetriNets 2020) ;
— *23th International Conference Information Visualisation* (IV 2019) [ED-2] ;
— *30th International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms* (AofA'19) ;
— *IX Latin and American Algorithms, Graphs and Optimization Symposium* (LAGOS 2017) ;
— 5 co-organisations (Publication Chair) du Workshop *Arabic Natural Language Processing (WANLP)* de l'ACL, (en 2017, 2019, 2020, 2021 et 2022) ;
— 2 co-organisations du workshop *Language Technology for Equality, Diversity, Inclusion (LT-EDI* (associé à EACL en 2021 et à ACL en 2022) ;

Responsabilités éditoriales. Le LIPN est impliqué dans des comités éditoriaux de journaux internationaux et mène un travail d'édition d'actes de conférences ou de numéros spéciaux de revues internationales.

Plus précisément, l'unité a à son actif :

— La fondation en 2017 du journal international Annales de l'Institut Henri Poincaré D Combinatorics, Physics and their Interactions, publié par l'EMS (European Mathematical Society) et la participation à son comité éditorial depuis lors (équipe CALIN).
— L'édition en chef du journal international de combinatoire Annals of Combinatorics (équipe CALIN).
— La participation aux comité éditoriaux des journaux internationaux *IEEE Transactions on Computers* (IEEE TC) et *Transactions on Petri Nets and Other Models of Concurrency* (ToPNOC) (équipes AOC et LoVe).
— La participation comme *Associate Editor* aux journaux internationaux Knowledge and Information Systems (KAIS) et Journal of Cloud Computing (équipes A³ et AOC).
— La participation à *l'Advisory Committee* du journal IEEE Cloud Computing (équipe AOC).
— La participation au comité scientifique de la collection *Dixit Grammatica*, une revue internationale électronique de grammaire, linguistique et langue françaises, publiée par L'Harmattan (équipe RCLN).

Le LIPN a également été responsable de l'édition de numéro spéciaux des revues internationales (*Mathematical Structures in Computer Science* [ED-2], *Networks* [ED-5], *Transportations Science* [ED-3], *Discrete Applied Mathematics* [ED-4, ED-2] (équipes LoVe, AOC et CALIN).

Le LIPN est régulièrement impliqué comme chair dans les comités de programme de conférences ou de workshop internationaux :

— PC chair de conférences internationales avec actes (LAGOS 2017 [ED-1, ED-2], FORMATS 2019 [ED-4], NPC 2021 [ED-6], PetriNets 2022 [ED-5]) (équipes CALIN, LoVe et AOC).
— PC Chair de workshop internationaux avec actes (FMICS-AVoCS 2017 [ED-1], DICE/FOPARA 2019 [ED-3]) (équipe LoVe).
— 5 Publication Chair du workshop *Arabic Natural Language Processing* de l'ACL en 2017, 2019, 2020, 2021 et 2022) [ED-1, ED-5, ED-8, **ED-15**] (équipe RCLN).

D'une manière générale, les activités éditoriales sont en nette progression par rapport à la période précédente et mettent en évidence l'amélioration de la reconnaissance internationale de l'unité.

Participation à des instances de pilotage ou évaluation de la recherche. L'équipe est très investie au niveau du pilotage de la recherche, plusieurs membres ayant des responsabilités administratives importantes.

Administration de la recherche :

- présidence de l'Université Sorbonne Paris Nord (depuis décembre 2020) ;
- vice-présidence formation et vie universitaire, USPN (2020–2021) ;
- vice-présidence systèmes d'information, USPN (2016–2020) ;
- vice-présidence numérique, USPN (2016–2020) ;

- direction adjointe scientifique (DAS) de l'INS2I CNRS (2016-2022) ;
- responsabilité de la mission pour l'expertise scientifique du CNRS (2022) ;
- direction de l'institut INS2I (depuis février 2023) ;
- adjointe au Directeur Scientifique Référent (ADSR) du CNRS pour le site Normandie (depuis 2019) ;
- membre de l'instance représentante du CNRS à l'ECSO European Cyber Security Organisation (2017–2022) ;

- direction du DIM RFSI (2018–2022) ;

- direction (adjointe) de l'ED Galilée (depuis 2018) ;
- direction de la fédération MathSTIC (sur toute la période) ;
- direction IREM Paris Nord (sur toute la période) ;
- responsable du cluster MAGI de l'USPN (depuis 2011) ;

Animation de structures de recherche nationales :

- directeur adjoint GDR Recherche Opérationnelle (GDR RO) depuis 2020 ;
- responsable du GT ALEA du GDR IM de 2012 à 2018 ;
- responsable du GDRI AleaNetwork de 2015 à 2018 ;
- responsable du GT « Structures Formelles pour le Calcul et les Preuves » (SCALP) du GDR IM (2018–2023) ;
- responsable de l'axe *Accès à l'information et fouille de textes* du GDR TAL (depuis 2018).

Participation à comités de pilotage de conférences ou workshop :

- 5 SC de conférences internationales (VECOS 2017–2022, FSCD 2019–2023, AofA 2021-présent, PetriNets 2022–présent, ETAPS 2022–présent) ;
- 1 SC de workshop international (Syncop 2017–présent) ;

En outre les membres du LIPN ont participé à une centaine de PC de conférences et workshops internationaux sur la période. Leur participation aux PC des conférences les plus prestigieuses restent toutefois en nombre un peu limité.

Participation à des comités d'expertise ou d'évaluation de la recherche :

- CoNRS - section 06 (membre et secrétaire scientifique, depuis janvier 2021) et CID 53 (membre et membre de bureau, depuis septembre 2021) ;
- conseil Scientifique de l'INS2I du CNRS (membre élu, 2018–2020) ;
- CNU section 27 (3 membres sur la période) ;

- groupe d'Evaluation des Sciences Informatiques pour le programme de subventions à la découverte (NSERC) (l'équivalent Canadien de l'ANR) de 2019 à 2022.
- comité d'évaluation des projets Marie Skłodowska-Curie en 2020 et 2021.
- 2 comités d'évaluation scientifique de ANR (membre du CE48, appels 2021 et 2022 et du CE23 pour les appels 2018, 2020 et 2022) ;

— 9 comités HCERES (3 présidences, 1 vice-présidence, 5 membres) ;
— 1 comité d'évaluation Inria, équipes affiliées au thème *Security and Confidentiality* (membre, 2019) ;

— Jurys IUF junior ou senior (3 membres sur toute la période à raison de 1 par an) ;

— Conseil scientifique PGMO (Programme Gaspard Monge pour l'Optimisation), Fondation Mathématique Jacques Hadamard (membre depuis 2018) ;
— Conseil scientifique du LabEx Milyon (membre depuis 2019) et du LabEx *Les passés dans le présent* (membre depuis 2019) ;

— Expertes pour les campagnes d'avancement des IT CNRS, commission régionale d'interclassement de la BAP D et BAP E (circonscription Île-de-France Villejuif).

Le LIPN est également impliqué dans les jurys de concours de recrutement du CAPES NSI et de l'agrégation d'informatique (cf. portfolio 6 du LIPN).

Prix et distinctions. Des membres du LIPN se sont vus décerner des prix et distinctions sur la période. Les principales sont indiquées ci-après.

— Prix « Paolo Gentilini » pour la logique mathématique et ses applications à l'informatique, attribué par l'Association Italienne de la Logique et ses Applications (AILA), 2020.
— *Distinguished Paper* POPL 2020 (CORE rang A*).
— *Best Paper Award* ICECCS 2018 (CORE rang A).
— *Best Paper Award* FORTE 2017 (CORE rang A).
— Award Recipient of IEEE Technical Committee on Cloud Computing Research Innovation Award en 2020.
— *SageMath Prize 2015-2021*.
— 2 Prix de la recherche médicale 2022 de la fondation de France.

3.2.2 Référence 2. L'unité est attractive par la qualité de sa politique d'accompagnement des personnels.

Le LIPN a su se rendre attractif à la fois sur les postes permanents et non-permanents. Sur la période d'évaluation, il y a eu

— 4 recrutement de PU dont 2 en mutation depuis d'autres postes PU ;
— 2 recrutements de MCF en poste à l'étranger ;
— 2 recrutements de CR CNRS ;
— 2 recrutement d'IR CNRS.

Le LIPN est particulièrement proactif dans sa recherche de candidat(e)s. Pour les postes de CR CNRS, chaque année de nombreux candidat(e)s pressenti (e)s sont contacté(e)s et conseillé(e)s avec une incitation à demander une affectation au LIPN. Pour les candidat(e)s auditionné(e)s, une simulation de jury d'audition est organisée.

Les candidat(e)s sur les postes MCF et PU sont également recherché(e)s. Les conditions de leur accueil sont présentées dans la section 3.1.3.

Formations niveau master. L'USPN propose un master Informatique avec deux spécialités en 2ème année pour environ quatre-vingt étudiants par an. Ces spécialités s'appuient sur des axes forts et des compétences bien développées au LIPN :

— Exploration Informatique des Données et Décisionnel (EID^2) avec un parcours international Science des Données
— Programmation et Logiciels Sûrs (PLS).

Le master de mathématiques propose également un parcours Mathématiques des données.

L'EUR M&CS (cf. section 1.5) qui a débuté en 2020 complète l'offre de formation à l'USPN en finançant de très bons étudiants incités à réfléchir à un parcours interdisciplinaire mathématiques-informatique.

De plus, dans l'optique du recrutement de doctorants, des membres du LIPN interviennent dans divers masters parisiens (MPRI, MPRO, Master PA-Info de l'école Polytechnique, Master LMFI, *etc.*).

Au niveau international, outre le parcours *Science des données* du master, le LIPN participe au master Mathématiques, Informatique et Statistiques Appliquées (MISA) de l'*Université d'Antananarivo* (Madagascar). Sur la période, avant la pandémie, trois étudiants ont effectué leur stage au LIPN et 2 y ont fait leur thèse. Le LIPN participe, depuis sa création en 2011, au master STIC à l'USTH, Hanoï, Vietnam. Plusieurs stagiaires de master 2 ont été accueillis au LIPN.

Accueil de stagiaires. Le LIPN accueille chaque année une trentaine de stagiaires dont une importante proportion font leur stage de fin de master. Malgré le manque de bureaux, ils sont accueillis dans les locaux du LIPN, si besoin un ordinateur portable leur est fourni et ils participent aux activités de l'unité.

Formation à la recherche. La formation à la recherche de type doctoral s'effectue au sein de l'École Doctorale Galilée (ED 146) dont la direction est assurée par un membre du LIPN. Cette école doctorale décline son activité en quatre groupes : (i) mathématique, informatique, signal, (ii) physique, matériaux, sciences pour l'ingénieur, (iii) santé, médecine, biologie humaine et chimie, (iv) éthologie. Les laboratoires de la fédération MathSTIC (LIPN, LAGA et L2TI) constituent le premier groupe. Le conseil de l'école doctorale détermine la politique de la formation doctorale. Le LIPN y a un représentant. Dans une volonté politique de soutien à la recherche, l'USPN finance 25 contrats doctoraux par an qui sont répartis par l'école doctorale. Ainsi, le LIPN bénéficie d'environ 5 contrats doctoraux par an. Les doctorants du LIPN sont inscrits à cette école doctorale l'exception d'une doctorante en chimie et d'une doctorante en double compétence (Linguistique-Informatique).

Le comité des thèses du LIPN organise le suivi des thèses et l'animation des activités doctorales au sein du laboratoire (cf. section 1.2.2).

Sur les 105 thèses soutenues ou en cours sur la période 2017–2022, 55 sont ou ont été financées par des CD USPN (dont 7 CDSN et 1 CD USPC), 18 par contrats CIFRE auxquels on peut ajouter 3 financements industriels, 17 ont été financées sur projets émanant d'agences françaises de financement public (ANR, LabEx, FUI), 9 par des bourses de gouvernements étrangers (Italie, Canada, Iran, Tunisie, Maroc, Vietnam) et 2 par des crédits de l'UE (Erasmus+).

Si l'on compare à la période précédente (2012-2017), le nombre de thèses est en augmentation (105 vs 96), le nombre de financements sur CD USPN (55 vs 63) est en baisse probablement du fait de la réduction de la durée des thèses, le nombre de contrats CIFRE est en nette progression (18 vs 11 soit de l'ordre de 70% de progression). Le montant de l'environnement des thèses en contrat CIFRE est également en nette progression. Le LIPN a acté en 2022 qu'il devait s'élever à 10k€ par an au minimum et préférablement à au moins 15k€ par an.

La durée moyenne des thèses soutenues sur la période est de 3,3 ans en nette diminution comparée au 3,7 ans de la période précédente.

Les docteurs issus du LIPN, pour ceux qui ne sont plus en postdoctorat, sont pour 1/3 MCF et pour 2/3 ingénieurs R&D majoritairement spécialisés en sciences des données, mais aussi en développement logiciel ou en optimisation. Le secteur est actuellement extrêmement attractif en termes de salaires et d'opportunité de carrière rapide.

Depuis l'allègement des restrictions sanitaires, le *séminaire junior* se tient à une fréquence assez élevée (toutes les 2 à 3 semaines) avec un forte participation.

Le LIPN a également participé à des jurys concernant les doctorants comme
— Jury pour les prix des thèses de la chancellerie de Paris en sciences en 2020 ;
— Jury pour le prix de thèse national PGMO en 2019.

Le LIPN s'investit dans l'organisation d'écoles en direction notamment des jeunes chercheurs comme
— L'*École de Printemps d'Informatique Théorique* (EPIT 2023) : Le Kaléidoscope de la Complexité. ;
— *International school and conference on Geometric Complexity Theory* (septembre 2021–janvier 2022) ;
— *Summer School "Complexity as a Kaleidscope"* (*Caleidoscope 2019*) ;
— *Column Generation School 2018* ;
— MathExp 2018 : "Mathématiques expérimentales: méthodes et pratiques".

HDR. Sur la période 2017–2022, 11 habilitations à diriger des recherches ont été soutenues. Sur ces 11 habilités, 2 CR CNRS sont devenus DR, 1 CR Inria est devenu DR et 4 MCF sont devenus PU (dont 2 dans la campagne de recrutement 2023).

Accompagnement des personnels d'appui à la recherche. Le LIPN compte 3 IR qui sont sous la responsabilité de leurs responsables d'équipe (qui procèdent notamment aux entretiens annuels). Bien intégrés dans leurs équipes de recherche, ils consacrent de l'ordre de 15% de leur temps à des tâches pour l'ensemble du laboratoire. Cela permet une certaine cohésion avec les personnels d'appui à la recherche, cela leur permet aussi de rayonner sur l'ensemble de l'unité et pas seulement sur leurs équipes.

Comme indiqué dans la Section 1.2.3 qui mentionne la question de la formation et des carrières des personnels de soutien à la recherche, les 3 personnels d'appui à la recherche (2 IR BAP E et 1 IR BAP D) ont suivi de nombreuses (plusieurs par an) formations au cours de la période. Il s'agissait principalement de formations pour acquérir des techniques ou connaissances de pointe et de formation de management de projet. Sur la période, l'une des IR a été successivement promue IR 1, puis IR hors-classe, tandis qu'un autre a été promu IR 1. Le troisième IR n'était pas promouvable.

Accueil de professeurs invités. La LIPN accueille de l'ordre d'un quinzaine de professeurs invités pour une durée de 2 à 4 semaines chaque année, à l'exception du coup d'arrêt de la pandémie en 2020 et d'une lente reprise de 2021. Cet accueil est rendu possible par les financements sur appels à projets de l'USPN, de la fédération MathSTIC et tout récemment de l'EUR.

Science ouverte. Le LIPN est impliqué dans la science ouverte, qui inclut non seulement l'accessibilité des publications, mais aussi l'accessibilité et l'ouverture des données de la recherche, *i.e.* la publicisation des données des articles scientifiques, permettant la reproductibilité et ainsi une plus grande confiance dans les résultats de la recherche. L'unité met en œuvre les préconisations de l'USPN et du CNRS, mais son investissement va également au-delà.

L'USPN a signé les accords de DORA (2021), a établi une charte de la science ouverte (2021) et depuis un groupe de travail Science ouverte auquel participent deux memebres du LIPN préconise des actions en lien avec la Commission de la Recherche de l'USPN. Le Coalition for Advancing Research Assessment - COARA a été signé en 2022. En 2021, une campagne de sensibilisation encourageant l'utilisation du numéro ORCID a été menée et des rappels sont régulièrement effectués auprès des enseignants-chercheurs et des doctorants.

Les publications du LIPN sont actuellement disponibles sur HAL ou arxiv. Certaines équipes ont également constitué des corpus et des jeux de données qui sont accessibles à tous et sont utilisés par la communauté scientifique.

Un membre du LIPN est aussi membre depuis 2021 de « l'EOSC (European Open Science Cloud) Future User Group »« qui prépare le lancement de la plateforme visant (parmi de nombreux autres objectifs) à ouvrir les données de la recherche des universitaires européen(ne)s.

Il est aussi membre depuis 2021 du Conseil scientifique DoRANum disciplinaire, qui vise à accompagner les universitaires français(es) dans l'ouverture de leurs données de recherche. Cette participation implique notamment la conception de matériel pédagogique spécifiquement dédié aux universitaires en informatique. Son matériel pédagogique est disponible sous la forme du cours en ligne « Ouvrir les données de recherche en informatique théorique : qu'a-t-on à y gagner ? ».

Éthique scientifique et intégrité scientifique. Un référent intégrité scientifique est en poste depuis 2017 à l'USPN. Il occupe également les fonctions de référent déontologie.

Depuis 2019, l'usage de COMPILATIO (dispositif anti-plagiat) a été rendu obligatoire pour l'analyse des manuscrits de thèse avant autorisation de soutenance. Les doctorant.e.s sont formé.e.s à l'usage de COMPILATIO et de la documentation est disponible. Chaque école doctorale peut demander l'analyse d'autres documents, en particulier le projet de thèse.

L'USPN a bénéficié, dans la suite de la COMUE USPC, du Comité d'Ethique de la Recherche de l'Université de Paris Cité. En 2022, l'USPN a créé son propre CER. Un réseau de correspondants Éthique et Intégrité Scientifique auquel participe un membre du LIPN a également été mis en place en 2022. Des groupes de travail préparent depuis une recension des pratiques et des recommandations sur le cahier de laboratoire et le stockage des données. Le réseau est en lien avec l'Office Français de l'intégrité scientifique et les référents intégrité des organismes partenaires. Deux séminaires de formation en résidence pour les membres du CER et le réseau Ethique et Intégrité Scientifique ont été organisés. La VP recherche et le référent intégrité scientifique organisent un cycle de trois conférences par an (validable en ECTS et ouvert à tous). L'éthique et l'intégrité scientifique font partie de la formation obligatoire des doctorant.e.s.

Dans ce contexte, un des enjeux importants pour le LIPN est la *reproductibilité* des expériences et des exécutions logicielles. L'unité travaille à rendre accessible l'ensemble de sa production (publication, production logicielle et de données). Les publications du LIPN sont disponibles via Arxiv ou HAL et la plupart des outils développés sont publiés sous licence libre.

La plateforme d'intégration continue dont dispose le LIPN permet de tester la reproductibilité des résultats et de montrer comment le code et les données génèrent les résultats publiés. De plus, l'outil GUIX HPC, qui permet de construire des environnements reproductibles sur GNU/Linux, est installé sur les serveurs GPUs et des tests d'utilisation sont en cours. L'unité participe également au projet Software Heritage d'archivage des logiciels. Plusieurs logiciels sont déjà sur l'archive comme par exemple, IMITATOR, CosyVerif, ADT2AMAS. GUIX étant couplé avec Software Héritage, il est possible de construire des environnements reproductibles avec les logiciels qui se trouvent dans l'archive.

3.2.3 Référence 3. L'unité est attractive par la reconnaissance de ses succès à des appels à projets compétitifs.

Le LIPN encourage ses membres à répondre aux appels à projets pertinents relativement à leur situation (début de carrière, existence d'une thématique émergente, projet ambitieux et mûr, *etc.*).

Les *nouveaux recrutés* répondent de manière presque systématique aux appels à projets internes des tutelles (projets d'établissement de l'USPN (ex-BQR), projets de l'INS2I), avec un excellent taux de réussite. L'étape suivante est la réponse aux appels JCJC de l'ANR, sur lesquels le LIPN a un assez bon taux de réussite (2 contrats sur la période et une demande en liste complémentaire en 2022).

Le LIPN répond également à des *projets régionaux* : DIM RFSI, appel Cofund postdoctorant régional, financement PGMO de la fondation Jacques Hadamard.

Les membres du LIPN plus expérimentés et avec des propositions suffisamment solides répondent aux appels multipartenaires de l'ANR comme porteurs de projets : 8 contrats ont démarré sur la période dont 2 coordonnés par le LIPN. Des financements FUI ont également été obtenus dans la période.

> L'évolution la plus notable, en termes de financements, est l'obtention de fonds issus des PIA : démarrage d'un LabCom, réussite à un appel AMI Cloud BPI, obtention de financement du LabEx EFL, et de manière plus ponctuelle du LabEx AMIES. Cela représente sur la période 1642 k€ des 5834 k€ (soit 28% du total) de crédits obtenus par l'unité dans des appels à projets.

Concernant les projets relevant de financements de l'UE, sur toute la période, le LIPN a porté des projets Erasmus+ K107 avec la Tunisie, le Maroc et le Vietnam qui ont permis des mobilités entrantes d'étudiants de master et de doctorants et des mobilités sortantes d'enseignants-chercheurs. Le LIPN a bénéficié de financements du projet H2020 Philhumanis mais sans en être partenaire et d'un financement de postdoc Cofund via un appel régionial. Un MCF a passé l'oral de l'ERC Starting Grant en 2017 et le LIPN compte un membre extérieur dans le projet ERC Advanced Grant Geoscape : *Advanced Grant From Geometry to Combinatorics and Back : Escaping the Curse of Dimensionality*. Enfin aucun financement Marie Skłodowska-Curie n'a été obtenu.

Au niveau international, hors projets financés par l'UE, sur la période d'évaluation le LIPN a obtenu le financement de projets de coopération ECOS Nord, PHC, IEA CNRS, INS2I, dont un a abouti à une demande de création d'IRN CNRS, soumise en 2022 et approuvée en 2023. Un projet ANR avec Taiwan (2015-2020) s'est également déroulé sur la période.

Les projets industriels ne sont pas abordés dans cette section mais sont discutés en section 3.4.1. Le mode de financement des doctorants est précisé à la section 3.2.2.

3.2.4 Référence 4. L'unité est attractive par la qualité de ses équipements et de ses compétences techniques.

Le développement logiciel fait l'objet du portfolio 4 du LIPN. La commission développement logiciel est, elle, décrite dans la section 1.2.2, les infrastructures et la **dev-team** dans la section 1.2.3. Enfin les enjeux d'avenir autour du développement logiciel sont évoqués dans la section 4.3.5.

Dans ce qui suit, les principaux éléments sont succinctement rappelés et les effets de levier liés aux plateformes sont indiqués.

Les activités de développement bénéficient du fort dynamisme de la **dev-team** dont la composition et l'activité est présentée dans la section 1.2.3. La commission développement logiciel (cf. section section 1.2.2) se charge, elle, d'organiser des formations en lien avec le développement logiciel et de diffuser les bonnes pratiques. Cette activité s'intensifie actuellement (cf. section 4.3.5).

Pour améliorer la qualité et la robustesse des outils développés au LIPN, les membres de la **dev-team**, en étroite collaboration avec l'équipe système réseau, ont mis en place un serveur d'intégration continue (cf. portfolio 4 du LIPN).

De plus, afin de rendre visibles les logiciels développés au LIPN et de simplifier leur utilisation par le public, chaque équipe dispose d'un serveur propre puissant dédié à leur « publication ». Ainsi

— l'équipe AOC met à disposition son solveur de problèmes quadratiques binaires BiqCrunch ;
— l'équipe CALIN utilise son serveur pour différentes tâches liées au logiciel Sage, à savoir : la compilation de binaires, l'hébergement de son forum de support et de son wiki, et le mirroring de *sage-on-gentoo* ;
— l'équipe LoVe met à disposition sa plateforme de spécification et vérification de systèmes complexes CosyVerif et l'interface web du logiciel d'analyse de systèmes paramétrés IMITATOR ;
— L'équipe RCLN déploie ses plateformes TAL, par exemple, neoveille et morfetik.

Les équipes LoVe et RCLN comptent 3 IR très investis dans le développement de leurs plateformes et ressources morphologiques. Un apprenti ingénieur CNRS est également impliqué dans le développement de CosyVerif.

Les plateformes neoveille et morfetik ont bénéficié de financements du ministère de la culture et de crédit de l'IdEx USPC pour leur développement, actuellement des crédits du LabEx EFL ont pris le relais.

Le LIPN s'est également doté d'une infrastructure (cf. section 1.2.3) constituée de serveurs de calcul à accélérateurs GPU pour entraîner des modèles basés sur de l'apprentissage profond. Cette infrastructure est accessible aux membres du LIPN, du L2TI et des laboratoires du LabEx EFL (une cinquantaine d'utilisateur.rice.s actuellement). Elle a été financée par un projet du LabEx EFL et par des crédits du LIPN et du L2TI. Les publications de l'équipe RCLN mêlant TAL et apprentissage profond sont toutes issues d'entraînements menés au moins en partie sur cette infrastructure.

Le logiciel CosyVerif permet de rendre accessibles les outils et artefacts développés aux utilisateurs et même aux reviewers via un navigateur web. C'est un outil utilisé dans de très nombreuses publications de l'axe Vérification de l'équipe LoVe. Par exemple, le logiciel ADT2AMAS étant intégré à CosyVerif, tous les benchmarks mentionnés dans les publications l'utilisant peuvent être lancés sans rien à installer comme le montre cette vidéo associée à [**CI-99**].

L'interface web imitator permet aussi aux utilisateurs de tester les différentes versions du logiciel imitator sans rien à installer. Cette interface est régulièrement utilisée pour publier des artefacts imitator comme celle de [**CI-100**].

3.2.5 Synthèse de l'auto-évaluation

- **Forces et Atouts**

— *Un investissement significatif dans l'administration de la recherche* : des membres du LIPN ont pris de très importantes responsabilités dans l'administration de la recherche tant au niveau local qu'au niveau national.
— *Une activité éditoriale de bon niveau* : toutes le équipes sont impliquées dans des comités éditoriaux de revues internationales et dans des activités d'édition en lien notamment avec des conférences internationales.
— *Un implication forte dans l'organisation de manifestations scientifiques internationales* : toutes le équipes participent à l'organisation de conférences et workshops internationaux, souvent de taille conséquente.
— *Une vraie dynamique autour du développement logiciel* portée par la **dev-team** et des investissements du LIPN dans des infrastructures adéquates.
— *Une sensible diminution de la durée moyenne des thèses* qui passe de 3,7 ans sur la période 2012-2017 à 3,3 ans (malgré la pandémie) sur la période évaluée.

- **Faiblesses**
  - — *Peu d'invitations dans les conférences vraiment prestigieuses.*
  - — *Faible participation aux PC de conférences de référence.*
  - — *Pas de projet européen H2020 - Horizon Europe, ERC ou Marie Skłodowska-Curie.*

3.3 Domaine 3. Production scientifique

3.3.1 Référence 1. La production scientifique de l'unité satisfait à des critères de qualité.

La production scientifique du LIPN repose sur des fondements théoriques et méthodologiques solides comme en témoignent notamment la qualité des média via lesquels les résultats obtenus sont diffusés. La liste complète des publications est donnée par équipe. Les informations mentionnées en Figure 3.3 quantifient cette production.

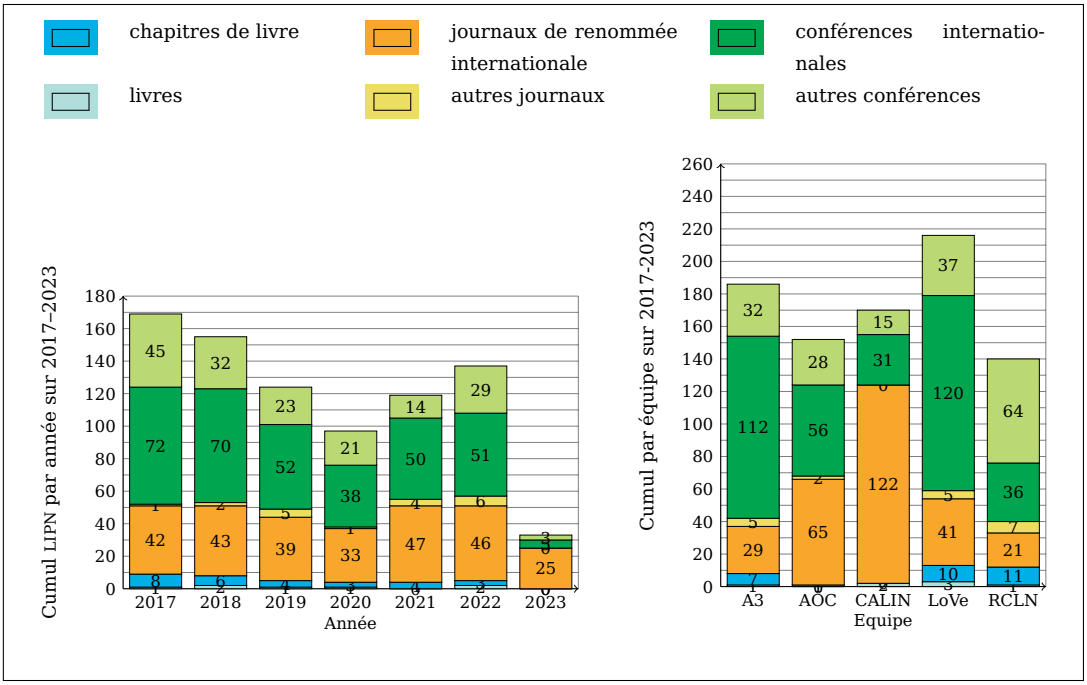


Figure 3.3: Publications scientifiques publiées ou acceptées entre 2017 et 2022

La production scientifique se traduit essentiellement en articles dans des journaux internationaux (275) et en communications dans des conférences internationales (338) avec actes et comités de lecture, ainsi qu'en 7 livres et 28 chapitres de livre. De plus, 64 logiciels, allant du prototype à de simples développements, ont été produits avec une méthodologie permettant la reproductibilité (cf. section 3.2.2).

Sur la période évaluée, on relève notamment la production de 3 livres remarquables :

- — un livre sur les avancées récentes en théorie de l'adaptation du domaine, *Advances in Domain Adaptation Theory* [LI-1] (cf. portfolio 3 de l'équipe A³) ;
- — un livre sur l'échantillonnage aléatoire : aspects probabilistes, combinatoires et géométriques, *Sampling in Combinatorial and Geometric Set Systems* [LI-2], publié par l'American Mathematical Society (cf. portfolio 4 de l'équipe CALIN) ;
- — un livre, *A Lambda Calculus Satellite* [LI-3] co-signé avec Henk Barendregt, « mise-à-jour » de la référence fondamentale pour le lambda-calcul (cf. portfolio 4 de l'équipe LoVe) ;

qui ont vocation à devenir des références dans leurs domaines respectifs.

Les supports de publication sont variés et de qualité.

En termes de revues internationales, on peut noter *Advances in Mathematics, Journal of the European Mathematical Society, Israel Journal of Mathematics Physical Review D, The Proceedings of the National Academy of Sciences (PNAS), Mathematical Programming, Mathematical Programming Computation, Transport Science, Operations Research, Pattern Recognition, Artificial Intelligence, Journal of Parallel and Distributed Computing* ou encore *ACM Transactions on Computational Logic*.

En termes de conférences, on retrouve parmi les plus reconnues : *ACM-SIGACT Symposium on Principles of Programming Languages (POPL)* avec 6 contributions sur la période ; *Computer Aided Verification (CAV)* avec 4 contributions sur la période ; *International Joint Conference on Artificial Intelligence (IJCAI), International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS), National Conference of the American Association for Artificial Intelligence (AAAI)* et *International World Wide Web Conference (WWW)* avec 3 contributions sur la période ; mais aussi *IEEE Symposium on Logic in Computer Science (LICS), International Conference on Machine Learning (ICML)* ou encore *International Semantic Web Conference (ISWC)*.

Les résultats marquants de chaque équipe sont détaillés dans leurs portfolios. Les contributions en complexité et en science des données sont évoquées succinctement dans les portfolios 1 et 2 du LIPN.

Toutes ces contributions ont été obtenues grâce à un important réseau de collaborations académiques et industrielles au niveau national et sur les six continents.

D'une manière générale, la qualité de la production s'est sensiblement améliorée sur la période comme en témoignent les média plus sélectifs via lesquels elle est diffusée et la longueur parfois conséquente de ces contributions.

3.3.2 Référence 2. La production scientifique de l'unité est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels.

La production scientifique est à estimer quantitativement relativement au temps consacré à la recherche par le personnel permanent de l'unité. Le tableau suivant indique les équivalents temps plein recherche (ETP) permanents, en tenant compte des personnels exerçant à temps partiel, par équipe et par an.

Année	2017	2018	2019	2020	2021	2022	Total
A^3	8.33	8	7.71	7.83	7.33	6.83	46.03
AOC	6.33	6.33	6.83	6.5	6.28	5.68	37.95
CALIN	8.5	8.33	8	8.33	8.83	8.66	50.65
LoVE	10.5	11.25	11	11.08	11.6	11.71	67.14
RCLN	6.5	5.83	4.79	5.96	6.17	6.5	35.75
LIPN	40.16	39.74	38.33	39.70	40.21	39.38	237.52

Le graphique 3.4 précise, pour chaque équipe, le nombre de publications par année selon leur type ainsi que l'évolution du nombre de membres permanents comptés en équivalent temps plein permanent (1 enseignant-chercheur comptabilisé comme 0,5 chercheur temps plein). Les modes de publication (revue *versus* conférence par exemple) dépendent du domaine d'activité de l'équipe.

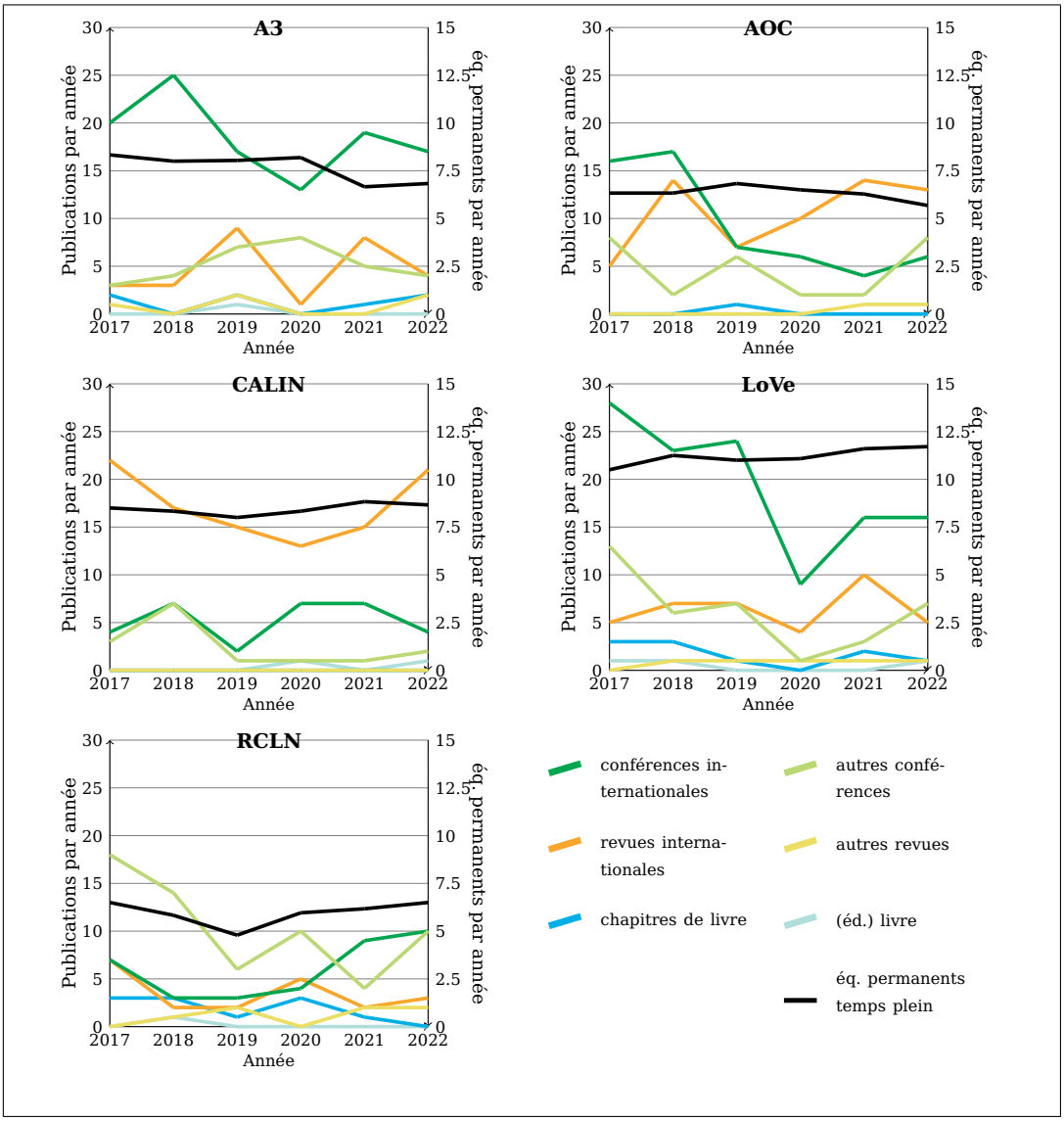


Figure 3.4: Principales publications par équipe

En termes quantitatifs, la production scientifique du LIPN correspond à :

— en ne tenant pass compte des multi-auteurs du LIPN pour une même publication
  — 1,15 revue internationale par ETP et par an
  — 1,42 conférence internationale par ETP et par an
— en tenant compte des multi-auteurs du LIPN pour une même publication
  — 1,55 revue internationale par ETP et par an
  — 2,72 conférences internationales par ETP et par an

Cette production semble raisonnablement proportionnée au potentiel de recherche de l'unité. Elle est correctement répartie entre les personnels s'il est tenu compte des lourdes responsabilités administratives assumées par certains. La production est quantitativement stable par rapport à la période précédente.

Il est à noter que, sur l'ensemble des auteurs des publications du LIPN (en tenant compte des multi-auteurs), 400 sont doctorants (pour 105 doctorants sur la période) et 49 sont postdoctorants.

Selon les domaines de recherche des équipes, les cultures du publications diffèrent fortement. Ainsi en combinatoire et en optimisation combinatoire, les publications en revue jouent un rôle primordial, tandis qu'en logique et programmation, en vérification

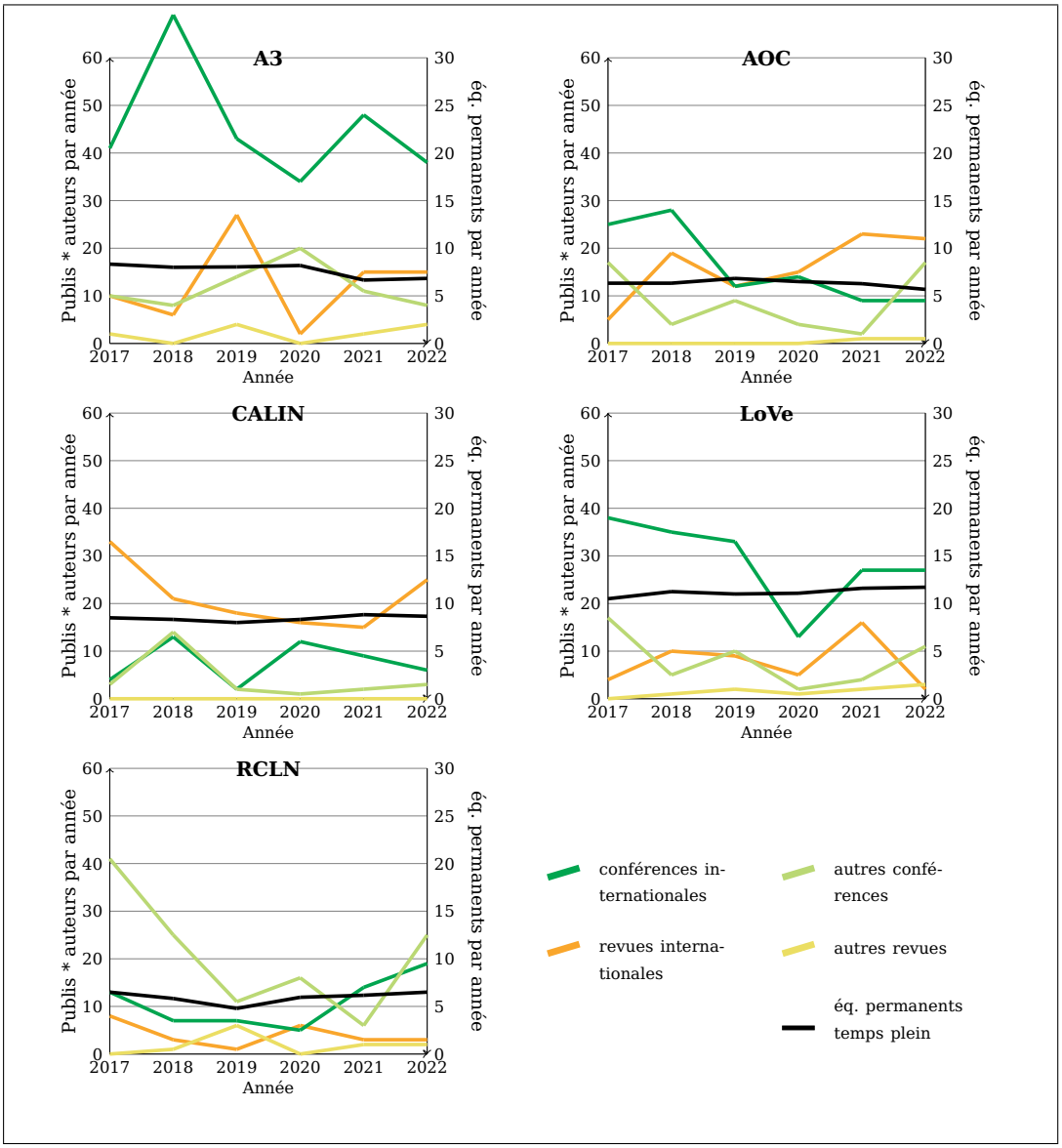


Figure 3.5: Nombre d'auteurs ayant contribué aux principales publications par équipe

ou en apprentissage les conférences internationales sont l'endroit prioritaire où diffuser ses résultats.

Certains membres de l'unité sont moins actifs. Si chaque cas est différent, cela correspond essentiellement à des prises de responsabilités administratives très lourdes et/ou à une volonté de réorienter sa carrière vers des directions alternatives à la recherche. Mettre en place des procédures d'accompagnement dans ces cas ne paraît adapté. Pour les quelques autres qui rencontrent des difficultés et souhaitent continuer à faire de la recherche, la participation à des séminaires, groupes de travail, à des projets collectifs leur permettent en général de remettre le pied à l'étrier.

Comme cela a été mentionné, les personnels d'appui à la recherche ont un apport essentiel dans tout ce qui concerne le développement logiciel et la constitution de corpus. Ces 3 IR CNRS ont néanmoins publié, dans la période évaluée, 6 articles dans des revues internationales et 9 communications dans des conférences internationales avec actes. Leur contribution à la production scientifique de l'unité est en tous points exemplaire.

3.3.3 Référence 3. La production scientifique de l'unité respecte les principes de l'intégrité scientifique, de l'éthique et de la science ouverte. Elle est conforme aux directives applicables dans ce domaine.

L'unité a présenté sa politique en matière de science ouverte et ses principes d'éthique scientifique et de reproductibilité de la science dans la section 3.2.2.

Suite à des décisions du conseil de laboratoire, le LIPN refuse toutes les publications en revue exigeant un paiement, et promeut les publications en *diamond open access*. L'équilibre entre médium de diffusion vertueux et bénéficiant d'une bonne visibilité n'est toutefois pas toujours facile à trouver. De même des discussions sont prévues (en conseil de laboratoire) au sujet des choix de conférences, certaines pratiquent en effet des montants prohibitifs de frais d'inscription depuis la pandémie, sans même garantir une très bonne visibilité aux communications qui y sont faites.

3.3.4 Synthèse de l'auto-évaluation

- **Forces et Atouts**
  - — *Une amélioration sensible de la qualité de la production scientifique* rendue visible par la qualité des supports de diffusion et la longueur conséquente d'une proportion importante de ces contributions.
  - — *Une bonne participation des jeunes chercheurs à la production scientifique.*

- **Faiblesses**
  - — *Un léger manque d'équilibre entre publications en revues et en conférences* : même si la culture de publications varie selon les domaines, certaines équipes gagneraient à rendre davantage visibles leurs meilleurs résultats dans des conférences de premier plan, et d'autres à publier plus d'articles sur leurs contributions marquantes dans les revues de référence.

3.4 Domaine 4. Inscription des activités de recherche dans la société

3.4.1 Référence 1. L'unité se distingue par la qualité et la quantité de ses interactions avec le monde non-académique.

Le LIPN développe des recherches fondamentales tout en menant une politique de coopération et de valorisation avec des industriels, politique qui s'est traduite notamment par 18 thèses en contrat CIFRE auxquelles s'ajoutent des thèses financées par Facebook et Deep Knowledge. L'unité a également tissé des relations au long cours avec Renault, Safran, la BPCE, EDF, RTE et Orange.

Les financements industriels (hors financements de la SATT Erganeo) représentent 22% des ressources propres de l'unité sur la période.

Les enjeux autour de la santé. Le LIPN s'implique dans les défis autour de l'IA de la santé. Ainsi deux étudiants en thèse au LIPN, ou sur le point de l'être, avec des financements de l'entreprise de rechercher médicale Deep Knowledge ont reçu le Prix de la recherche médicale 2022 de la Fondation de France pour le travail qu'ils ont réalisé durant leur stage de master. Il s'agit de

- — Leo Dechaumet pour de l'apprentissage auto-supervisé de représentations pour la super-résolution d'images issues de la microchirurgie de la main ;

— Nosseiba Ben Salem pour de l'apprentissage profond géométrique de représentations pour la détection d'objets dans des images échographiques du canal carpien.

Le LIPN est également partenaire de l'IHU PROMETHEUS sur le sepsis qui vient d'être accepté en mai 2023.

Les enjeux autour de la fiabilité. Le LIPN collabore avec l'industrie aérospatiale et automobile sur des questions de fiabilité des systèmes, cela inclut notamment

— une traduction d'un formalisme (Time4sys) proposé par Thales, afin de permettre une vérification automatisée de systèmes temps-réel [CI-75, CI-69] et la vérification [**CI-70**] d'un algorithme distribué mis au point par Thales ;

— une vérification [RI-33] d'un ordonnanceur d'Ariane avec le logiciel IMITATOR, ainsi que des algorithmes de monitoring efficaces avec application à des exemples issus de l'industrie automobile [**CI-51**, CI-53, **CI-52**, ↑RI-13].

Les enjeux environnementaux. Des travaux de recherche en lien avec l'éco-responsabilité démarrent au LIPN notamment sur le green cloud computing et sur la logistique de l'économie circulaire. Un membre du LIPN a également écrit un article destiné à un large audience dans *The Conversation* sur l'ÉcoIndex, un outil destiné à mesurer le score environnemental des sites web. Une communication sur la même thématique est acceptée à la conférence IEEE Compsac 2023 et le projet FogSLA (AMI Cloud BPI) en lien avec cette question et impliquant de nombreuses entreprises a démarré en 2021.

3.4.2 Référence 2. L'unité développe des produits à destination du monde culturel, économique et social.

Le LIPN a une importante activité à destination du monde socio-économique dans la période et s'est impliqué dans la valorisation de ses activités (cf. portfolio 5 du LIPN). Cela inclut notamment les opérations soutenues par la SATT Erganeo : la création de la start-up HephIA et une preuve de concept d'un détecteur de malware. À cela s'ajoute l'édition de la norme ISO-IEC 15909-3, publiée en novembre 2021.

Le LIPN et l'USPN ont également soutenu des mobilités vers le secteur industriel

— Céline Rouveirol et Henry Soldano ont bénéficié d'une année de délégation en entreprise chez NukkAI dont Henry Soldano, maintenant MCF émérite, a intégré l'équipe - un contrat CIFRE est également en cours ;

— Julien David a passé une année comme ingénieur de recherche chez Qwant ;

— César Rodriguez est en disponibilité et travaille comme ingénieur de recherche à Oxford, UK, dans la société Cadence Design Systems) ;

— Éric Alphonse est en disponibilité et dirige la société UBIQUITY AI.

3.4.3 Référence 3. L'unité partage ses connaissances avec le grand public et intervient dans des débats de société.

Promotion de l'enseignement de l'informatique dans le secondaire. Le LIPN s'investit très fortement dans la promotion de l'enseignement de l'informatique dans le secondaire (cf. portfolio 6 du LIPN). Cela se traduit notamment par :

— le montage et le pilotage d'un Diplôme Inter-Universitaire « EIL » (Enseigner l'Informatique au Lycée) pour 24 enseignants de l'académie de Créteil ;

— l'animation du groupe informatique avec participation de plusieurs chercheurs du LIPN ;

— l'organisation du stage « Coq en math ou l'art de raisonner », pour lycéennes et lycéens de 2nde su Lycée Joséphine Baker de Pierrefite (2022) ;

— la participation aux jurys du CAPES NSI et de l'agrégation d'informatique récemment créés.

Diffusion de la culture scientifique en direction du grand public. Le LIPN a également mené de nombreuses actions de partage de la connaissance avec le grand public et le jeune public :

- Interview sur la détection de malware pour France Culture (2018) ;
- Interview dans le cadre du programme « Incroyables recherches » de la Cité des sciences et de l'industrie en décembre 2020 : *Ma recherche : comment trouver la meilleure combinaison ?* ;
- Édition d'un numéro spécial de la revue *Tangente* consacré à la Recherche Opérationnelle [ED-2] ;
- Participation via conférences et interviews à la « Fête de la science »/« Savante banlieue » ;
- Mise en place d'animations scientifiques tous publics et tous continents, comme celles autour des puzzles (cf. Fig.3.6) ;
- Participation à la création de plusieurs laboratoires de mathématiques du plan Villani-Torossian ;
- Présentations concernant le métier de chercheur/chercheuse et/ou les thématiques du LIPN à des élèves de collège et lycée (toute la période) ;
- Participation aux journées « Filles et Maths : une équation lumineuse », organisées à l'USPN et s'adressant à des lycéennes (de la 2nde à la terminale) avec pour but de les encourager à s'orienter vers des études scientifiques et techniques (toute la période).
- animation d'ateliers Math.en.Jean (2017–2020).

Figure 3.6: Écoliers et étudiants jouant avec des puzzles inspirés de la théorie des pavages apériodiques.

Promotion des logiciels libres et de l'ouverture des données de la recherche. Comme cela a déjà été évoqué dans ce rapport, le LIPN est aussi impliqué dans des actions en faveur de la promotion des logiciels libres et de l'ouverture des données de la recherche de l'échelle locale au niveau national et international. Cela se traduit notamment par

- la participation (1 membre) au Conseil scientifique DoRANum disciplinaire, qui vise à accompagner les universitaires français(es) dans l'ouverture de leurs données de recherche (depuis 2021) ;
- la participation (1 membre) du groupe « EOSC (European Open Science Cloud) Future User Group », qui vise à préparer l'ouverture de la plateforme EOSC ayant, parmi d'autres, l'objectif d'ouvrir les données de la recherche des universitaires européen(ne)s (depuis 2021) ;

— la participation (1 ambassadeur) au projet Software Heritage, ayant pour ambition de collecter, préserver et partager tous les logiciels qui sont publiquement disponibles sous forme de code source.

3.4.4 Synthèse de l'auto-évaluation

- **Forces et Atouts**
  - — *Une sensible augmentation (de 70%) du nombre de thèses en contrat* CIFRE : il y a eu 18 contrats CIFRE sur la période contre 11 sur la période précédente. Le montant de l'environnement associé à ces contrats est également en nette augmentation et peut atteindre jusqu'à 73k€.
  - — *Une fort investissement dans la promotion de l'enseignement de l'informatique dans le secondaire.*
  - — *Un renforcement des partenariats industriels et des activités de valorisation.*

- **Faiblesses**
  - — *Une valorisation des logiciels encore trop limitée* : la diffusion des logiciels a besoin d'être améliorée et rationalisée.

Chapitre 4
Trajectoire de l'unité

Sommaire

4.1 Atouts, Faiblesses, Opportunités et Menaces 54
4.2 Positionnement relativement au projet 2019-2023 56
4.3 Projet pour la période 2023-2028 56
4.3.1 Complexités 56
4.3.2 Interdisciplinarité 58
4.3.3 Collaborations internationales et projets européens 58
4.3.4 Visibilité nationale et internationale du LIPN 58
4.3.5 Le développement logiciel et la **dev-team** 58
4.3.6 Science ouverte, reproductibilité et diffusion 59
4.3.7 Structuration du LIPN : Équipe LoVe 59
4.3.8 Parité et inclusion 60
4.3.9 Développement durable 60
4.3.10 Communiquer plus et mieux 61

4.1 Atouts, Faiblesses, Opportunités et Menaces

- **Forces et Atouts**
  - *Bonne reconnaissance scientifique sur tous les axes majeurs* : cela se concrétise notamment à travers la responsabilité de projets et l'organisation de conférences reconnues, d'écoles d'été, de fortes responsabilités dans des GDR (Recherche Opérationnelle, Informatique Mathématique, Traitement Automatique des Langues), la co-responsabilité d'un GDR international (AleaNetwork) et d'un IRN qui démarre avec le Danemark.
  - *Large spectre scientifique* : les thématiques scientifiques du laboratoire couvrent un spectre scientifique large, alliant des compétences théoriques et appliquées. Cette couverture thématique favorise les interactions entre différents domaines de l'informatique.
  - *Bonne interdisciplinarité* : les recherches menées au LIPN bénéficient d'une fructueuse interdisciplinarité avec des domaines variés (Mathématiques, Physique, Linguistique, Philosophie, Santé, Arts).
  - *Attractivité en termes de recrutements* : chaque année des candidats CR et DR demandent une affectation au LIPN, deux CR et un DR sont ainsi arrivés au LIPN sur la période ; les recrutements d'enseignants-chercheurs ont notamment attiré des candidatures internationales au niveau MCF et des mutations au niveau PU.
  - *Master informatique, école d'ingénieur* : les formations sont pérennisées et bien reconnues.
  - *Lien avec les entreprises* : de nombreux projets ont été réalisés en partenariat avec des entreprises de toutes tailles – notamment via des contrats CIFRE –, sur presque toutes les thématiques présentes au LIPN.
  - *Valorisation des recherches* : une start-up a été créée et des membres du LIPN sont fortement impliqués dans une autre. Un LabCom vient de démarrer.
  - *Présence dans les instances nationales* : directrice adjointe scientifique (maintenant directrice) de l'INS2I du CNRS, secrétaire scientifique de la section 6 du CoNRS, membres de la section 27 du CNU, membres du jury de l'IUF, ...
  - *Cohérence du laboratoire* : gestion collective et collaborations scientifiques inter-équipes en nette augmentation en dépit de la diversité des thématiques.
  - *Dynamisme du développement logiciel* : le développement logiciel a beaucoup progressé dans la période avec l'appui de la **`dev-team`** qui s'est constituée et grâce des investissements en matériel (serveur d'intégration continue, achats de cartes GPU). Deux IR CNRS (BAP E et BAP D) ont rejoint le laboratoire, plusieurs apprentis CNRS ont été et sont encore accueillis dans la **`dev-team`**.
  - *Gestion souple et maîtrisée du budget* : la gestion financière mutualisée et centralisée profite à tous les membres du laboratoire.

- **Faiblesses**
  - *Absence de projets européens collaboratifs* : les relations internationales se traduisent encore insuffisamment dans des projets européens. Le LIPN a participé à un projet H2020 mais sans être partenaire. Un projet Horizon Europe impliquant le LIPN est en cours d'évaluation.
  - *Manque de chercheurs CNRS sur certaines thématiques* : le LIPN compte deux DR et sept CR dans seulement deux de ses cinq équipes (CALIN et LoVe). La possible arrivée d'une CR CNRS dans l'équipe AOC pourrait améliorer la situation.
  - *Couverture du laboratoire par les masters encore à améliorer* : le master d'informatique local ne fournissait pas un vivier suffisant de doctorants et ne concerne que deux équipes (A^3 et LoVe). Le démarrage de l'EUR M&CS commence à amélio-

rer la situation. La participation à des masters à l'étranger (Madagascar, Vietnam, Maroc) ouvre des possibilités supplémentaires.
— *Fortes charges administratives* : des enseignants-chercheurs du LIPN ont d'importantes responsabilités (Président de l'USPN, trois VP de l'USPN, directeur et directeurs-adjoints pour l'IG dont celui chargé de l'école d'ingénieur, directeur de l'école doctorale, directrice adjointe scientifique de l'INS2I du CNRS, secrétaire scientifique de la section 6 du CoNRS). Ces charges diverses affectent *de facto* le potentiel de recherche.
— *Manque d'un personnel dans l'équipe système réseau* : une seule administratrice système réseau gère l'ensemble de l'infrastructure pour le LIPN et LAGA. Le recrutement d'un personnel supplémentaire est indispensable vu l'ampleur de la tâche. Un poste d'IR CNRS est ouvert au recrutement, reste à trouver la perle rare.

- **Opportunités**

— *Un pôle de recherche MathSTIC au nord de Paris* : le LIPN est moteur dans la fédératin MathSTIC qui, avec un potentiel de 200 permanents, contribue au développement des recherches à l'interface Math-STIC. L'EUR M&CS renforce cette dynamique via la formation d'étudiants et de doctorants à cette interface Enfin, le déménagement des 3 laboratoires de la fédération dans le nouveau bâtiment devrait encore améliorer les collaborations et la visibilité des recherches ainsi menées.
— *Coopérations internationales* : De nombreuses collaborations internationales existent et peuvent servir de terreau à la construction de projets internationaux de plus grande ampleur.
— *Diffusion de logiciels* : le dynamisme de la **`dev-team`** devrait permettre d'améliorer la visibilité et la diffusion des logiciels développés au LIPN.

- **Menaces**

— *Fin du LabEx EFL* : L'équipe RCLN est très impliquée dans le LabEx EFL qui se termine fin 2024. Elle a bénéficié de l'environnement scientifique ainsi créé, et de financements conséquents directs et indirects (ANR DALIH). Un nouveau projet est en cours de rédaction. Son financement et la place laissée aux partenaires hors IdEx sont des enjeux importants pour le LIPN qui dépendent beaucoup de la future présidence de l'Université Paris Cité.
— *Pénurie de postes* : La situation actuelle en terme de postes est inquiétante. Au cours de la période, le nombre de rangs A a tout juste été maintenu alors qu'il était faible, et le nombre de MCF s'est érodé. In est indispensable qu'*a minima* l'ensemble des postes permanents libérés soient republiés. Il est également impératif de consolider les forces vives, **en particulier en rangs A**.
— *Nombre d'HDR* : Le LIPN compte de nombreux MCF titulaires d'une HDR. Cela présente deux risques de natures très différentes : (i) il pourrait y avoir de nombreux départs en promotion - ce qui s'annonce pour la rentrée 2023, (ii) malgré les repyramidages la difficulté à obtenir un poste de professeur.e pourrait conduire à une démotivation de certains MCF habilités.
— *Détachements, mises à disposition et disponibilités* : Le LIPN connaît une recrudescence de départs en détachements (INS2I, UE), de mises à disposition et de disponibilités. Cela fragilise l'unité d'autant que l'USPN n'est pas en mesure d'ouvrir des postes permanents en contre-partie, les personnels étant susceptibles de réintégrer l'université.
— *Départs dans l'équipe administrative* : Les 2 gestionnaires CNRS partiront en retraite au cours du prochain quinquennat et la gestionnaire USPN est en CDD. Le renouvellement sur des postes permanents de ces personnels sera un réel enjeu pour que l'équipe administrative ne soit pas déstabilisée.

4.2 Positionnement relativement au projet 2019-2023

Durant la période, le LIPN a, pour l'essentiel, conforté ses recherches autour de ses axes forts malgré de profondes transformations du paysage de la recherche sur les thématiques liées à l'IA.

Le LIPN a beaucoup décloisonné les activités de ses équipes notamment via des co-encadrements, des projets communs ou des co-publications, mais surtout avec la création de l'axe transverse Complexités (portfolio 1 du LIPN) et l'émergence d'activités impliquant toutes les équipes autour de la science des données (portfolio 2 du LIPN). Il est à noter que l'USPN participe au projet d'alliance européenne Uninovis centré autour de la science des données. En ce sens le LIPN s'inscrit pleinement dans la stratégie de l'USPN.

Le LIPN a également amplifié son parti pris d'interdisciplinarité dans la période (portfolio 3 du LIPN), a diversifié ses activités partenariales (avec le monde académique et non-académique) et a fait d'importants efforts de diffusion des résultats des ses recherches via du développement logiciel (portfolio 4 du LIPN) et de la valorisation (portfolio 5 du LIPN). Au delà de l'alliance européenne, la stratégie propre de l'USPN repose sur l'interdisciplinarité à tous les niveaux et une forte ouverture vers la société. La stratégie du LIPN est à nouveau en pleine cohérence avec ces objectifs.

Si les collaborations internationales du LIPN sont nombreuses, et si l'unité est attractive et compte des membres de grande qualité, l'objectif de décrocher des projets européens n'a pas été atteint. Le LIPN a seulement bénéficié de financements du projet H2020 Philhumanis sans en être partenaire et d'un financement de postdoc Cofund via un appel régional. Un MCF a passé l'oral de l'ERC Starting Grant en 2017 et le LIPN compte un membre extérieur dans le projet ERC Advanced Grant Geoscape. Mais malgré un certain nombre de dépôts de projets ERC et Marie Skłodowska-Curie aucun financement de ce type n'a été obtenu. Cela reste un objectif de la période à venir.

4.3 Projet pour la période 2023-2028

Le LIPN compte amplifier la mise en œuvre de sa stratégie de recherche ambitieuse (cf. section 1.3.1), s'appuyant sur de l'interdisciplinarité et sur de nombreuses collaborations avec l'environnement proche, ainsi qu'avec des partenariats (inter)nationaux.

L'idée est encore de soutenir les axes forts des équipes, certains bien sûr yant évolué depuis la dernière évaluation. Le LIPN tient également à renforcer les thématiques de recherche transverses notamment autour de la complexité et de la science des données.

4.3.1 Complexités

L'un des objectifs du LIPN dans les prochaines années est de consolider les activités de l'axe transverse Complexités (voir portfolio 1 du LIPN), qui constitue une originalité en France, avec pour ambition d'en faire le point d'ancrage national sur ce thème. En effet, la thématique souffre en France d'un manque d'unité et de visibilité : absence d'équipe dont la recherche serait principalement sur ce thème, absence de structuration (*i.e.*, séminaire national, GT de GDR, *etc.*). Les postes publiés en complexité ne trouvent pas de candidat.e.s car peu de thèses et de HDR sont soutenues sur le sujet.

Le LIPN souhaite pallier ces manques de plusieurs façons.

— Pour consolider l'axe Complexités et améliorer son rayonnement à l'échelle nationale, le séminaire hybride sera poursuivi et davantage de place sera laissée à des exposés de personnes extérieures au groupe de travail. L'hétérogénéité des membres de l'axe impose toutefois des formats vraiment interactifs afin d'aboutir à de réelles collaborations. Ainsi, une majeure partie des activités de l'axe consistera en des séances de travail en commun autour d'un article ou d'une série d'articles afin de comprendre ceux-ci en profondeur.

— Pour mieux structurer la communauté nationale et améliorer la formation des jeunes chercheurs sur ces questions, des évènements nationaux et internationaux seront organisés (p.e. les écoles Caleidoscope) et des séminaires ou journées nationales seront mis en place.
— Pour donner une plus forte visibilité à l'international à la recherche autour de la complexité qui se fait au LIPN et en France, des séjours de longue durée au LIPN seront particulièrement recherchés (par exemple via les accueils en délégation au niveau national, des financements de professeur.e.s invité.e.s, l'IRN avec le Danemark ou des financements postdoctoraux au niveau international). Ceci devrait permettre à plus long terme d'attirer des candidat.e.s pour des doctorats, des postdocs, ou même des postes permanents, que ce soit au LIPN ou plus globalement en France.

Pour mettre en place l'ensemble de ces activités, des financements seront recherchés au travers de programmes nationaux ou européens.

- **La science des données**

Le thème de la science des données est naturellement présent dans les équipes A³ et RCLN : d'une part, par définition, l'apprentissage artificiel repose de manière essentielle sur les données ; d'autre part, la plupart des recherches de l'équipe RCLN consistent en l'exploration, le traitement et l'extraction de connaissances à partir de grande masses de données textuelles. La différence principale est que l'équipe RCLN conçoit et développe des modèles d'apprentissage profond et symbolique dédiés à des contextes spécifiques, en rapport avec le traitement des langues, tandis que dans l'approche générale de l'équipe A³ il n'y a pas de tâche fixée *a priori*.

Les autres équipes du laboratoire, quoique moins directement impliquées dans la science des données, ont toutes des contributions à l'interface de cette thématique, ce qui la rend véritablement transverse à l'unité. Il s'agit même d'une transversalité qui va au delà du LIPN, car les interactions entre la science des données et les statistiques ou la topologie algébrique relèvent plutôt du LAGA. Dans ce contexte, le LIPN envisage de donner davantage de visibilité à cette thématique et de créer un axe transverse « Science des Données », suivant le même principe que l'axe Complexités créé récemment.

Les interactions autour de la science des données au LIPN peuvent se faire à au moins deux niveaux, pas nécessairement mutuellement exclusifs :

— l'exploration de techniques pour l'exploitation de données, ou l'étude des propriétés (du point de vue informatique mais aussi mathématique) de ces techniques. Les équipes impliquées ici sont A³, AOC, CALIN et LoVe.
— L'adaptation et le développement de ces techniques à des domaines divers, typiquement pour extraire des connaissances. Les équipes impliquées ici sont A³, RCLN et AOC.

Concrètement, les activités de cet axe peuvent prendre la forme d'un séminaire inter-équipes. Quelque chose de similaire a déjà été fait en février 2023 : l'équipe RCLN a animé un séminaire sur le fonctionnement de ChatGPT, auquel ont participé des membres de toutes les équipes du laboratoire. L'organisation de journées thématiques, sur le modèle de celles déjà organisées en 2018, 2019 et janvier 2023, sera un autre moyen de faire vivre cet axe. De nouveaux co-encadrements inter-équipes de stages ou de thèses (certains sont en cours) et la venue de professeurs invités pourraient également contribuer à l'activité sur cette thématique de la science des données.

La dynamique instaurée pourra, à terme, conduire les recherches de cet axe transverse à avoir une influence sur les recrutements, en particulier au cas où des interactions fécondes devraient émerger sur ce thème, qui nécessiteraient des moyens supplémentaires par rapport à ceux déjà existants dans l'unité.

4.3.2 Interdisciplinarité

Le LIPN souhaite continuer à fortement soutenir les recherches interdisciplinaires.

Les recherches à l'interface Math-STIC devraient être dynamisées par l'élargissement des axes de collaboration (un axe de recherche autour des catégories à vu le jour dans la fédération en 2022), par l'augmentation des effectifs de l'EUR, ainsi que certains recrutements récents et le regroupement dans un bâtiment commun.

L'avenir du LabEx EFL est un enjeu important notamment pour l'équipe RCLN, et le LIPN soutient pleinement la préparation d'un nouveau projet sur le même périmètre d'unités avec l'espoir que les conditions de portage des projets qui en émaneront seront maintenues.

Les interactions avec la philosophie sont en plein développement, notamment autour du concept d'algorithme (séminaire interdisciplinaire, stage, possible thèse). Conforter cette dynamique est un objectif de la période à venir.

Enfin, dans l'interface avec la santé, il est à noter que le LIPN est partenaire de l'IHU Prometheus « Approche par médecine de précision pour la prise en charge thérapeutique du sepsis communautaire et lié aux soins » qui vient d'être accepté.

4.3.3 Collaborations internationales et projets européens

La LIPN continuera à s'appuyer sur son réseau de collaborations internationales pour mener ses recherches et cherchera à les consolider via des activités contractuelles. Ainsi un IRN avec le Danemark démarre en 2023, et des projets, que le LIPN cherche à pérenniser, sont en cours avec des laboratoires de NYU Abu Dhabi. Plusieurs projets Erasmus+ (Vietnam, Tunisie, Maroc) ont également permis et permettent encore la mobilité entrante de doctorants et la mobilité sortante d'(enseignants)-chercheurs viennent de démarrer. Le LIPN souhaite élargir son réseau de collaborations internationales financées via Erasmus+.

En termes de projets internationaux, la priorité sera toutefois de travailler, en lien avec la cellule relations internationales, sur les dépôts de projets européens Horizon Europe, ERC et Actions Marie Skłodowska-Curie. Il est à noter qu'un projet Horizon Europe dont le LIPN est partenaire est en cours d'évaluation.

4.3.4 Visibilité nationale et internationale du LIPN

Le LIPN souhaite améliorer sa visibilité aux niveaux national et international pour attirer de bonnes recrues, pour étoffer son réseau de collaborations, pour mieux diffuser les résultats de ses recherches.

Dans cette perspective, l'unité prévoit de continuer à organiser des conférences internationales prestigieuses (ETAPS 2023, ...) et à s'impliquer dans les réseaux nationaux et internationaux d'animation de la recherche. A destination des jeunes chercheurs, elle encouragera l'organisation d'Ecoles (thématiques) à fort rayonnement. (EPIT/Caleidoscope 2023, CIMPA ...) et favorisera la présence de jeunes chercheurs dans les évènements locaux.

4.3.5 Le développement logiciel et la `dev-team`

D'une manière générale, la **`dev-team`** a pour ambition de poursuivre ses activités d'encadrement et de développement avec l'objectif d'intégrer tous les membres du LIPN qui développent des logiciels dans le laboratoire afin de constituer une équipe de développement scientifique transverse. Les activités envisagées par la commission développement logiciel s'inscrivent dans la même dynamique.

Un.e apprenti.e et un.e IR CNRS sont en cours de recrutement. Le premier recrutement est destiné principalement au développement d'une version stable de la plateforme Chênetal. Le deuxième sera mutualisé avec deux autres laboratoires franciliens (IRIF et LMF) et travaillera principalement sur des logiciels de vérification formelle. Le recrutement d'un autre apprenti pourrait être envisagé à la fin du contrat de l'apprenti actuellement présent en 2024. La **dev-team** cherchera aussi à améliorer la parité de genre : elle ne compte actuellement qu'une seule femme permanente et une stagiaire.

En termes de développement, la **dev-team** continuera son développement de l'intranet de l'unité et travaillera à la robustesse des plateformes Chênetal et CosyVerif de manière à ce qu'elles atteignent un niveau de robustesse qui permettra leur valorisation avec un dépôt APP, par exemple.

4.3.6 Science ouverte, reproductibilité et diffusion

Concernant la science ouverte, le LIPN travaille à rendre accessibles (presque) toutes ses publications, mais aussi ses données et ses logiciels.

En terme d'éthique scientifique, la reproductibilité des tests logiciels est un point essentiel. Dans cette perspective, la commission développement logiciel s'est donné pour objectif d'encourager les membres du laboratoire

- à utiliser le serveur d'intégration continue et les technologies de conteneurisation et GUIX pour créer de pipelines reproductibles ;
- à développer des logiciels libres et à les archiver sur Software Heritage pour leur préservation ;
- à utiliser le serveur d'intégration continue pour garantir la robustesse des logiciels (tests).

Pour améliorer la diffusion des logiciels produits, la commission logiciel prévoit d'alimenter automatiquement le site web du laboratoire avec tous les logiciels développés via l'API développée pour recenser les publications. Une feuille de route devrait être rédigée afin de diffuser à l'ensemble du personnel les bonnes pratiques de développement des logiciels de leur implémentation jusqu'à leur publication (compilation, testing, packaging, documentation et publication sur le serveur d'artefacts et de façon automatique grâce au serveur d'intégration continue). De même, il sera précisé quelles sont licences open-source et comment générer/publier des sites web avec GitLab pour afficher les logiciels. Chaque étape de la feuille de route sera accompagnée d'une formation proposée par la commission.

La commission logiciel a d'ores et déjà prévu les formations suivantes (classées par thématique) :

- *Cycle de développement logiciel* : Introduction aux licences de logiciels ; Le système de contrôle de version `Git` ; La technologie de conteneurisation `Docker` ; Serveur d'intégration continue et déploiement continu (CI/CD) ; Création de pages statiques avec `Jekyll` et leur déploiement avec GitLab pages.
- *Cycle de frameworks* : Programmation parallèle ; Le langage de programmation `Julia` ; Réseaux de neurones avec le framework `PyTorch` ; Slurm et son utilisation dans le cluster de GPUs du LIPN.
- *Cycle de développement reproductible* : Notebooks pour la recherche (Org Mode et Jupyter notebooks) ; Création des figures interactives avec Python ; Environnements logiciels reproductibles avec Guix ; Archivage pérenne des logiciels et leur citation avec Software Heritage.

4.3.7 Structuration du LIPN : Équipe LoVe

En terme de structuration de l'unité, une équipe, la plus importante numériquement, a un statut un peu particulier. L'équipe Logique et Vérification (LoVe) regroupe en effet deux groupes dynamiques qui mènent de manière autonome des recherches sur deux

thématiques indépendantes. La bipolarité de cette équipe incite à réfléchir à la création d'une sixième équipe pour donner plus de visibilité à ces thématiques.

L'axe Vérification, après avoir connu des départs en série, s'est très récemment renforcé : trois de ses huit membres sont arrivés depuis fin 2021. Si la situation de cet axe se stabilise, il sera profitable à tous de restructurer cette équipe.

4.3.8 Parité et inclusion

La commission pour la parité femmes-hommes qui s'est récemment constituée prévoit notamment

- de communiquer autour du harcèlement moral ou sexuel,
- de lancer des actions de médiation à destination des lycéennes ou des collégiennes,
- d'organiser du mentorat à destination des doctorantes, pour les aider à préparer leur projets de post-doctorat ou de candidature à des postes permanents.

Plus généralement, cette commission souhaite élargir son travail sur la parité femmes-hommes et devenir une **commission traitant de l'inclusion** des différentes catégories de membres du LIPN, que ce soit en terme de discriminations de genre, d'orientation sexuelle, de handicap ou d'origine. Une pièce de **théâtre forum** à ce sujet sera présentée aux membres du laboratoire lors de l'assemblée générale de rentrée du laboratoire en octobre 2023.

4.3.9 Développement durable

La *commission développement durable*, très récemment créée, a pour objectif de réfléchir à des propositions d'actions, de formations, d'initiatives pour améliorer l'impact environnemental du LIPN. Ces propositions ont vocation à être régulièrement discutées en conseil de laboratoire et au besoin en assemblée générale, et surtout à être suivies d'effets.

La commission est d'ores et déjà en train de mettre en place diverses actions :

- sensibilisation des membres du laboratoire au quotidien (économie d'énergie, de gobelets, de papier, *etc.*),
- inscription à l'initiative Labo1.5 (en cours de remplissage) : chiffrage des missions, bilan carbone, bâtiments, *etc.*,
- évaluation du coût énergétique de l'utilisation des serveurs GPU (action envisagée). Il est important que chaque chercheur soit en mesure d'estimer l'empreinte carbone de son programme informatique en calculant la quantité de CO2 produite lors de son exécution,
- limitation de la multiplication d'équipements nouveaux (machines personnelles, serveurs), et promotion du recyclage interne ou externe (action envisagée),
- réduction de l'empreinte carbone en terme de transports (missions) : réduction de la place de l'avion, promotion du train, proposition d'alternatives aux déplacements longue distance et surtout courte durée (action envisagée),
- réduction de l'impact des achats : groupement de commandes, exigence de fournisseurs eco-responsables (action envisagée),
- introduction aux enjeux environnementaux dans les formations (action en cours),
- organisation d'évènements en marge des conférences habituelles pour alerter les chercheurs sur l'urgence climatique et la crise écologique,
- invitation au LIPN de membres de la fresque du numérique et/ou du climat (action envisagée).

A cela s'ajoutent les initiatives de la commission développement durable de l'IG qui prévoit, par exemple, d'organiser un cycle de conférences sur le climat pour l'année 2023/2024.

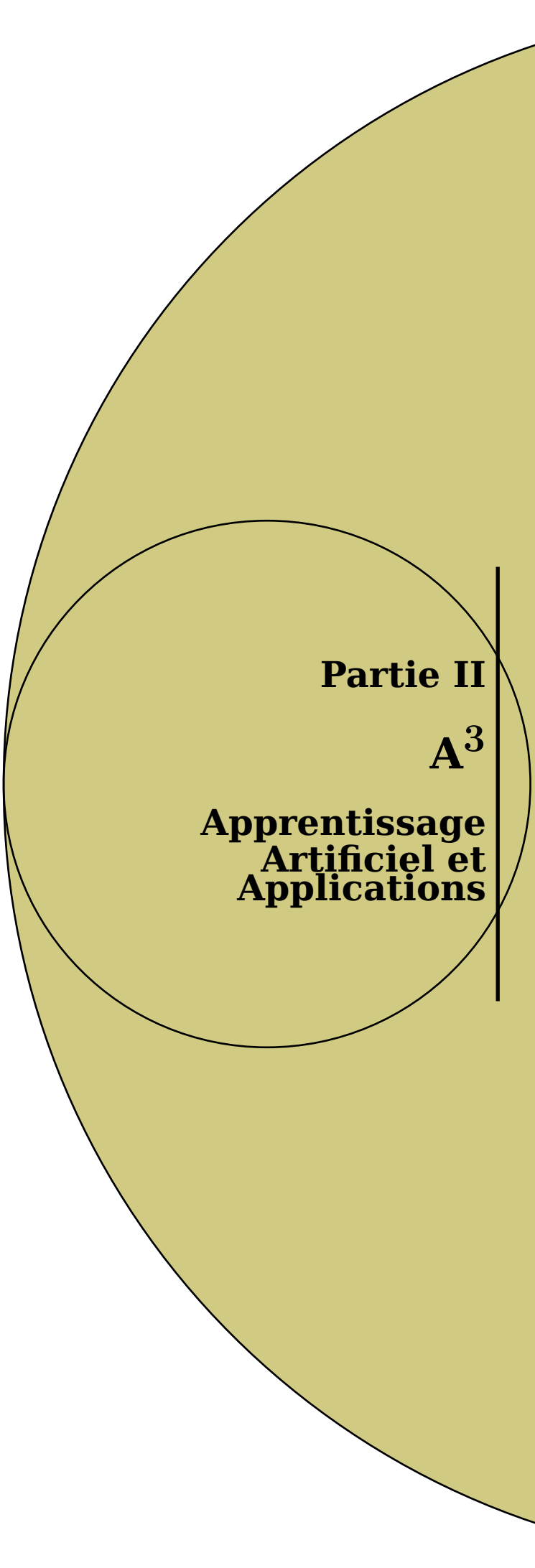
Il est à noter que des travaux de recherche en lien avec l'éco-responsabilité démarrent au LIPN notamment sur le green cloud computing et sur la logistique de l'économie circulaire. Un membre du LIPN a également écrit un article destiné à une large audience dans *The Conversation* sur l'ÉcoIndex, un outil destiné à mesurer le score environnemental des sites web. Une communication sur la même thématique est acceptée à la conférence IEEE Compsac 2023 et le projet FogSLA (AMI Cloud BPI) en lien avec cette question et impliquant de nombreuses entreprises a démarré en 2021.

4.3.10 Communiquer plus et mieux

La cellule com-web s'est donné pour objectif d'améliorer la communication interne et externe de l'unité.

Pour dynamiser la diffusion de l'information au sein du LIPN, outre l'amélioration continue du site web, elle réfléchit à la mise en place d'une lettre bi-hebdomadaire qui regrouperait informations scientifiques et récapitulatif des informations reçues sur la période (appels à projets, annonces de conférences, etc.).

En termes de communication externe, l'ambition est notamment de mieux faire remonter l'information pour une diffusion sur les différents canaux disponibles (réseaux sociaux, communication des tutelles) et de communiquer sur les activités du LIPN en s'appuyant sur les nombreuses photos qui ont été réalisées lors du reportage réalisé en partenariat avec la cellule de communication de l'INS2I.



Partie II

A^3

Apprentissage Artificiel et Applications

Sommaire

Informations générales sur l'équipe A3 **69**
1.1 Membres de l'équipe A3 70
1.2 Les thématiques scientifiques et leurs enjeux 72
1.3 Profil d'activités liées à la recherche 80
1.4 Prise en compte des recommandations du précédent rapport 80

Introduction du portfolio **83**

Auto-évaluation de l'équipe A3 **85**
3.1 Domaine 2. Attractivité 86
3.2 Domaine 3. Production scientifique 88
3.3 Domaine 4. Inscription des activités de recherche dans la société 105

Trajectoire de l'équipe A3 **107**
4.1 Organisation et ambitions de l'équipe 107
4.2 Analyse AFOM/SWOT 108
4.3 Perspectives et projet de recherche 109

Overview

Machine learning is a subfield of artificial intelligence that involves the design and development of algorithms and models, allowing computers to learn from observations and prior knowledge, without being explicitly programmed, to improve their performance in the execution of tasks (predictions, decision-making, etc.). Machine learning borrows techniques and theoretical backgrounds from several fields, including mathematics (Linear algebra, calculus, probability theory, geometry, optimization, information theory, etc.), statistics providing the framework for evaluating the accuracy and reliability of machine learning models, computer science, and cognitive psychology.

The latest advances in machine learning, particularly those initiated by transformers and more recently ChatGPT, are profoundly changing all areas of society. They are transforming relationships with work, education, research, health, entertainment, control systems, people, and access to personal data, among others.

In the context of classical machine learning and metalearning, our group covers a broad range of topics, from symbolic and statistical Learning to graph mining and reinforcement learning. This allows us to be quite reactive to cope with new challenges raised by emerging applications of Machine Learning. It also makes it possible to study innovative combinations of learning methods for tackling complex problems.

The structure of the group has evolved over the years, and now consists of three research axes :

- *Learning from Data and Learners*, which gathers research that deals with collaboration and transfer learning;
- *Relational Learning and Graphs*, motivated by learning from models expressed as logical, explicit and explainable programs and from data structured in graphs;
- *Meta-learning and Structures Learning*, which generalizes the learning problem by treating the learning process itself. It also allows simultaneous modeling of data and clustering (deep unsupervised learning).

In each of these axes, theoretical issues are addressed, and algorithms and software are developed, in the context of collaborative projects, including both academic and industrial partners. The team publishes in main Machine Learning international conferences (IJCNN, ICML, ECML, AAAI, IJCAI to name a few) as well as main international journals (Artificial Intelligence, Machine Learning, Pattern Recognition, . . .).

The strong links of the group with industrial partners is in particular testified by the high number of PhD in the group that take place in an industrial context.

The research topics of the team for the next five years aim at investigating promising areas where the team may emerge as a leader :

- Using emerging mathematical formalisms: algebraic topology, optimal transport
- Learning to learn including biais learning, decision boundaries, weak convergence and selection of admissible functions.
- Interpretability and symbolic machine learning
- Mining augmented and heterogeneous graphs
- New challenges for for simultaneous modeling of data and clustering (deep unsupervised learning) using multi-varied temporal data, large-scale learning, images and audio or video documents.

The group actively participates in national projects, and has several international collaborations. The researchers of the team are involved in teaching at the undergraduate and graduate levels, in USPN and in other institutions. In particular, the group has created a highly visible second year master specialization on Data Mining, Analytics, and Knowledge Discovery (*Exploration Informatique des Données et Décisionnel* (EID^2)) of the USPN master in computer science.

Chapitre 1

Informations générales sur l'équipe A3

Sommaire

1.1 Membres de l'équipe A3 70
1.1.1 Membres permanents 70
1.1.2 Membres non-permanents au 01/01/2023 71
1.1.3 Autres personnels temporaires du 01/01/2017 au 31/12/2022 71
1.1.4 Évolution de l'équipe (permanents) 72
1.2 Les thématiques scientifiques et leurs enjeux 72
1.2.1 Axe ADA : Apprentissage à partir de Données et d'Apprenants . . . 72
1.2.2 Axe ARG : Apprentissage Relationnel et Graphes 75
1.2.3 Axe MAR : Métaapprentissage et Apprentissage de Structures . . . 77
1.3 Profil d'activités liées à la recherche 80
1.4 Prise en compte des recommandations du précédent rapport . . . 80

Nom de l'équipe : Apprentissage Artificiel et Applications

Responsable de l'équipe : Aomar OSMANI

1.1 Membres de l'équipe A3

Tableau synthétique de l'évolution des membres de l'équipe (nombre de membres au 1er janvier de chaque année) :

Informations générales sur l'équipe A3

Corps	2017	2018	2019	2020	2021	2022	(2023)
DR	0	0	0	0	0	0	(0)
CR	0	0	0	0	0	0	(0)
PU	3	2	2	2	2	2	(2)
MCF	18	16	16	15	14	12	(12)
Total permanents	*21*	*18*	*18*	*17*	*16*	*14*	*(14)*
Post-docs et ATER	4	5	3	2	1	1	(1)
Ingénieurs (CDD)	0	0	0	0	0	1	(0)
Doctorants	11	14	15	14	11	14	(9)
Total	*36*	*37*	*36*	*33*	*28*	*30*	*(24)*

1.1.1 Membres permanents

(membres ou, *en italique, anciens membres*, et, **en gras, responsable de l'équipe** au 01/01/2023)

	Nom	*Prénom*	*Situation*	*Institution*
	Azzag	Hanane	HDR	IUTV
	Bennani	Younès	PU	IG
	Bouthinon	Dominique	MCF	IUTV
	Cabanes	Guénaël	MCF	IG
	Champesme	Marc	MCF	IG
	Chevaleyre	*Yann*	PU	IG
	Gérard	Pierre	MCF	IUTV
	Grozavu	*Nistor*	MCF	IG
	Guérif	Sébastien	MCF	IG
	Kanawati	Rushed	MCF	IUTV
	Lebbah	*Mustapha*	MCF	IUTV
	Matei	Basarab	HDR	IG
	Osmani	**Aomar**	HDR	IUTV
	Rouveirol	Céline	PU	IG

	Santini	Guillaume	MCF	IUTV
	Soldano	*Henry*	MCF	IG

1.1.2 Membres non-permanents au 01/01/2023

Nom	*Prénom*	*Situation*	*Financement*
Akouba	Meliha	doctorant	CD USPN
Amekeo	Kodjo M.	doctorant	CIFRE Groupe BPCE
Andriantsiory	Dina F.	doctorant	CD USPN
Barbara	Abir	doctorant	DeepKnowledge
Ben Salem	Nosseiba	doctorant	EUR M&CS
Faye	Bilal	doctorant	CD USPN
Khoufeche	Reda M.	doctorant	CIFRE DIM AI4DF
Kazi Aoual	Malik	doctorant	CIFRE Nukkai
Khalafaoui	Yasser	doctorant	CIFRE Alteca
Leclercq	Etienne	doctorant	CIFRE Alterway-Smile
Madane	Abdellah	doctorant	CIFRE SAFRAN
Modaresi	Seyed M.R.	doctorant	Ambassade de France
Zangato	Théo	doctorant	Ministère des Armées

1.1.3 Autres personnels temporaires du 01/01/2017 au 31/12/2022

Nom	*Prénom*	*Situation*	*Début*	*Fin*	*Financement*
Attaoui	Mohammed Oualid	doctorant	01/09/2018	30/06/2021	Campus France
Beck	Gaël	doctorant	01/03/2016	14/10/2019	CIFREKameleoon
Ben Bouazza	Fatima-Ezzahraa	doctorant	01/09/2017	26/11/2020	Bourse Maroc
Benlamine	Kaoutar	doctorant	01/09/2017	3/11/2020	CD USPN
Benyettou	Assia	doctorant	15/10/2014	02/11/2017	Bourse Algérie
Brahmi	Ibtissam	Postdoc	01/09/2016	30/11/2019	Bourse CNRST
Chahdi	Hatim	doctorant	01/09/2014	04/07/2017	ANR Coclico
Chelly	Zaineb	Postdoc	01/03/2016	28/02/2018	MARIE SKLODOWSKA-CURIE
Chevallier	Marc	doctorant	15/07/2019	23/11/2022	CIFRE Synaltic
Coutant	Anthony	Postdoc	15/01/2016	30/06/2018	ChistEra Adalab

Dilmi	Djallel	Postdoc	01/02/2022	02/06/2023	Amies
El Hamri	Mourad	doctorant	01/09/2019	15/12/2022	CD USPN
El Hamri	Mourad	ATER	01/09/2022	30/08/2023	USPN
Falih	Issam	doctorant	01/10/2015	08/03/2018	CD USPN
Forest	Florent	doctorant	01/12/2017	22/03/2021	CIFRE SAFRAN
Foucade	Yohan	doctorant	01/03/2019	10/06/2022	Contrat Facebook
Goffinet	Etienne	doctorant	01/12/2018	15/12/2021	CIFRE Renault
Hamidi	Massinissa	doctorant	01/10/2017	20/12/2022	CD USPN
Khoufeche	Reda M.	Ing.d'étude CNRS	01/01/2022	31/12/2022	France Relance ACOMI
Léger	Hippolyte	doctorant	01/09/2015	14/02/2019	CD USPN
Rastin	Parisa	doctorant	01/09/2015	26/06/2018	CIFRE Mindlytix
Rifi	Mouna	doctorant	01/02/2016	03/05/2019	CIFRE EDF
Sarazin	Tugdual	doctorant	04/02/2013	26/06/2018	CIFRE Smokewatchers
Veillon	Anne-Lise	doctorant	01/10/2014	08/03/2018	CD USPN
Veillon	Anne-Lise	Postdoc	01/02/2018	30/06/2018	ChistEra Adalab
Viard	Tiphaine	Postdoc	01/10/2019	31/08/2020	FUI-PCU
Zaiou	Ahmed	doctorant	10/10/2020	03/12/2022	CIFRE EDF
Zevio	Stella	doctorante	15/09/2017	23/02/2021	FUI-PCU
Zouinina	Sarah	doctorant	23/03/2017	06/10/2020	Bourse Maroc

1.1.4 Évolution de l'équipe (permanents)

Les effectifs de l'équipe A³ ont subi une forte baisse depuis 2017 avec le départ de plusieurs personnes actives. Il n'y a eu aucune arrivée dans l'équipe durant cette période. Un recrutement MCF est prévu pour la rentrée 2023/2024. Un fait remarquable concernant la composition de l'équipe est qu'elle continue d'accueillir de nombreux doctorants, en particulier avec des financements CIFRE, malgré la baisse de l'effectif des permanents.

Départs	Arrivées
2017 Y. Chevaleyre, PU, IG (mutation PU Univ. Paris-Dauphine)	Aucune arrivée
2021 H. Soldano, MCF, IG, (retraite)	
2022 N. Grozavu, MCF, IG, (PU, Université de Cergy)	
2022 M. Lebbah, MCF, IUTV, (PU, Université Versailles)	

1.2 Les thématiques scientifiques et leurs enjeux

En 2017, l'équipe A^3 était structurée en quatre axes : un axe apprentissage pour l'aide à la décision, un axe apprentissage non supervisé pour le transfert et la collaboration, un axe apprentissage dans les graphes et un axe apprentissage de modèles topologiques à partir de données massives.
Le départ de plusieurs membres actifs de l'équipe a conduit à la fermeture ou à la réorientation des axes sur lesquels ils étaient actifs. Le retour d'un membre actif d'une délégation pour création d'entreprise à conduit à la création d'un nouvel axe. En 2023, l'organisation de l'équipe a trouvé un nouvel équilibre et ses activités sont structurées en trois axes : l'axe ADA (Apprentissage à partir de Données et d'Apprenants), l'axe ARG (Apprentissage Relationnel et Graphes) et l'axe MAS (Méta-apprentissage et Apprentissage de Structures).

1.2.1 Axe ADA : Apprentissage à partir de Données et d'Apprenants

Le thème central de l'axe de recherche ADA est l'apprentissage de représentations à partir de données et d'apprenants, pour la collaboration et le transfert. Il se focalise sur trois paradigmes d'apprentissage multi-modèles fondés sur trois formalismes mathématiques : apprentissage collaboratif/fédératif, apprentissage par transfert et apprentissage à partir de données multimodales à travers des formalismes émergents (pour l'apprentissage) tels que le formalisme quantique, le transport optimal, et la topologie algébrique. Les contributions de l'axe ADA abordent à la fois des recherches à caractère fondamental et des recherches plus appliquées, le plus souvent

soutenues par des projets collaboratifs académiques et industriels dans différents domaines tels que : santé, marketing digital et recommandation, diagnostic des systèmes complexes, qualité et anonymisation de données, analyse des réseaux sociaux, etc.

Dans le cas de la collaboration, nous nous intéressons à l'apprentissage non supervisé collaboratif (ou *collaborative clustering*). Il s'agit d'un paradigme d'apprentissage automatique permettant de préserver notamment la confidentialité des données en utilisant d'autres résultats de classification non supervisée, sans avoir recours aux données de ces dernières. Ayant une collection de bases de données distribuées sur plusieurs sites différents, le problème consiste à partitionner chacune de ces bases en considérant les données locales et les résultats des classifications distantes des autres bases collaboratrices, sans partage de données entre les différents sites. Le principe fondamental du clustering collaboratif est d'appliquer des algorithmes de clustering localement sur les différents sites, puis de faire collaborer les sites en partageant les résultats obtenus lors de la phase locale [RI-6]. Il existe deux grandes familles de clustering collaboratif : l'approche horizontale pour la collaboration des bases de données qui décrivent les mêmes individus avec des variables différentes ; l'approche collaborative dite verticale pour la collaboration de plusieurs bases de données décrites par les mêmes variables représentant des populations différentes. Parmi les cadres théoriques qui formalisent nos études, on peut citer le transport optimal (TO) qui définit un ensemble d'outils géométriques pour comparer les mesures de probabilité prises sur un espace métrique. Ces outils ont trouvé de multiples applications dans des domaines tels que l'économie, l'imagerie et la mécanique des fluides pour modéliser les transferts, les morphings et les déplacements. La géométrie du TO a des propriétés qui la rendent particulièrement adaptée pour résoudre des problèmes d'apprentissage artificiel, pour lesquels la capacité de comparer des mesures de probabilité empirique en grande dimension est souvent cruciale. Cette direction de recherche vise à établir un pont entre les problèmes de TO et le fondement de la science des données. Il cherche à développer de nouveaux outils méthodologiques pour l'analyse et la représentation de données accessibles sous la forme de distributions empiriques ainsi qu'à revisiter plusieurs cadres classiques de l'apprentissage artificiel (apprentissage profond, factorisation matricielle, ...) sous le point de vue du TO. Pour le contexte de l'apprentissage par transfert, les membres de l'axe ADA cherchent à utiliser un jeu de tâches pour influencer l'apprentissage et améliorer les performances sur une autre tâche. Cependant, ce paradigme d'apprentissage peut en réalité gêner les performances si les tâches (sources et cibles) sont trop dissemblables. Un défi pour l'apprentissage par transfert est donc de développer des approches qui détectent et évitent le transfert négatif des connaissances utilisant très peu d'informations sur la tâche cible. Un cas particulier de ce type d'apprentissage est l'adaptation de domaine [**LI-1**]. C'est une situation où les tâches sources et cibles sont identiques mais dans des domaines différents. Bien que l'informatique quantique et l'apprentissage artificiel puissent sembler a priori très différents, ils ont déjà pu se rencontrer et interagir de multiples façons par le passé. L'apprentissage quantique est le domaine qui est né de l'intersection et des rencontres entre l'informatique quantique et l'apprentissage automatique. Il s'agit d'une discipline en émergence présentant de nombreuses perspectives pertinentes pour la science. Parallèlement aux autres thèmes de l'axe ADA, nous tentons de modifier les algorithmes traditionnels d'apprentissage automatique pour les reformuler en termes quantiques. Aussi, nous avons proposé des algorithmes pour les ordinateurs quantiques pour résoudre efficacement certains problèmes qui ne peuvent pas être traités par des ordinateurs conventionnels.

- **Théorie du transport optimal pour le clustering collaboratif, le co-clustering et l'apprentissage par transfert**

Clustering collaboratif : Le travail de recherche développé concerne le développement d'approches de clustering multi-modèles à partir de données distribuées. Nous avons travaillé sur deux volets principaux en apprentissage automatique multi-modèles. Le premier concerne le clustering multi-vues où nous visons à apprendre un modèle optimal global à partir des modèles locaux des différentes vues. Le second volet porte sur le clustering collaboratif où l'idée principale est de trouver un modèle d'échange de connaissances entre les différents collaborateurs afin d'améliorer leurs propres performances. Pour le clustering multi-vues, nous avons proposé deux approches basées sur la théorie du transport optimal [CI-61]. La première approche (CPA : Consensus Projection Approach) consiste à trouver un modèle de consensus à partir des modèles locaux, en projetant les distributions des différentes vues sur un espace global. Il s'agit ici d'un apprentissage coopératif. La seconde approche (CNR : Consensus with New Representation) vise à apprendre une nouvelle représentation consensuelle à partir de la représentation

locale des distributions. Dans le cas du clustering collaboratif, nous avons introduit une nouvelle approche basée sur la théorie du transport optimal (Co-OT : Collaborative clustering through Optimal Transport) qui vise à améliorer le mécanisme de la collaboration et la manière de transporter les informations entre les collaborateurs avec un coût minimum [CI-75]. Pour ce faire, nous avons proposé une fonction objective pour la collaboration basée sur la distance de Wasserstein. Nous avons aussi proposé une solution pour choisir les meilleurs collaborateurs par comparaison de la distribution locale des prototypes et l'analyse de la diversité entre les collaborateurs. Pour une autre approche dans ce cadre, nous avons proposé un nouveau modèle de collaboration guidé par la sélection des caractéristiques, où l'idée principale est de choisir les caractéristiques qui donnent la meilleure représentation pour chaque collaborateur et garantissant la meilleure communication entre eux, tout en préservant la confidentialité des données de chaque collaborateur [CI-74]. Cette dernière approche collaborative est aussi développée dans le cadre de la théorie du transport optimal.

Co-clustering : Nous avons proposé une nouvelle méthode de co-clustering, une approche d'apprentissage non supervisé qui vise à découvrir des groupes homogènes d'instances de données et de caractéristiques en les regroupant simultanément [**CI-14**]. La méthode proposée utilise le transport optimal régularisé par l'entropie entre les mesures empiriques définies sur les instances de données et les caractéristiques afin d'obtenir une fonction de densité de probabilité conjointe représentée par la matrice de couplage optimale. Cette matrice est ensuite factorisée pour obtenir les partitions de lignes et de colonnes induites en utilisant l'approche des représentations multi-échelles. Pour justifier notre méthode sur le plan théorique, nous montrons comment la solution du transport optimal régularisé peut être vue du point de vue de l'inférence variationnelle, ce qui motive son utilisation pour le co-clustering. L'algorithme dérivé de la méthode proposée et sa version à base de noyaux utilisant la notion de distance de Gromov-Wasserstein sont rapides, précis et peuvent déterminer automatiquement le nombre de clusters en ligne et en colonne. Ces caractéristiques sont démontrées de façon éclatante par des évaluations expérimentales approfondies.

Apprentissage par transfert et adaptation de domaine avec garanties théoriques : Nous avons proposé une approche consacrée au transport optimal hiérarchique pour l'adaptation de domaine [CI-107]. Elle apporte quelques réponses à la question de l'apprentissage des structures cibles cachées en utilisant le transport optimal, et leur incorporation ensuite dans le processus du transport pour résoudre le problème d'adaptation de domaine. Il s'agit d'une formulation structurelle du transport optimal qui exploite, au-delà des informations géométriques capturées par la métrique de base, des informations structurelles plus riches dans les domaines source et cible. L'information structurelle dans le domaine source étiqueté est formée instinctivement en regroupant les échantillons dans des structures en fonction de leurs étiquettes de classe. L'exploitation des structures cachées dans le domaine cible non étiqueté est réduite au problème d'apprentissage des mesures de probabilité à travers le barycentre de *Wasserstein*, que nous prouvons théoriquement être équivalent au *clustering spectral*. Pour l'approche proposée, nous avons développé un nouveau cadre théorique d'adaptation de domaine à travers le transport optimal hiérarchique [**CI-14**]. Ce paradigme fournit des garanties théoriques sous la forme de bornes de généralisation en permettant de considérer l'organisation structurelle implicite des échantillons dans les deux domaines en classes ou *clusters*. De plus, nous avons fourni une nouvelle mesure de divergence entre les domaines source et cible, appelée la distance de *Wasserstein* Hiérarchique, qui indique, sous des hypothèses modérées, quelles structures doivent être alignées pour mener à une adaptation réussie.

- **Apprentissage quantique à partir de données séquentielles**

Dans ce thème, notre objectif est de fournir des solutions moins complexes en utilisant des algorithmes d'apprentissage automatique quantiques pour traiter des problèmes, notamment dans le domaine des études Probabilistes de Sécurité (*PSA : Probabilistic Safety Assessment*) avec un temps raisonnable. Notre travail a été guidé par les besoins et la complexité du problème traité. Nous avons abordé les deux aspects du problème des PSA, statique et dynamique. Pour le cas statique, nous nous sommes intéressés à trouver toutes les combinaisons d'événements de base du système qui peuvent générer des accidents graves, nous avons proposé d'obtenir ces combinaisons d'événements de base par un algorithme quantique, en utilisant des graphes orientés au lieu

de chercher toutes les solutions d'un problème SAT. Notre proposition est un algorithme quantique qui utilise un nombre linéaire de $qubits$ et grâce à un filtre classique, nous pouvons trouver toutes les combinaisons d'événements de base qui peuvent générer ces accidents [CI-96]. Dans le cas dynamique, nous nous sommes intéressés à trouver toutes les séquences accidentelles d'un système. Notre principal intérêt est le traitement de ces séquences. Dans le cas classique, afin de trouver toutes ces séquences, nous utilisons le graphe d'état du système et nous cherchons tous les chemins entre l'état courant et tous les états critiques. Comme ce problème est NP-complet, nous avons proposé une solution quantique pour trouver tous ces chemins. Nous avons proposé deux algorithmes quantiques, tous les deux basés sur la philosophie des marches quantiques. Le premier algorithme permet de trouver tous les chemins entre un sommet source et plusieurs sommets destination dans un graphe orienté acyclique. Cet algorithme utilise $Nqubits$ et M portes afin de trouver tous les chemins. Le second est une version hybride du premier, capable de traiter de grands graphes même avec un nombre réduit de $qubits$. Une autre contribution propose une approche quantique de l'algorithme *Dynamic Time Warping (DTW)* afin de calculer la similarité entre ces séquences, ainsi qu'une version capable de trouver la meilleure correspondance entre les séquences en utilisant des sous-séquences dont la longueur varie dynamiquement. Nous avons proposé également une stratégie d'apprentissage pour les Modèles de *Markov Cachés Quantiques (QHMM)* afin de générer des scénarios accidentels à partir de n'importe quel état initial du système et également pour gérer le système en temps réel [CI-96]. Nous avons proposé enfin une version améliorée de *k-means* quantique [CI-73].

1.2.2 Axe ARG : Apprentissage Relationnel et Graphes

Cet axe rassemble des recherches dans le domaine de l'apprentissage symbolique supervisé (apprentissage de règles dans des langages logiques, notamment en Datalog) ou non supervisé (fouille de données, et notamment fouille de graphes). C'est un thème « historique » de l'équipe A³ qui s'est retrouvé vivifié par l'intérêt récent dans la communauté d'IA pour l'explication de modèles d'apprentissage.

- **Fouille de graphes attribués**

Cet axe rassemble nos travaux portant sur le développement d'algorithmes d'extraction de connaissances dans les graphes. Ils s'inscrivent dans la suite de nos travaux de formalisation théorique et de développement d'algorithmes efficaces pour l'analyse de réseaux attribués et de graphes biologiques, et dans le développement de l'approche de fouille de motifs *fermés* dans les graphes *attribués*. Cette approche s'appuie sur le développement d'opérateurs d'*abstraction de graphe*. L'extension d'un motif d'attribut (classiquement un sous-ensemble d'items) est constituée de l'ensemble des sommets qui sont des occurrences de ce motif. En appliquant un opérateur *d'abstraction* on réduit cet ensemble à son cœur, c'est-à-dire la partie dont le sous-graphe induit vérifie une certaine propriété topologique. Un motif fermé abstrait est alors le plus spécifique parmi ceux partageant la même extension abstraite.

Nous inspirant des travaux de J. Kleinberg sur l'importance relative des sommets selon leurs rôles nous avons d'abord introduit la définition d'un nouveau cœur de graphe (*Hub-Autorité*), spécifique des graphes orientés, dans lequel un sommet est retenu en raison de son degré sortant suffisamment grand (il a alors un rôle de *hub*) ou en raison d'un degré entrant suffisamment grand (il a alors un rôle *d'autorité*) dans le sous-graphe induit. Nous avons alors étendu notre algorithme d'énumération de motifs fermés abstraits dans ce cadre. Lorsque les sommets peuvent ainsi avoir plusieurs rôles dans un sous-graphe, il est intéressant de considérer des bi-motifs, constitué de deux composantes chacune associée à un rôle dans le sous-graphe. Nous avons ainsi défini de nouveaux sous-graphes cœurs, défini les bi-motifs fermés abstraits associés, et étendu notre algorithme d'énumération à ces bi-motifs. Ceux-ci peuvent être définis et utilisés pour analyser des graphes orientés ou des graphes bi-modaux, graphes bi-partis liant deux types de nœuds comme par exemple auteurs et articles dans des données bibliographiques, ou encore pour faire apparaître des structures cœur/périphérie dans des graphes non-orientés [CO-7, CO-11].

Nous nous sommes également intéressés à la sélection d'un sous-ensemble de ces motifs fermés abstraits d'abord en considérant les motifs individuellement. La *modularité locale* d'un sous-graphe, issue des travaux sur la détection de communautés dans les graphes, peut-être utilisée pour sélectionner des motifs intéressants au sens où ils ont plus d'arêtes internes qu'attendu. Une telle contrainte, non anti-monotone, avait auparavant été introduite dans l'énumération de motifs via un estimateur anti-monotone optimiste. Pendant la visite de M. Atzmueller dans le laboratoire

nous avons appliqué cette idée à notre algorithme d'énumération de motifs fermés abstraits, évitant ainsi une coûteuse étape de sélection des motifs de modularité locale suffisante postérieure à l'énumération[TU-1, CI-54].

Cependant, si notre approche vise à minimiser le nombre de motifs énumérés en ne considérant que les motifs fermés supportés par un cœur de graphe, et en sélectionnant des motifs individuellement intéressants pendant l'énumération, il reste des situations où leur nombre est très important. C'est notamment le cas avec les bi-motifs dont la contrepartie au gain d'expressivité est une explosion combinatoire du nombre de motifs. C'est une question omniprésente en fouille de données : peut on sélectionner un nombre suffisamment restreint de motifs pour qu'ils soient présentés aux utilisateurs et experts du problème abordé ? Nous nous sommes intéressés à la sélection de sous-ensembles de motifs respectant collectivement une contrainte de diversité et avons développé un algorithme glouton de sélection d'un sous-ensemble de motifs divers [CI-4].

Nous avons appliqué et adapté ces méthodes à l'exploration de plusieurs graphes de terrain. En collaboration avec le sociologue E. Lazega nous avons exploré les relations de conseils entre pairs dans un réseaux professionnel d'avocats sous l'angle des hubs et des autorités [CI-4] et avec S. Bary nous avons étudié un jeu de données épistémologiques. La densité du graphe construit par S. Bary nous a conduit en particulier à proposer une méthode de partitionnement des arêtes, permettant notamment de faciliter l'exploration et la visualisation de graphes denses [CI-27]. La fouille de bi-motifs a été développée lors de l'étude de ces deux domaines, pour rendre compte de la diversité des structures de conseils en ce qui concerne les avocats (ceux qui demandent un avis n'ont pas le même profil que ceux à qui on le demande) et pour examiner les liens entre campagnes et participants dans l'étude épistémologique de campagnes d'exploration des fonds sous-marins.

Dans le cadre d'un projet international CHIST-ERA, nous avons, par l'analyse de réseaux de co-régulateurs de l'expression des gènes de la levure, identifié des motifs dont les gènes formant l'ensemble support sont des candidats pertinents à l'expérimentation biologique pour l'évaluation de leur implication dans la transition diauxic [**RI-11**]. En collaboration avec l'équipe RCLN du LIPN, dans le cadre du projet FUI-PCU, nous avons exploré les réseaux de collaborations scientifiques pour identifier des réseaux d'experts et d'expertises [CO-19], [CO-33]. Dans ce contexte nous avons été conduits à développer une approche permettant d'imposer aux deux parties d'un bi-motif de partager une description commune sur une partie du langage de description [CI-21].

Pour finir, nous avons profité du post-doc de Tiphaine Viard pour généraliser notre approche de fouille de motifs aux flux de liens. Ceux-ci offrent une représentation de l'évolution temporelle des interactions entre individus là où les graphes ne peuvent que capturer les interactions à un instant donné. Nous avons développé pour cela un cadre d'énumération, un algorithme et une implémentation de la fouille de motifs dans les flux de liens attribués [CI-112].

• Programmation Logique Inductive

La Programmation Logique Inductive (PLI dans la suite) est un domaine de l'apprentissage symbolique qui s'intéresse à l'apprentissage, le plus souvent supervisé, de modèles exprimés dans des restrictions de la logique des prédicats, comme Datalog. Les travaux réalisés autour de la Programmation Logique Inductive se sont poursuivis et développés pendant cette période, notamment en collaboration avec *Nukkai* et dans le cadre de la délégation de H. Soldano (2019-2021), de la demi-délégation de C. Rouveirol (2019-2020).

Exemples ambigus Cette recherche fait suite au cadre *d'apprentissage relationnel à partir d'exemples ambigus* que nous avions conçu, où chaque exemple est un ensemble de descriptions $\{x_1, \cdots, x_n\}$, une seule description, non déterminée, représentant la réalité. Nous appelons cette description l'*exemple caché*. Nous avions montré que pour tenir compte de l'ambiguïté des exemples il fallait remplacer une classique relation de couverture entre une hypothèse h et un exemple e, positif ou négatif par deux relations dites de *compatibilité* : l'une pour les exemples positifs, h *compatible*$^+$ avec e si h *couvre* au moins une description de e, l'autre pour les exemples négatifs, h *compatible*$^-$ avec e si h *ne couvre pas* au moins une description de e.

Suivant ce cadre, nous avons implémenté un système nommé *Lear*, où les descriptions d'un exemple sont les modèles d'une théorie clausale composée de clauses générales, les hypothèses des clauses de Horn et la relation *couvre* la théta-subsumption. *Lear* cherche une hypothèse *compatible*$^+$ avec les exemples positifs et *compatible*$^-$ avec les exemples négatifs.

Nous avons poursuivi cette recherche en associant à chaque description d'un exemple ambigu la probabilité d'être l'exemple caché. Ainsi nous pouvons définir la probabilité qu'une hypothèse

h couvre l'exemple caché dans e : c'est la probabilité que l'exemple caché soit dans les possibilités de e couvertes par h. Les relations de compatibilité ne tenant pas compte des probabilités, nous utilisons un critère de *justesse* pour guider la recherche d'une solution : la justesse d'une hypothèse h est l'espérance de la probabilité que h classe correctement un exemple e tiré au hasard, c'est-à-dire couvre l'exemple caché dans e si e est un exemple positif, et ne le couvre pas si e est un exemple négatif. Le but de l'apprentissage est de trouver une hypothèse de justesse maximale.

Nous avons conçu un nouveau système, nommé *nukLear*[1], qui adopte la représentation des hypothèses et exemples de *Lear*, mais associe des probabilités aux possibilités et utilise le critère de justesse pour chercher une hypothèse solution.

ILP et explications Nous avons étudié, également en collaboration avec *Nukkai*, l'utilisation de systèmes d'apprentissage supervisé de PLI dans le cadre d'une boucle d'interaction avec des experts de Bridge, dans le but de produire des explications associées à une prise de décision. Dans un premier temps, nous l'avons étudié une méthodologie pour résoudre un problème relativement simple de Bridge (décision d'ouverture dans le cas de mains « limites ») [CI-32], puis à un problème plus complexe d'enchères [CO-19] où les classes des exemples ont été attribuées par un processus boîte noire utilisant notamment des simulations. Dans les deux cas, des modèles à la fois explicables et précis ont été appris, notamment grâce à une interaction avec des experts de Bridge et au fait que les modèles produits par PLI soient explicites.

D'autres travaux en cours dans le cadre de la thèse CIFRE de M. Kazi Aoual abordent le problème de l'explication d'une décision dans un cadre *MDP* : étant donné un état initial, une séquence d'actions est effectuée et la récompense associée à la séquence d'actions est perçue dans l'état final. Dans le cadre du jeu de Bridge, les actions sont des cartes jouées, l'état initial est une donne ; un joueur (déclarant) joue contre un adversaire (défenseur) dont on connaît *a posteriori* les cartes, le but est d'expliquer quelles décisions (cartes jouées) ont mené à un gain de la partie, et pourquoi. Nous avons choisi de représenter les trajectoires (séquences d'actions) comme des interprétations, décrivant à la fois les actions effectuées, mais également les propriétés des états atteints par ces trajectoires, certaines de ces propriétés perdurant sur des intervalles de temps le long de la trajectoire. Etant donné un groupe T^+ de trajectoires optimales et un groupe T^- de trajectoires non optimales, nous avons conçu un algorithme permettant de construire des explications *pour ce groupe*, comme les sous-ensembles minimaux et corrects de la généralisation maximalement spécifique de T^+. Une implémentation est proposée permettant de résoudre les problèmes de complexité de ces étapes pour des langages relationnels (Datalog).

Les jeux sont un terrain d'applications privilégié pour identifier et expliquer des trajectoires optimales dans un cadre *MDP*, mais ces questions peuvent être transférées à d'autres problèmes réels. Par exemple, nous sous sommes également intéressés à des applications d'enseignement numérique. Pour ce travail, un exercice d'algorithmique est posé à un étudiant, qui va le résoudre progressivement en appliquant des actions prédéfinies, et où le problème est d'identifier à des fins de recommandations et d'explication pourquoi une séquence d'actions est optimale pour résoudre le problème. Ce travail, très exploratoire, a été initié dans le cadre d'un BQR financé par l'Université Sorbonne Paris Nord, qui a permis de tester une première représentation des traces de résolution d'exercices et d'évaluer les principales difficultés pour apprendre une politique définissant quelle action doit être appliquée dans un état – une solution partielle à l'exercice – donné par des techniques d'apprentissage par renforcement classique.

Enfin, dans le cadre d'une collaboration avec l'équipe RCLN, nous sous sommes intéressés à l'application de l'apprentissage relationnel à l'apprentissage de règles discriminantes dans le cadre de données graphes [CI-24]. Ces travaux ont permis d'évaluer les avantages et limites de chacune des méthodes en présence (arbre de décision relationnel et fouille de règles relationnelles « à la AMIE » sur le problème donné et d'initier une collaboration sur l'apprentissage de règles en présence de très peu d'exemples étiquetés et d'une théorie du domaine de grande taille représentée sous forme de graphe de connaissances.

1.2.3 Axe MAR : Métaapprentissage et Apprentissage de Structures

Cet axe regroupe les activités en métaapprentissage et en modélisation simultanée des données et du clustering (apprentissage non supervisé profond) à large échelle appliqués aux données temporelles multivariées, aux images et aux documents audios et vidéos.

1. *NukLear* est développé par *Nukkai*, dans le cadre d'une collaboration avec le LIPN.

Le metaapprentissage comprend l'ensemble des processus qui permettent d'apprendre à apprendre. Il inclut, en particulier, l'analyse des algorithmes d'apprentissage pour sélectionner le plus approprié en optimisant l'efficacité et l'efficience, la conciliation, la composition ou le renforcement de modèles de solutions issus de perspectives différentes, le transfert partiel ou total et le contrôle du changement d'espaces de représentations, le choix des fonctions d'évaluation et la prise en compte de toutes les formes de connaissances a priori et leur relativité, en plus des données, pour améliorer l'apprentissage courant et futur en minimisant l'effort, le temps et les ressources.

Mieux apprendre ne s'évalue pas uniquement en fonction des critères d'efficacité sur les données du problème spécifique, mais aussi sur l'efficience de la solution globale incluant, par exemple, l'évaluation des critères tels que le respect de l'éthique, le respect de la loi, le respect de la vie privée, la qualité de l'explication ou de l'interprétation, l'effort en ressources consommées (facteur écologique, etc.) ou le nombre d'expériences nécessaires pour apprendre.

Un modèle intégré unique, basé souvent sur des modèles de fonctions non linéaires complexes optimisées par des biais de choix des hypothèses minimisant des mesures générales ne suffisent pas pour répondre aux autres exigences, qui s'évaluent sur d'autres critères souvent impossibles à exprimer dans ces modèles. Ceci nécessite le développement de nouveaux paradigmes d'apprentissage multimodèles chacun responsable de certifier un sous-ensemble de fonctions d'évaluations du système global. Notre définition du métaapprentissage c'est aussi la proposition de nouvelles architectures de conception capablent d'auditer, de qualifier et plus généralement d'évaluer les modèles d'apprentissage selon des normes préétablies par de larges communautés, dépassant celles des experts en intelligence artificielle.

Cet axe regroupe les recherches qui tentent de proposer des solutions et modèles génériques pour ces problèmes à travers l'exploration de divers paradigmes complémentaires pour améliorer les performances avec moins d'exemples en tenant compte de la structure explicite ou non, en recherchant des régularités par la minimisation du risque structurel notamment, en modélisant le contexte et les divers biais, en explicitant la nature des données (car les données n'ont pas toutes le même impact, ni les mêmes caractéristiques -véracité, certitude, imprécision, etc.-) [RI-18] et en modélisant les interactions entre les concepts à apprendre et concepts appris selon divers points de vue. C'est aussi tenir compte de l'évolution des théories apprises et de la métamodélisation des hyperparamètres. L'objectif ultime est de construire un cadre général et générique pour le processus d'apprentissage lui-même. Nous détaillons ci-dessous les principales contributions.

Intégration de biais de contraintes structurelles Durant cette période exploratoire de l'impact du métaapprentissage (MML) et de la structuration du risque structurel sur la qualité de l'apprentissage, nous avons choisi de travailler essentiellement sur les applications issues de l'internet des objets [**CI-81**]. Ce choix, très ciblé, répond à plusieurs problématiques fondamentales du MML dont la présence de plusieurs points de vue sur les phénomènes d'intérêt (relativité), la coexistence de plusieurs théories cohérentes, mais d'apparence incompatibles (apprentissage fédéré, distribué, etc.), la multiplicité des modalités peu étudiées ensemble (son, image et les divers types de capteurs), existence de modèles équationnels qui régissent une partie des phénomènes observés, le mode de disponibilité des données (souvent par fenêtres temporelles), leur hétérogénéité et leurs distributions, etc. qui s'ajoute aux problématiques habituelles liées au choix de meilleures hypothèses qui représentent au mieux les exemples d'apprentissage. Cette famille d'applications impose de nouveaux paradigmes d'apprentissage automatique qui doivent inclure la modélisation à la fois des données, des structures des espaces d'hypothèses, des fonctions d'évaluation, de l'adaptation, du renforcement, du transfert, de l'explicabilité, de l'interprétabilité des résultats et des décisions, de l'incrémentalité de l'apprentissage, de la quantification des méthodes utilisées pour apprendre en plus de la quantification liée aux données, du respect de diverses normes et contraintes, etc. Nous avons proposé plusieurs approches basées sur les biais de contraintes structurelles et sur la structuration du processus d'apprentissage. La première catégorie de travaux porte sur la prise en compte de différentes formes de structures du domaine et leur intégration dans le processus d'apprentissage [**CI-52** , **CI-86**, **CI-80** , 1]. Alors que la deuxième utilise les connaissances et la sémantique de l'espace des catégories pour organiser le processus d'apprentissage [**CI-81**, **CI-97**]. L'objectif est la fois la réduction du nombre d'exemples nécessaires pour apprendre et la recherche de modèle de prise en compte de ce type de contraintes dans une modélisation générique du MML.

Abstraction du contexte et relativité Les données réelles qu'on manipule sont essentiellement issues de processus de transformation via des capteurs en utilisant des fonctions de

transferts. Des biais intrinsèques entachent les données par ce simple fait [**CI-81**]. De plus, l'emplacement des sources de données apportent des biais de point de vue dont la connaissance améliore sensiblement la qualité de l'apprentissage. Nous avons travaillé sur les mécanismes de conciliation des apprenants par l'abstraction de contexte et également, quand la connaissance du domaine le permet, la modélisation explicite ou implicite des données relatives afin de distinguer les données qui dépendent du contexte local des données générales utilisables dans un contexte de centralisation de la théorie de l'apprentissage [**CI-98**, **CI-105**].

Quantification des biais Quantifier les biais consiste à prendre en compte tout le pipeline d'apprentissage, depuis la modélisation de l'environnement pour définir la typologie puis la topologie des capteurs des données, à l'acquisition de l'information via les fonctions de transfert, à la prise en compte des biais dans les processus de transformation et de transmission des données et au choix des fonctions d'évaluation. Dans nos derniers travaux, en plus de ceux plus théoriques, nous avons exploré l'impact de ces biais dans diverses applications [TU-3] sur le biais des modèles d'évaluation ou l'article [TU-3] sur l'étude de la segmentation et de l'évaluation comme problèmes eux-mêmes à apprendre. L'outil développé est disponible ici : (accéder au Code). ou la redéfinition de la segmentation comme un problème de décomposition nécessitant de quantifier la perte d'information issue de la segmentation, l'apprentissage après segmentation et le lien qui existe entre la solution du problème d'apprentissage sur le modèle segmenté et le problème initial [CI-101]. L'outil développé est accessible ici (accéder au Code) et ici pour l'évaluation de la décomposition (Code) [**CI-99**]. Dans le domaine de l'analyse du son, l'étude du problème d'évaluation d'évènements sonores selon l'angle du métaapprentissage, nous a permis de présenter une méthode multimodale simple et visuelle pour l'évaluation des évènements sonores [**CI-100**] (accéder au Code) qui permet d'analyser les biais liés au choix des modalités et les résultats obtenus).

Multi-Coclustering topologique et visualisation En apprentissage automatique, les modèles de mélanges topologiques et leurs variantes stochastiques sont très présents en statistique ; plus généralement, l'hypothèse markovienne est souvent invoquée à chaque fois que des probabilités sont utilisées pour modéliser l'état d'un processus dynamique. Ces modèles de Markov sont souvent combinés avec d'autres algorithmes de classification. Nous avons étudié la possibilité d'utiliser conjointement les modèles de mélange et/ou les modèles topologiques pour traiter des séries temporelles multivariées (e.g., fonctionnelles). Cela consiste à estimer automatiquement et simultanément le modèle, la structure cachée ainsi que le nombre de classes, supposé lui aussi inconnu. Les mélanges de lois permettent de modéliser des distributions de probabilités d'une manière à la fois flexible (une grande diversité de distributions peuvent être modélisées sous cette forme) et parcimonieuse (la distribution ainsi modélisée est décrite par un nombre réduit de paramètres) [CI-87]. Parmi les résultats obtenus, nous citons le développement d'un modèle de Multi-Coclustering non paramétrique. Ces travaux ont été valorisés dans le cadre d'applications concrètes dans le contexte d'un projet $CIFRE$ avec $Renault$.
Pour améliorer les résultats du clustering, nous avons également abordé d'autres travaux liés aux changements de représentations tel que le travail sur le regroupement de sous-espaces (subspace clustering), nous définissons un double système de pondération pour détecter les attributs et les sous-espaces d'attributs "influenceurs" dans un clustering. Nous avons étudié deux types de pondération, la pondération globale dans laquelle un seul vecteur de pondération est attribué pour chaque attribut et sous-espace, et la pondération locale où un vecteur de pondération différent est attribué pour chaque attribut et sous-espace d'un cluster [CI-93].
Dans la même direction, nous nous sommes également intéressés, au clustering multiobjectifs en utilisant des méthodes évolutionnaires. Contrairement aux méthodes classiques qui détectent les clusters en se basant sur une seule mesure de similarité, les approches évolutionnaires élargissent l'espace de recherche et permettent de détecter les clusters via des fonctions objectifs et un algorithme génétique [CI-93]

Apprentissage profond et Clustering Les succès récents des architectures profondes pour le clustering (unsupervised deep learning) nous a conduit à développer un modèle qui apprend conjointement la représentation des données et la représentation des référents des cartes topologiques (SOM :Self-Organizing Map). Notre modèle est composé d'un autoencoder et une couche SOM optimisée lors d'une procédure d'apprentissage conjointe (DESOM : Deep Embedded Self-Organizing Map) [CI-57, CI-35]. les résultats obtenus dans le cadre de la $CIFRE$ avec $Safran$ sur les séries temporelles nous ont confronté dans l'intérêt de coupler l'apprentissage profond aux méthodes d'apprentissage non supervisé afin d'encoder/vectoriser des données

brutes non formatées (par exemple les séries temporelles, les images, etc.). Les données disponibles pour le clustering sont issues de sources hétérogènes, ces problèmes d'encodage multimodales constituent une dimension intéressante de recherche qui reste encore à explorer.
De plus, la considération de la contrainte de scalabilité des modèles nous a conduit à la proposition d'un nouvel algorithme utilisant un système de hachage (LSH : Locality Sensitive Hashing) tout en bénéficiant du récent paradigme MapReduce appliqué aux systèmes distribués. Nous avons proposé également la variante de l'algorithme de clustering pour les données catégorielles en utilisant le codage binaire et la distance de Hamming [CI-45], aussi nous avons introduit une amélioration du Clusterwise en combinant l'algorithme de clustering et un modèle de régression [RI-26]. L'amélioration de sa complexité en temps, en appliquant le clustering avant une tâche de régression PLS, a permis d'assurer le passage à l'échelle dans le domaine de la sélection de variables. Nos résultats sont mis en œuvre et testés dans le cadre d'applications réelles et ont conduit au projet opensource Clustering4Ever et une collaboration CIFRE avec Kameleoon.

Références complémentaires

[1] M. Hamidi and A. Osmani. Description of Structural Biases and Associated Data in Sensor-Rich Environments. *CoRR*, abs/2104.04885, 2023.

1.3 Profil d'activités liées à la recherche

Activités (répartir 100 points sur ces 7 items)	
Administration et animation de la recherche : pilotage de la recherche (VP, direction d'institut, DAS, par exemple), participation à des instances d'évaluation (CNU, CoNRS, CSS, Hcéres, par exemple), responsabilité de dispositifs Idex ou Isite, direction de projets (ANR, Horizon Europe, ERC, CPER, PIA, France 2030, par exemple), responsabilités éditoriales dans des revues ou collections nationales et internationales.	10
Aide aux politiques publiques et expertise technique : pouvoirs publics aux niveaux européen, national et régional, entreprises, instances internationales comme FAO, OMS, etc.	5
Contribution à l'adossement d'enseignements innovants à la recherche : EUR, SFRI, etc.	5
Dissémination de la recherche : partage de connaissances avec le grand public, médiation scientifique, interface sciences et société	5
Recherche et encadrement de la recherche.	50
Valorisation, transfert, innovation.	20
Autres activités. (à préciser en une ligne maximum).	5

1.4 Prise en compte des recommandations du précédent rapport

Lors de la précédente évaluation avait jugé que : *"le contexte, favorable à l'IA et à l'apprentissage automatique, devrait permettre à l'équipe de gagner en visibilité et attractivité"... "Les productions sont de très bon, d'excellent niveau, pourraient être renforcées par la considération de conférences plus centrales de l'apprentissage automatique, (ECML/PKDD) ou (NIPS). L'équipe doit se poser la question des conférences qu'elle souhaite cibler... pour renforcer son attractivité et sa visibilité internationale."*
Les dernières années le contexte est devenu encore plus favorable à l'IA et à l'apprentissage automatique, au point que les postes universitaires deviennent largement moins compétitifs que ceux proposés par les grands groupes internationaux. L'équipe a perdu plusieurs de ses membres depuis 2017 et n'a eu aucun recrutement. Malgré ce contexte, la production de l'équipe reste soutenue et plusieurs papiers sont acceptés dans des conférences centrales de l'apprentissage automatique comme ECML/PKDD, PAKDD, NIPS, IJCAI ou AAAI.

Le comité avait également indiqué que *"les thématiques mises en avant par l'équipe sont parfois peu "lisibles" et gagneraient à être simplifiées.* De plus, le comité a indiqué que l'équipe

jouit d'une *Excellente production scientifique, avec des publications dans des conférences et des journaux de tout premier plan. Elle a une visibilité et une attractivité fortes au niveau national et participe de façon importante à l'animation de la communauté apprentissage en France. L'attractivité et la présence de l'équipe à l'international peuvent toutefois être renforcées.*

Un effort a été fait pour structurer les activités de l'équipe en passant de quatre axes à trois axes avec un affichage beaucoup plus clair et une identité renforcée. Malgré l'amélioration des collaborations à l'international avec, notamment, plusieurs (8) séjours dans des laboratoires étrangers par des membres de l'équipe, la participation ou l'organisation à plusieurs conférences et workshops internationaux, l'équipe doit fournir des efforts supplémentaires pour améliorer sa présence et sa visibilité à l'international.

Le comité a indiqué que *la vie de l'équipe mérite d'être renforcée avec des activités plus régulières permettant de mieux partager l'information et d'impliquer tous ses membres. Il est important de faire en sorte que tous les membres participent à la production et aux collaborations de l'équipe. L'équipe devra aussi être attentive dans ses recrutements futurs et veiller à un bon équilibre F/H... le ratio rang A/rang B, extrêmement défavorable*, et les responsabilités administratives importantes assurées par plusieurs membres seniors représentent toutefois des facteurs de risque pour le futur. Il est important que l'équipe A3 puisse s'appuyer sur plus de membres de rang A et se "décharge" de certaines responsabilités administratives actuelles.

Plusieurs membres de l'équipe ont été et sont encore très impliqués dans des responsabilités administratives (VP USPN, directrice adjointe de l'IG en charge de l'école d'ingénieurs, directeurs de département d'IUT, chargé de mission auprès de la présidence). Ce facteur et l'absence de recrutement de membres de rang A constituent un facteur de risque pour l'équipe.

L'équilibre homme/femme s'est amélioré du fait du départ de plusieurs membres et grâce au soin pris à recruter des doctorantes (sur la période, il y a eu 11 doctorantes sur un total de 23 doctorant.e.s). L'équipe compte une professeure et un professeur et une maîtresse de conférences contre neuf maîtres de conférences.

Chapitre 2
Introduction du portfolio

Le portfolio qui suit met en exergue une partie de l'expertise de l'équipe à travers cinq réalisations :

Transport optimal Des avancées récentes en adaptation du domaine sont dues à la théorie du transport optimal. Nous présentons nos résultats publiés en 2022 dans *Machine Learning*.

Données relatives et apprentissage de représentation Le métaapprentissage s'intéresse à la structuration et à l'apprentissage de tous les biais permettant d'apprendre à apprendre. Il s'agit ici de décrire une série de résultats [**CI-105**, **CI-97**, **CI-99**] qui présentent deux biais importants et leurs apports pour l'apprentissage : la relativité des données et leur impact sur le succès de l'apprentissage et l'utilisation d'invariants pour la structuration de l'espace de représentation des concepts à apprendre.

Un livre sur la théorie de l'adaptation du domaine écrit en collaboration avec un enseignant chercheur de l'équipe, ce livre présente les résultats les plus récents en apprentissage par transfert.

La startup HelphIA Les collaborations industrielles de membres de l'équipe ont conduit au développement, en 2020, d'un projet de création d'une startup à succès mise à l'honneur dans ce portfolio.

Apprentissage de modèles structurés et projet CHIST-ERA Adalab Des membres de l'équipe ont participé à ce projet européen Dont l'objectif était d'accélérer les découvertes scientifiques grâce à la conception et au développement de robots scientifiques autonomes capables de mettre à jour un modèle, de réaliser et de choisir des expériences plus rapidement et de manière fiable [**RI-11**]. Les bons résultats obtenus, dans le cadre de ce projet, sont détaillés dans ce portfolio.

Chapitre 3
Auto-évaluation de l'équipe A3

Sommaire

3.1 Domaine 2. Attractivité 86

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche. 86

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels. 87

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs. 88

3.2 Domaine 3. Production scientifique 88

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité. 89

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels. 89

3.2.3 Publications 90

3.3 Domaine 4. Inscription des activités de recherche dans la société . 105

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique. 105

3.3.2 Référence 2. L'équipe développe des produits à destination du monde culturel, économique et social. 105

Auto-évaluation de l'équipe A3

3.1 Domaine 2. Attractivité

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche.

L'équipe est active dans sa communauté scientifique, à la fois en termes de publications de haut niveau, de participation aux comités de programme de conférences de premier plan ou d'organisation ou d'invitations à des conférences et workshops internationaux. La nouvelle structuration de l'équipe commence à donner ses fruits. A la fois au niveau théorique (transport optimal, métaaprentissage ...) et au niveau des applications de haut niveau.

Invitations dans des institutions académiques ou des congrès internationaux L'équipe a été invitée au Colloque international de l'UM6SS sur le thème *Artificial Intelligence for health care*, à la journée FHU MOSAIC' sur le thème *Multiscale Optimised Strategy for Artificial intelligence-based Imaging biomarkers in digestive Cancer*, à la conférence CIFSD'18-Tanger sur *Limits of Artificial Intelligence*, par le centre de conférences CURI-USMBA de Fès sur *Behind Deep Learning*. Elle a également participé au colloque *sciences et sens de l'intelligence artificielle* avec deux conférences invitées l'une sur la *Chirurgie augmentée par l'Intelligence Artificielle* et l'autre sur l'*Intelligence Artificielle : entre confiance et défiance* qui ont eu lieu au campus Condorcet.

Durant cette période, les membres de l'équipe ont effectué huit séjours dans des laboratoires étrangers. Il y a eu deux séjours en Italie, à Caserta, à l'Université Campania *Luigi Vanvitelli* et à Naples ; un séjour à l'université du Texas ; deux séjours au Maroc, à Tanger (ENSAT) et Fès (centre CURI) ; un séjour eu Vietnam (USTH) ; un séjour au Sénégal et deux séjours aux États-Unis (université de Californie et université du Texas).

Conférences où des membres de l'équipe sont membres du PC Chaque année, les membres de l'équipe participent aux comités de programmes de plusieurs conférences et journaux internationaux de premier plan. Parmi ces conférences, on trouve régulièrement : American Conference on Artificial Intelligence (AAAI), International Joint Conference on Artificial Intelligence (IJCAI), ICONIP (International Conference on Neural Information Processing), International Conference on Machine Learning (ICML), European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD), IJCNN (International Joint Conference on Neural Networks), Inductive Logic Programming conference (ILP), ISCV (International Conference on Intelligent Systems and Computer Vision).

Plusieurs membres de l'équipe participent aux comités de programmes et l'organisation de la recherche au niveau national. Ils participent aux comités de programmes des conférences tels que : Conférence Francophone d'Apprentissage (CAP), Conférence Internationale Francophone Extraction et Gestion des Connaissances (EGC), Rencontres annuelles de la Société Francophone de Classification (SFC), Apprentissage Artificiel et Fouille de Données (AAFD), French Regional Conference on Complex Systems (FRCC), APIA (Applications Pratiques de l'Intelligence Artificielle) sur la plateforme de l'AFIA (Association Française de l'Intelligence Artificielle).

Organisation de conférences Une année sur deux plusieurs l'équipe organise le workshop international ALML (International Workshop on Advances in Learning from/-

with Multiple Learners). Des membres de l'équipe ont participé à l'organisation de la 23ème conférence IV 2019 (23th International Conference Information Visualisation), de la conférence ICHS (International Conference 24 hours Holographic Surgery), de la conférence internationale ACONTA (International Conference on Augmented Complex Networks), du symposium intenational CS2 (IEEE International Symposium on Cloud and Services Computing), de la conférence EGC (Extraction et de Gestion des Connaissances), de la conférence CIFSD (Conférence Internationale Francophone sur la Science des données), de deux workshop LSISA (Workshop on Large-scale Industrial Time Series Analysis) en marge de la conference ICDM et du workshop international IWMAN (International workshop on Mining Attributed Networks).

Elle organise des sessions adossées à des conférences internationales comme la session *Applying Complex Network Analysis for Cybersecurity* à la conférence internationale EICC à deux reprises, la session *Data Science Approaches or Modeling, Analyzing and Mining Networks on Networks* à la conférence DSAA, la session *Applying Complex Network Analysis to Industrial Risk Studies* à la conférence Networks et la session *Representation learning for Transfer and Collaborative approaches* associée au Symposium IEEE SSCI (Series on Computational Intelligence).

L'équipe organise également régulièrement l'école thématique "Deep Learning & Data Science".

Responsabilités éditoriales dans des revues et des collections de haut niveau L'équipe a eu des activités de responsabilité éditoriales durant la période. Younès Bannani est *Associate Editor* du journal international Knowledge and Information Systems (Kais) édité par les éditions Springer et est *Deputy/Managing Editor* du journal Marocain *Pure and Applied Analysis* .Hanane Azzag et Mustapha Lebbah ont participé à l'édition des actes de la conférence internationale IV en 2019 et à ceux de la conférence EGC en 2018.

Participation à des instances de pilotage de la recherche et d'expertise scientifique Plusieurs membres de l'équipe ont participé à l'expertise de projets ANT, ANRT. Younès Bennani était notamment nommé expert sur des dossiers Masters pour l'HCERES et auprès du Pôle de Compétitivité Cap Digital. Rushed Kanawati était également nommé expert d'un projet pour la Czech Science Foundation en 2020. Mustapha Lebbah était nommé expert scientifique pour le ministère de l'enseignement, de la recherche et l'innovation et membre externe du conseil scientifique du centre de développement des technologies avancées en Algérie (CDTA). Hanane Azzag est membre du bureau de l'association EGC et membre de son comité de pilotage depuis 2019. Céline Rouveirol est membre du conseil d'administration de l'AFIA depuis 2020 et responsable du *collège Apprentissage Automatique*.

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels.

Le laboratoire a mis en place des dispositifs d'accompagnement des doctorants dont profite totalement l'équipe. L'équipe bénéficie des allocations de recherche de l'école doctorale réparties en favorisant les meilleurs dossiers et une équité entre les équipes. Par ailleurs, les membres de l'équipe disposent d'un bon réseau de partenaires industriels qui compte des entreprises comme SAFRAN, Renault, Meta, EDF, Orange, BPCE et Nukkai leur permettant de nombreux financements. Sur la période concernée par l'évaluation, le nombre de doctorants dans l'équipe est passé à 33 dont 14 thèses CIFRE (dont 18 thèses soutenues), en augmentation substantielle par rapport à la période précédente (23 doctorants dont 10 thèses CIFRE). Cette augmentation du nombre s'est également accompagnée de l'amélioration de la qualité du recrutement des doctorants. Tous les membres HDR de l'équipe participent également à des co-encadrements de

thèses en collaboration avec des membres des autres équipes (Calin, AOC, RCLN), avec les autres laboratoires de la fédération MathStic (L2TI et LAGA) et avec d'autres universités (UPEC, UVSQ, Cergy).
La durée moyenne des thèses est d'environ 41 mois contre 45 mois lors de la précédente période d'évaluation.

Auto-évaluation de l'équipe A3

L'équipe a accueilli également six post-doctorants entre 2017 et 2022. Deux dans le cadre du projet Européen ChistEra Adalab, une dans le cadre de la bourse Marie SKLODOWSKA-CURIE, une dans le cadre du projet FUI-PCU et un dans le cadre de la startup HelphIA.

L'équipe n'a accueilli qu'un ingénieur d'étude CNRS durant un an, financé dans le cadre du projet France relance ACOMI. Ce manque d'ingénieurs constitue une des faiblesses de l'équipe. Plusieurs développements de logiciels sont faits, plusieurs plateformes d'expérimentations sont développées sans arriver à leurs termes par manque de compétences techniques et permanentes dans l'équipe pour assurer le maintien, le suivi et la diffusion appropriée.

L'équipe compte également plusieurs membres associés qui sont au nombre de douze en janvier 2023.

Les membres de l'équipe se font inviter dans plusieurs laboratoires dans le monde. Il y a eu huit séjours durant cette période : en Italie (l'Université Campania "Luigi Vanvitelli", Caserta et à l'université de Naples), aux USA (université du Texas et université de Californie), au Vietnam (USTH), au Sénégal (AIMS) et au Maroc (Tanger et au centre CURI-USMBA de Fès). Sept chercheurs ont été invités dans l'équipe entre 2017 et 2022. Ils participent à l'animation de la vie scientifique et collaborent à divers projets.

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs.

Les membres de l'équipe participent à plusieurs projets nationaux. Ces projets représentent la source principale de financement des doctorants et post-doctorants.

Durant cette période, l'équipe a eu deux contrats campus France, un PROFAS et une bourse d'excellence de l'ambassade de France en Iran qui a permis le co-encadrement d'un doctorant. Elle a participé à deux projets Européen, un projet CHIST-ERA (Adalab) et un ERASMUS K107 pour un montant total de 243k€, un Labcom ANR IRISER (363 k€). Les membres de l'équipe ont participé à un contrat avec un laboratoire public du ministère de la défense pour un montant de 150k€ et un projet AMIES (99k€), un projet France Relance ACOMI (122k€), et un projet DIM4IDF (122k€), Un projet IUH, dont équipe est partenaire, est accepté en mai 2023. La source principale de financements des activités de l'équipe vient des projets de partenariats avec l'environnement socio-économique. Les membres de l'équipe ont participé à 19 contrats durant la période de référence pour un montant total de 2,13 millions d'euros. Il s'agit des contrats CIFRE MINDLYTIX 6k€), EDF-2020 (45k€), ALTER WAY (15k€), NUKKAI(20k€), SYNALTIC (28k€), RENAULT CIFRE (45k€), RENAULT PostDoc (56k€), BPCE (79k€), SAFRAN (57k€), REGIENOV (45k€), EDF (45k€), FaceBook (120k€), Deep Knowledge (120k€), SAFRAN (69k€) et Novedia (30k€). Elle a été sollicité pour une prestations de service pour un montant par REGIENOV pour un montant de 126k€ puis pour un montant de 122k€. Elle a, également, bénéficié d'un projet de maturation d'un montant de 470 k€. Le montant total des financements obtenus par l'équipe durant cette période d'évaluation avoisine les 3 millions d'euros.

3.2 Domaine 3. Production scientifique

Le tableau suivant résume la production scientifique de l'équipe durant la période.

	2017	2018	2019	2020	2021	2022	2023	*Total*
Livres	0	0	1	0	0	0	0	*1*
Chapitres de livre	2	0	2	0	1	2	0	*7*
Revues internationales	3	3	9	1	8	4	1	*29*
Autres revues	1	0	2	0	0	2	0	*5*
Conférences internationales	20	25	17	13	19	17	1	*112*
Autres communications	3	4	7	8	5	4	1	*32*
Total	*29*	*32*	*38*	*22*	*33*	*29*	*3*	*186*

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité.

La production scientifique est importante et de très grande qualité. Elle repose sur des fondements théoriques et méthodologiques solides et s'appuie sur des applications réelles nombreuses et de qualité. L'équipe travaille sur diverses thématiques allant de la programmation logique inductive, aux modèles théoriques d'apprentissage collaboratif et par transfert et aux mécanismes qui permettent d'apprendre pour apprendre. Ces thématiques, tout en utilisant les dernières avancées en apprentissage automatique, restent originales et constituent des orientations différenciantes au niveau régional et national.

Plusieurs membres de l'équipe sont impliqués dans des responsabilités administratives et d'enseignement chronophages qui limitent considérablement leurs implications en recherche durant de nombreuses années. Durant la période les membres de l'équipe ont pu publié plus d'une centaine d'articles dans des conférences internationales dont un article à IJCAI, deux articles AAAI, 2 articles ECML/PKDD, 2 articles PAKDD et une trentaine d'articles dans les revues internationales dont un article dans machine learning journal, un article dans transactions of the American Mathematical Society, un article dans les Proceeding National Academy of Sciences (USA) et un article dans Neurocomputing.

L'accent est moins mis sur les revues.

Les réalisations logicielles sont nombreuses, mais souvent mal mises en valeur. L'équipe gagnerait à mettre ses productions logicielles et ses démonstrateurs dans une plateforme visible et accessible. L'absence de personnel dédié dans l'équipe explique en partie cette situation.

Les diverses participations aux projets nationaux démontrent la vigueur et l'activité intense des membres de l'équipe. Cependant, il y a encore peu de participations aux projets internationaux.

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels.

Depuis la dernière évaluation, plusieurs membres sont partie en mutation ou en promotion dans d'autres universités fragilisant l'équipe et son équilibre. Dans ce contexte, l'arrivée de nouveaux collègues dynamiserait le potentiel de l'équipe. Il n'y a pas eu de recrutement sur la période, mais, mais un MCF est en cours de recrutement. Il est à noter que le recrutement de profils répondant à nos besoins est devenu très difficile à cause de la concurrence avec le secteur privé qui offre de meilleures conditions de travail et de meilleures opportunités.

L'accueil de doctorant est très actif et s'améliore en qualité et en production scientifique.

L'équivalent temps plein travaillé (ETP) de l'équipe est passé de 8,33 en 2017 à 6,83 en 2022. La moyenne annuelle sur la période est de 7,80.La production annuelle s'établit de ce fait à 26/46,03=0,56 revues par personne et par an et 111/46,03=2,4 conférences internationales par personne et par an. La production totale, incluant chapitres de livres et autres publications s'établit à 168/46,03=3,65 par personne et par an. Cette situation cache cependant, des disparités importantes entre membres de l'équipe. Il est à noter que les doctorants participent à la production scientifique.

Les restructurations internes de l'équipe n'ont pas permis l'établissement de plans d'accompagnement des enseignants chercheurs qui souhaitent reprendre leur activité de recherche. Il s'agit d'un challenge à relever dans les prochaines années. Cet accompagnement peut se faire, en définissant des projets fédérateurs intégrant les compétences des collègues souhaitant reprendre l'activité de recherche à la fois dans le montage des projets et l'encadrement de doctorants.

3.2.3 Publications

Les références
- comportant « * » sont acceptées et à paraître.
- indiquées sur fond gris sont les productions phare de chaque catégorie.
- en gras constituent les 20% les meilleures de la production scientifique de chaque catégorie.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2023 [RI-29] M. O. Attaoui, N. Dif, H. Azzag, M. Lebbah et K. Nabil. Regions of interest selection in histopathological images using subspace and multi-objective stream clustering. *The Visual Computer*, 39 :1683–1701, 2023.

2022 [RI-28] K. Benlamine, Y. Bennani, B. Matei, N. Grozavu et I. Falih. Collaborative Learning to Improve the Non-uniqueness of NMF. *International Journal of Computational Intelligence and Applications*, 21(1) :2250001 :1–24, 2022.

[RI-27] N. Dif, M. O. Attaoui, Z. Elberrichi, M. Lebbah et H. Azzag. Transfer learning from synthetic labels for histopathological images classification. *Applied Intelligence*, 52(1) :358–377, 2022.

[**RI-26**] M. El Hamri, Y. Bennani et I. Falih. Hierarchical Optimal Transport for Unsupervised Domain Adaptation. *Machine Learning*, 111(11) :4159–4182, 2022.

[**RI-25**] E. Goffinet, M. Lebbah, H. Azzag, G. Loïc et A. Coutant. Functional non-parametric latent block model : A multivariate time series clustering approach for autonomous driving validation. *Computational Statistics & Data Analysis*, 176 :107565 (23 pages), 2022.

2021 [RI-24] P. Alizadeh, A. Osmani, M. Essaid, A. Chibani et Y. Amirat. Reinforcement Learning for Interactive QoS-Aware Services Composition. *IEEE Systems Journal*, 15(1) :1098–1108, 2021.

[RI-23] P. Alizadeh, E. Traversi et A. Osmani. Deterministic policies based on maximum regrets in MDPs with imprecise rewards. *AI Communications*, 34(4) :229–244, 2021.

[RI-22] M. O. Attaoui, H. Azzag, M. Lebbah et N. Keskes. Subspace data stream clustering with global and local weighting models. *Neural Computing and Applications*, 33(8) :3691–3712, 2021.

[RI-21] F. E. Ben Bouazza, Y. Bennani, G. Cabanes et A. Touzani. Unsupervised collaborative learning based on Optimal Transport theory. *Journal of Intelligent Systems*, 30(1) :698–719, 2021.

[RI-20] F. Forest, M. Lebbah, H. Azzag et J. Lacaille. Deep embedded self-organizing maps for joint representation learning and topology-preserving clustering. *Neural Computing and Applications*, 33(24) :17439–17469, 2021.

[**RI-19**] M. Hamidi et A. Osmani. Domain models for data sources integration in HAR. *Neurocomputing*, 444 :244–259, 2021.

[RI-18] M. Hamidi et A. Osmani. Human Activity Recognition : A Dynamic Inductive Bias Selection Perspective. *Sensors*, 21(21) :7278 (42 pages), 2021.

[RI-17] S. Zouinina, Y. Bennani, N. Rogovschi et A. Lyhyaoui. Data Anonymization through Collaborative Multi-view Microaggregation. *Journal of Intelligent Systems*, 30(1) :327–345, 2021.

2020 [RI-16] Z. C. Dagdia, C. Zarges, G. Beck et M. Lebbah. A Scalable and Effective Rough Set Theory based Approach for Big Data Pre-processing. *Knowledge and Information Systems*, 62 :3321–3386, 2020.

2019 [RI-15] E. Agora, J. Aantezana, C. Cabrelli et B. Matei. Existence of quasicrystals and universal stable sampling and interpolation in LCA groups. *Transactions of the American Mathematical Society*, 372(7) :4647–4674, 2019.

[RI-14] M. Atzmueller, H. Soldano, G. Santini et D. Bouthinon. MinerLSD : efficient mining of local patterns on attributed networks. *Applied Network Science*, 4 :Article 43 (33 pages), 2019.

[RI-13] G. Beck, T. Duong, M. Lebbah, H. Azzag et C. Cérin. A distributed approximate nearest neighbors algorithm for efficient large scale mean shift clustering. *Journal of Parallel and Distributed Computing*, 134 :128–139, 2019.

[RI-12] B. Chien, A. Mokraoui et B. Matei. Contrast enhancement and details preservation of tone mapped high dynamic range images. *Journal of Visual Communication and Image Representation*, 58 :589–599, 2019.

[**RI-11**] A. Coutant, K. Roper, D. Trejo-Banos, D. Bouthinon, M. Carpenter, J. Grzebyta, G. Santini, H. Soldano, M. Elati, J. Ramon, C. Rouveirol, L. N. Soldatova et R. D. King. Closed-loop cycles of experiment design, execution, and learning accelerate systems biology model development in yeast. *The Proceedings of the National Academy of Sciences (PNAS)*, 116(36) :18142–181, 2019.

[RI-10] R. Interdonato, M. Atzmueller, S. Gaito, R. Kanawati, C. Largeron et A. Sala. Feature-rich networks : going beyond complex network topologies. *Applied Network Science*, 4 :Article 4 (13 pages), 2019.

[**RI-9**] P. Rastin, G. Cabanes, B. Matei, Y. Bennani et J.-M. Marty. A new sparse representation learning of complex data : Application to dynamic clustering of web navigation. *Pattern Recognition*, 91 :291–307, 2019.

[RI-8] P. Rastin, B. Matei, G. Cabanes, N. Grozavu et Y. Bennani. Impact of Learners' Quality and Diversity in Collaborative Clustering. *Journal of Artificial Intelligence and Soft Computing Research*, 9(2) :149–165, 2019.

[RI-7] H. Soldano, G. Santini, D. Bouthinon, S. Bary et E. Lazega. Bi-pattern mining of attributed networks. *Applied Network Science*, 4 :Article 37 (26 pages), 2019.

2018 [RI-6] A. Cornuéjols, C. Wemmert, P. Gançarski et Y. Bennani. Collaborative clustering : Why, when, what and how. *Information Fusion*, 39 :81–95, 2018.

[RI-5] H. Léger, D. Bouthinon, M. Lebbah et H. Azzag. An Instance Based Model for Scalable θ-Subsumption. *International Journal on Artificial Intelligence Tools*, 26(7) :1860011 (13 pages), 2018.

[RI-4] L. Picchi, G. Cabanes, C. Ricci-Bonot et M. Cristina. Quantitative Matching of Clutch Size in Reciprocating Hermaphroditic Worms. *Current Biology*, 28(20) :3254–3259 E3, 2018.

2017 [**RI-3**] Y. Chevaleyre, U. Endriss et N. Maudet. Distributed Fair Allocation of Indivisible Goods. *Artificial Intelligence*, 242 :1–22, 2017.

[RI-2] M. Ghesmoune, H. Azzag, S. Benbernou, M. Lebbah, T. Duong et M. Ouziri. Big Data : From Collection to Visualization. *Machine Learning*, 106 :837–862, 2017.

[**RI-1**] J. Sublime, B. Matei, G. Cabanes, N. Grozavu, Y. Bennani et A. Cornuéjols. Entropy based probabilistic collaborative clustering. *Pattern Recognition*, 72 :144 – 157, 2017.

Articles dans des revues nationales ou internationales

2022 [RE-5] M. Chevallier, N. Rogovschi, F. Boufarès et N. Grozavu. Detecting near duplicate dataset with machine learning. *International Journal of Computer Information Systems and Industrial Management Applications*, 14(32) :12, 2022.

[RE-4] A. F. Dos Reis, Y. Medjahdi, B. Sens, J. Sublime, G. Brante et F. Bader. Low Complexity LSTM-NN-Based Receiver for Vehicular Communications in the Presence of High-Power Amplifier Distortions. *IEEE Access*, 10 :121985–122000, 2022.

2019 [RE-3] S. Oujdi, H. Belbachir et F. Boufarès. C4.5 Decision Tree Algorithm for Spatial Data, Alternatives and Performances. *Journal of Computing and Information Technology*, 27(3) :29–43, 2019.

[RE-2] J. Sublime, G. Cabanes et B. Matei. Study on the Influence of Diversity and Quality in Entropy Based Collaborative Clustering. *Entropy*, 21(10) :1–21, 2019.

2017 [RE-1] A. Benyettou, Y. Bennani, A. Benyettou, A. Bendahmane et G. Cabanes. Semi-Supervised Multi-Label Classification Through Topological Active Learning. *International Journal on Communications Antenna and Propagation (IRECAP)*, 7(3) :222–232, 2017.

Chapitres de livre

2022 [CL-7] N. Grozavu, Y. Bennani, B. Matei et K. Benlamine. *Multi-view clustering based on non-negative matrix factorization*, Pedrycz, éditeur, *Recent Advancements in Multi-View Data Analytics*, volume 106 de *Studies in Big Data*, pages 177–200. Springer, 2022.

[CL-6] P.-A. Murena, J. Sublime et B. Matei. *Rethinking Collaborative Clustering : A Practical and Theoretical Study Within the Realm of Multi-view Clustering*, *Recent Advancements in Multi-View Data Analytics*, volume 106 de *Studies in Big Data*, pages 97–130. Witold Pedrycz, springer édition, 2022.

2021 [CL-5] M. Hamidi et A. Osmani. *Combining Domain Knowledge and Deep Learning to Improve HAR Models*, Z. Chen, M. Wu et X. Li, éditeurs, *Generalization With Deep Learning : For Improvement On Sensing Capability*, pages 55–77. World Scientific, 2021.

2019 [CL-4] S. Benbernou, M. Bentounsi, C. Cérin, M. Ouziri, P. Garteiser, L. Abidi, M. Lebbah, H. Azzag et M. Smadja. *A Big Data Platform for Enhancing Life Imaging Activities*, J. Darmont et S. Loudcher, éditeurs, *Utilizing Big Data Paradigms for Business Intelligence*, pages 39–71. IGI Global, 2019.

[CL-3] A. Osmani et M. Hamidi. *Bayesian optimization of neural architectures for human activity recognition*, N. Kawaguchi, N. Nishio, D. Roggen, S. Inoue, S. Pirttikangas et K. Van Laerhoven, éditeurs, *Human Activity Sensing*, pages 171–195. Springer, 2019.

2017 [CL-2] M. Pujari et R. Kanawati. *Link prediction in Large-scale multiplex networks*, S. Cordier, N. Debarsy, C. Ertur, F. Nemo, D. Nourrit-Lucas, G. Poisson et C. Vrain, éditeurs, *Understanding Interactions in complex systems*, Chapitre 7, pages 107–136. Cambridge Scholar Publishing, 2017.

[CL-1] H. Soldano, G. Santini et D. Bouthinon. *Formal Concept Analysis of Attributed Networks*, R. Missaoui, S. Obiedkov et S. Kuznetsov, éditeurs, *Formal Concept Analysis of Social Networks*, pages 143–170. Lecture Notes in Social Network. Springer, 2017.

Tutoriaux dans des conférences internationales

2023 [TU-3] S. M. R. Modaresi, A. Osmani, M. Razzazi et A. Chibani. Multi-Modal Evaluation Approach for Medical Image Segmentation. *arXiv*, 2023.

2017 [TU-2] N. Grozavu et R. Kanawati. Topological & Graph-based Clustering : Recent Algorithmic Advances : A tutorial. *The 2017 International Joint Conference on Neural Networks (IJCNN 2017), Anchorage, Alaska, USA, May 14-19, 2017*, pages 120–120, 2017.

[TU-1] R. Kanawati et M. Atzmueller. Mining Attributed Networks : A Tutorial. *The 4th IEEE International conference on Data Science and Advanced Analytics (DSAA 2017), Tokyo, Japon, October 17-21, 2017*, 2017.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2023 [CI-112] T. Viard, H. Soldano et G. Santini. Exploring and mining attributed sequences of interactions. *The 11th International Conference on Complex Networks and their Applications (COMPLEX NETWORKS XI), Palermo, Italy, November 8-10, 2022*, volume 1078 de *Studies in Computational Intelligence*, pages 537–549, 2023.

2022 [**CI-111**] P. Alizadeh, A. Osmani et S. Taleb. Improving Imitation Learning by Merging Experts Trajectories. *The 31st ACM International Conference on Information & Knowledge Management (CIKM'22), Atlanta, GA, USA, October 17-21, 2022*, pages 3796–3800, 2022.

[CI-110] M. Chevallier, N. Grozavu, F. Boufarès et N. Rogovschi. Detecting Near Duplicate Dataset. *The 13th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2022), Seattle, WA, United States, December 14-16, 2022*, volume 417 de *Lecture Notes in Networks and Systems*, pages 394–403, 2022.

[CI-109] M. Chevallier, N. Grozavu, N. Rogovschi, F. Boufarès et C. Clairmont. Trade Between Population Size and Mutation Rate for GAAM (Genetic Algorithm with Aggressive Mutation) for Feature Selection. *The 18th IFIP WG 12.5 International Conference on Artificial Intelligence Applications and Innovations (AIAI 2022), Hersonissos, Crete, Greece, June 17-20, 2022*, volume 646 de *IFIP Advances in Information and Communication Technology*, pages 432–444, 2022.

[CI-108] M. El Hamri, Y. Bennani et I. Falih. Incremental Unsupervised Domain Adaptation Through Optimal Transport. *The 2022 International Joint Conference on Neural Networks (IJCNN 2022) Padua, Italy, July 18-23, 2022*, pages 1–8, 2022.

[CI-107] M. El Hamri, Y. Bennani et I. Falih. When Domain Adaptation Meets Semi-supervised Learning Through Optimal Transport. *The 18th IFIP WG 12.5 International Conference on Artificial Intelligence Applications and Innovations (AIAI 2022), Hersonissos, Crete, Greece, June 17-20, 2022*, volume 646 de *IFIP Advances in Information and Communication Technology*, pages 58–69, 2022.

[CI-106] W. Fogos da Rocha, H. Azzag, M. Lebbah et A. Mokraoui. Super-Resolution GAN Improving YOLO's Performance Benchmark. *The 21st IEEE International Conference on Machine Learning and Applications (ICMLA 2022), Nassau, The Bahamas, December 12-14, 2022*, pages 1–8, 2022.

[**CI-105**] M. Hamidi et A. Osmani. Context Abstraction to Improve Decentralized Machine Learning in Structured Sensing Environments. *The 2022 Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2022), Grenoble, France, September 19-23, 2022*, volume 13715 de *Lecture Notes in Artificial Intelligence*, pages 647–663, 2022.

[CI-104] M. Hamidi, A. Osmani, L. Rasaq, G. Dogan et N. Alotaibi. Multiview Representation Learning for Human Activity Recognition. *The 9th IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA 2022), Chemnitz, Germany, June 15-17, 2022*, pages 1–5, 2022.

[CI-103] Y. Khalafaoui, N. Grozavu, B. Matei et L.-W. Goix. Multi-modal Multi-view Clustering based on Non-negative Matrix Factorization. *The 2022 IEEE Symposium Series on Computational Intelligence (SSCI 2022), Singapore, Singapore, December 4-7, 2022*, pages 1386–1391, 2022.

[CI-102] J.-V. Loddo et R. Kanawati. Mariotel : A Web-Based Virtual Remote Computer Science Lab. *The 20th International Conference on Information Technology Based Higher Education and Training (ITHET'22), Antalya, Turkey, November 7-9, 2022*, pages 1–6, 2022.

[CI-101] S. M. R. Modaresi, A. Osmani, M. Razzazi et A. Chibani. Evaluation of Early Diagnosis of COVID-19 Algorithms. *The 34th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2022), Macau, China, October 31-November, 2022*, pages 1–8, 2022.

[**CI-100**] S. M. R. Modaresi, A. Osmani, M. Razzazi et A. Chibani. Multimodal Evaluation Method for Sound Event Detection. *The IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2022), Singapore, 23-27 May 2022*, pages 1026–1030, 2022.

[**CI-99**] S. M. R. Modaresi, A. Osmani, M. Razzazi et A. Chibani. Uniform Evaluation of Properties in Activity Recognition. *The 26th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD 2022), Chengdu, China, May 16-19, 2022,*, volume 13281 de *Lecture Notes in Computer Science*, pages 83–95, 2022.

[**CI-98**] A. Osmani et M. Hamidi. Reduction of the Position Bias via Multi-level Learning for Activity Recognition. *The 26th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD 2022), Chengdu, China, May 16-19, 2022*, volume 13281 de *Lecture Notes in Computer Science*, pages 289–302, 2022.

[**CI-97**] A. Osmani, M. Hamidi et P. Alizadeh. Clustering Approach to Solve Hierarchical Classification Problem Complexity. *The 36th AAAI Conference on Artificial Intelligence (AAAI 2022), Virtual Conference, February 22-March 1, 2022*, volume 36(7) de *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 7904–7912, 2022.

[CI-96] A. Zaiou, Y. Bennani, M. Hibti et B. Matei. Quantum Approach for Vertex Separator Problem in Directed Graphs. *The 18th IFIP WG 12.5 International Conference on Artificial Intelligence Applications and Innovations (AIAI 2022), Hersonissos, Crete, Greece, June 17-20, 2022*, volume 646 de *IFIP Advances in Information and Communication Technology*, pages 495–506, 2022.

[CI-95] A. Zaiou, Y. Bennani, B. Matei et M. Hibti. A quantum learning approach based on Hidden Markov Models for failure scenarios generation. *The 2022 Asia Conference on Algorithms, Computing and Machine Learning (CACML 2022), Hangzhou, China, March 25-27, 2022*, pages 62–67, 2022.

2021

[CI-94] D. F. Andriantsiory, J. Ben Geloun et M. Lebbah. Multi-Slice Clustering for 3-order Tensor Data. *The 20th IEEE International Conference on Machine Learning and Applications (ICMLA 2021), Pasadena, CA, USA, December 13-16, 2021*, pages 173–178, 2021.

[CI-93] M. O. Attaoui, H. Azzag, N. Keskes et M. Lebbah. A New Subspace Multi-Objective Approach for the Clustering and Selection of Regions of Interests in Histopathological Images. *The 2021 IEEE Congress on Evolutionary Computation (CEC 2021), Kraków, Poland, June 28 - July 1, 2021*, pages 556–563, 2021.

[CI-92] G. Beck, M. Lebbah, H. Azzag et T. Duong. A New Nearest Neighbor Median Shift Clustering for Binary Data. *The 30th International Conference on Artificial Neural Networks (ICANN 2021), Bratislava, Slovakia, September 14-17, 2021*, volume 12895 de *Lecture Notes in Computer Science*, pages 101–112, 2021.

[CI-91] M. Chevallier, F. Boufarès, N. Grozavu et N. Rogovschi. Near duplicate Column identification : a machine learning approach. *The 2021 IEEE Symposium Series on Computational Intelligence (SSCI 2021), Orlando, FL, USA, December 5-7, 2021*, pages 1–7, 2021.

[CI-90] M. El Hamri, Y. Bennani et I. Falih. Inductive Semi-supervised Learning Through Optimal Transport. *The 28th International Conference on Neural Information Processing (ICONIP 2021), Sanur, Bali, Indonesia, December 8-12, 2021*, volume 1516 de *Communications in Computer and Information Science*, pages 668–675, 2021.

[CI-89] M. El Hamri, Y. Bennani et I. Falih. Label Propagation Through Optimal Transport. *The 2021 International Joint Conference on Neural Networks (IJCNN 2021), Shenzhen, China, July 18-22, 2021*, pages 1–8, 2021.

[CI-88] Y. Foucade, Y. Bennani et Z. Aabbou. Collaborative Random Forests Learning. *The 2021 IEEE 23rd International Conference on High Performance Computing & Communications ; 7th International Conference on Data Science & Systems ; 19th International Conference on Smart City ; 7th International Conference on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys 2021), Haikou, Hainan, China, December 20-22, 2021*, pages 1147–1154, 2021.

[CI-87] E. Goffinet, M. Lebbah, H. Azzag, L. Giraldi et A. Coutant. Multivariate Time Series Multi-Coclustering. application to advanced driving assistance system validation. *The 9th European Symposium on Artificial Neural Networks (ESANN 2021), Bruges, Belgium, October 6-8, 2021*, pages 463–468, 2021.

[**CI-86**] M. Hamidi et A. Osmani. Data generation process modeling for activity recognition. *The 2020 Joint European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2020), Ghent, Belgium, Septemebr 14-18, 2021*, volume 12460 de *Lecture Notes in Artificial Intelligence*, pages 374–390. Springer, 2021.

[CI-85] M. Kaddouh, G. Cabanes et Y. Bennani. Incorporating Neighborhood Information During NMF Learning. *The 28th International Conference on Neural Information Processing (ICONIP 2021), Sanur, Bali, Indonesia, December 8-12, 2021*, volume 1516 de *Communications in Computer and Information Science*, pages 591–598, 2021.

[CI-84] P. Le Jeune, M. Lebbah, A. Mokraoui et H. Azzag. Experience feedback using Representation Learning for Few-Shot Object Detection on Aerial Images. *20th IEEE International Conference on Machine Learning and Applications (ICMLA 2021), Pasadena, CA, USA, December 13-16, 2021*, pages 662–667, 2021.

[CI-83] T. Mecharnia, N. Pernelle, C. Rouveirol, F. Hamdi et L. Chibout. Mining Contextual Rules to Predict Asbestos in Buildings. *The 26th International Conference on Conceptual Structures (ICCS 2021), Berlin, Germany, September 20-22, 2021*, volume 12879 de *Lecture Notes in Artificial Intelligence*, pages 170–184, 2021.

[CI-82] A. Osmani, P. Alizadeh et C. Rodrigues. Equational Model Guided by Real-time Sensor Data to Monitor Industrial Robots. *17th IEEE International Conference on Automation Science and Engineering, CASE 2021, Lyon, France, August 23-27, 2021*, pages 1848–1854, 2021.

[**CI-81**] A. Osmani et M. Hamidi. Hierarchical Learning of Dependent Concepts for Human Activity Recognition. *The 25th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD 2021), Virtual Event, May 11-14, 2021*, volume 12713 de *Lecture Notes in Artificial Intelligence*, pages 79–92, 2021.

[**CI-80**] A. Osmani, M. Hamidi et S. Bouhouche. Augmented Experiment in Material Engineering Using Machine Learning. *The 35th AAAI Conference on Artificial Intelligence (AAAI 2021), Virtual Event, February 2-9, 2021*, volume 35 de *10*, pages 9251–9258, 2021.

[CI-79] G. Santini, H. Soldano et S. Zevio. Experiments on F-Restricted Bi-pattern Mining. *The 10th International Conference on Complex Networks and Their Applications (COMPLEX NETWORKS X), Madrid, Spain, November 30-December 2, 2021*, volume 1015 de *Studies in Computational Intelligence*, pages 539–550, 2021.

[CI-78] G. Santini, H. Soldano et S. Zevio. Restricted Bi-pattern Mining of Attributed Networks. *The 26th International Conference on Conceptual Structures (ICCS 2021), Virtual Event, September 20-22, 2021*, volume 12879 de *Lecture Notes in Artificial Intelligence*, pages 200–207, 2021.

[CI-77] A. Zaiou, Y. Bennani, B. Matei et M. Hibti. Balanced K-means using Quantum annealing. *The 2021 IEEE Symposium Series on Computational Intelligence (SSCI 2021), Orlando, FL, USA, December 5-7, 2021*, pages 1–7, 2021.

[CI-76] A. Zaiou, B. Matei, Y. Bennani et M. Hibti. Convex Non-negative Matrix Factorization Through Quantum Annealing. *The 23rd IEEE International Conference on High Performance Computing & Communications ; 7th International Conference on Data Science & Systems ; 19th International Conference on Smart City ; 7th International Conference on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys 2021), Haikou, Hainan, China, December 20-22, 2021*, pages 1253–1258, 2021.

2020 [CI-75] F. E. Ben Bouazza, Y. Bennani, G. Cabanes et A. Touzani. Collaborative Clustering Through Optimal Transport. *29th International Conference on Artificial Neural Networks (ICANN 2020), Bratislava, Slovakia, September 15-18, 2020*, volume 12397 de *Lecture Notes in Computer Science*, pages 873–885, 2020.

[CI-74] F. E. Ben Bouazza, Y. Bennani, A. Touzani et G. Cabanes. Subspace Guided Collaborative Clustering Based on Optimal Transport. *The 12th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2020), Hyderabad, India, December 15-18, 2020*, volume 1383 de *Advances in Intelligent Systems and Computing*, pages 113–124. Springer, 2020.

[CI-73] K. Benlamine, Y. Bennani, N. Grozavu et B. Matei. Quantum Collaborative K-means. *The 2020 International Joint Conference on Neural Networks (IJCNN 2020), Glasgow, United Kingdom, July 19-24, 2020*, pages 1–7, 2020.

[CI-72] K. Benlamine, Y. Bennani, B. Matei et N. Grozavu. Quantum Semi Non-negative Matrix Factorization. *The 12th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2020), Hyderabad, India, December 15–18, 2020*, volume 1383 de *Advances in Intelligent Systems and Computing*, pages 135–144, 2020.

[CI-71] M. Chevallier, N. Rogovschi, N. Grozavu et F. Boufarès. Seeding Initial Population, in Genetic Algorithm for Features Selection. *The 12th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2020), Online, India, December 15–18, 2020*, volume 1383 de *Advances in Intelligent Systems and Computing*, pages 572–582, 2020.

[CI-70] G. Dupont, E. Kalinicheva, J. Sublime, F. Rossant et M. Pâques. Unsupervised Change Detection Using Joint Autoencoders for Age-Related Macular Degeneration Progression. *The 29th International Conference on Artificial Neural Networks (ICANN 2020), Bratislava, Slovakia, September 15-18, 2020*, volume 12397 de *Lecture Notes in Computer Science*, pages 813–824, 2020.

[CI-69] F. Forest, Q. Cochard, C. Noyer, A. Cadut, M. Joncour, J. Lacaille, M. Lebbah et H. Azzag. Large-scale Vibration Monitoring of Aircraft Engines from Operational Data using Self-organized Models. *The 14th Annual Conference of the Prognostics and Health Management Society (PHM22), Nashville, Tennessee, November 1-4, 2022*, pages 1–6, 2020.

[**CI-68**] F. Forest, A. Mourer, M. Lebbah, H. Azzag et J. Lacaille. An Invariance-guided Stability Criterion for Time Series Clustering Validation. *The 25th International Conference on Pattern Recognition (ICPR 2020), Virtual Event, Milan, Italy, January 10-15, 2021*, pages 9296–9303, 2020.

[CI-67] Y. Foucade et Y. Bennani. Unsupervised Learning from Data and Learners. *The 12th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2020), Hyderabad, India, December 15-18, 2020*, volume 1383 de *Advances in Intelligent Systems and Computing*, pages 496–507, 2020.

[**CI-66**] E. Goffinet, M. Lebbah, H. Azzag et L. Giraldi. Autonomous Driving Validation with Model-Based Dictionary Clustering. *The 2020 European Conference Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2020), Ghent, Belgium, September 14-18, 2020*, volume 12460 de *Lecture Notes in Computer Science*, pages 323–338, 2020.

[CI-65] M. Hamidi, A. Osmani et P. Alizadeh. A Multi-View Architecture for the SHL Challenge. *The 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and 2020 ACM International Symposium on Wearable Computers (UbiComp/ISWC '20), Virtual Event, Mexico, September 12-17, 2020*, pages 317–322, 2020.

[CI-64] J. Zamora et J. Sublime. A New Information Theory Based Clustering Fusion Method for Multi-view Representations of Text Documents. *The 12th International Conference on Social Computing and Social Media (SCSM 2020) held as Part of the 22nd HCI International Conference (HCI 2020), Copenhagen, Denmark, July 19-24, 2020*, volume 12194 de *Lecture Notes in Computer Science*, pages 156–167, 2020.

[CI-63] S. Zouinina, Y. Bennani, N. Rogovschi et A. Lyhyaoui. A Two-Levels Data Anonymization Approach. *The 16th IFIP WG 12.5 International Conference on The Artificial Intelligence Applications and Innovations (AIAI 2020), Neos Marmaras, Greece, June 5-7, 2020, Proceedings, Part I*, volume 583 de *IFIP Advances in Information and Communication Technology*, pages 85–95, 2020.

2019 [CI-62] M. O. Attaoui, M. Lebbah, N. Keskes, H. Azzag et M. Ghesmoune. Soft Subspace Growing Neural Gas for Data Stream Clustering. *The 28th International Conference on Artificial Neural Networks (ICANN 2019), Munich, Germany, September 17-19, 2019*, pages 569–580, 2019.

[CI-61] F. E. Ben Bouazza, Y. Bennani, M. El, G. Cabanes, B. Matei et A. Touzani. Multi-view Clustering through Optimal Transport. *The 26th International Conference on Neural Information Processing of the Asia-Pacific Neural Network Society (ICONIP 2019) Sydney, Australia, December 12-15 2019*, volume 15 de *Australian Journal of Intelligent Information Processing Systems*, pages 1–9, 2019.

[CI-60] K. Benlamine, Y. Bennani, A. Zaiou, M. Hibti, B. Matei et N. Grozavu. Distance Estimation for Quantum Prototypes Based Clustering. *The 26th International Conference on Neural Information Processing (ICONIP 2019), Sydney, NSW, Australia, December 12-15, 2019*, volume 11955 de *Lecture Notes in Computer Science*, pages 561–572, 2019.

[CI-59] K. Benlamine, N. Grozavu, Y. Bennani et B. Matei. Collaborative Non-negative Matrix Factorization. *The 28th International Conference on Artificial Neural Networks (ICANN 2019), Munich, Germany, September 17-19, 2019*, volume 11730 de *Lecture Notes in Computer Science*, pages 655–666, 2019.

[CI-58] I. Brahmi, G. Cabanes, Y. Bennani et A. Touzani. Deepness Analysis of Learned Factors in Multilayer NMF. *The 26th International Conference on Neural Information Processing of the Asia-Pacific Neural Network Society (ICONIP 2019) Sydney, Australia, December 12-15 2019*, volume 16 de *Australian Journal of Intelligent Information Processing Systems*, pages 24–31, 2019.

[CI-57] F. Forest, M. Lebbah, H. Azzag et J. Lacaille. Deep Embedded SOM : joint representation learning and self-organization. *The 27th European Symposium on Artificial Neural Networks (ESANN 2019), Bruges, Belgium, April 24-26, 2019*, pages 1–6, 2019.

[CI-56] E. Kalinicheva, J. Sublime et M. Trocan. Change Detection in Satellite Images Using Reconstruction Errors of Joint Autoencoders. *The 28th International Conference on Artificial Neural Networks (ICANN 2019), Munich, Germany, September 17-19, 2019*, volume 11729 de *Lecture Notes in Computer Science*, pages 637–648, 2019.

[CI-55] E. Kalinicheva, J. Sublime et M. Trocan. Object-Based Change Detection in Satellite Images Combined with Neural Network Autoencoder Feature Extraction. *The 9th International Conference on Image Processing Theory, Tools and Applications (IPTA 2019), Istanbul, Turkey, November 6-9, 2019*, pages 1–6, 2019.

[CI-54] R. Kanawati et M. Atzmueller. Modeling and Mining Feature-Rich Networks. *The 2019 World Wide Web Conference (WWW 2019), San Francisco, CA, USA, May 13-17, 2019*, pages 1306–1307, 2019.

[CI-53] H. Kordestani, R. Mojarad, A. Chibani, A. Osmani, Y. Amirat, K. Barkaoui et W. Zahran. Hapicare : A Healthcare Monitoring System with Self-Adaptive Coaching using Probabilistic Reasoning. *The 16th IEEE/ACS International Conference on Computer Systems and Applications (AICCSA 2019), Abu Dhabi, UAE, November 3-7, 2019*, pages 1–8, 2019.

[**CI-52**] A. Osmani, M. Hamidi et S. Bouhouche. Monitoring of a Dynamic System Based on Autoencoders. *The 28th International Joint Conference on Artificial Intelligence (IJCAI 2019), Macao, China, August 10-16, 2019*, pages 1836–1843, 2019.

[CI-51] P. Rastin, Y. Bennani et R. Verde. Automatic detection of the support points in relational clustering. *The 2019 International Joint Conference on Neural Networks (IJCNN 2019), Budapest, Hungary, July 14-19, 2019*, pages 1–8, 2019.

[CI-50] P. Rastin, G. Cabanes, R. Verde, Y. Bennani et T. Couronne. Generative Histogram-Based Model Using Unsupervised Learning. *The 26th International Conference on Neural Information Processing (ICONIP 2019), Sydney, NSW, Australia, December 12-15, 2019*, volume 11955 de *Lecture Notes in Computer Science*, pages 634–646. Springer, 2019.

[CI-49] N. Rogovschi, S. Zouinina, B. Matei, I. Falih, N. Grozavu et S. Ozawa. t-Distributed Stochastic Neighbor Embedding Spectral Clustering using higher order approximations. *The 26th International Conference on Neural Information Processing of the Asia-Pacific Neural Network Society (ICONIP 2019) Sydney, Australia, December 12-15 2019*, volume 17 de *Australian Journal of Intelligent Information Processing Systems*, pages 78–86, 2019.

[CI-48] H. Soldano, G. Santini et D. Bouthinon. Attributed Graph Pattern Set Selection Under a Distance Constraint. *The 8th International Conference on Complex Networks and Their Applications (COMPLEX NETWORKS VIII), Lisbon, Portugal, December 10-12, 2019*, volume 882 de *Studies in Computational Intelligence*, pages 228–241, 2019.

[CI-47] J. Sublime. Incremental Collaborative Clustering Using Information Theory and Information Compression. *The 5th International Conference on Fuzzy Systems and Data Mining (FSDM 2019), Kitakyushu City, Japan, October 18-21, 2019*, volume 320 de *Frontiers in Artificial Intelligence and Applications*, pages 457–464, 2019.

[CI-46] S. Zouinina, Y. Bennani, M. Ben-Fares, N. Rogovschi et A. Lyhyaoui. Preserving Utility during Attribute-oriented Data Anonymization Process. *26th International Conference on Neural Information Processing of the Asia-Pacific Neural Network Society (ICONIP 2019) Sydney, Australia, December 12-15 2019*, volume 16 de *Australian Journal of Intelligent Information Processing Systems*, pages 25–35, 2019.

2018 [CI-45] G. Beck, H. Azzag, S. Bougeard, N. Niang et M. Lebbah. A New Micro-Batch Approach for Partial Least Square Clusterwise Regression. *The 3rd INNS Conference on Big Data and Deep Learning (INNS BDDL 2018), Sanur, Bali, Indonesia, April 17-19, 2018*, volume 144 de *Procedia Computer Science*, pages 239–250, 2018.

[CI-44] K. Benlamine, N. Grozavu, Y. Bennani, N. Rogovschi, K. Haddadou et A. Amamou. Domain Name Recommendation based on Neural Network. *The 2018 INNS Conference on Big Data and Deep Learning 2018, Sanur, Bali, Indonesia, 17-19 April 2018*, volume 144 de *Procedia Computer Science*, pages 60–70, 2018.

[CI-43] S. Boulbazine, G. Cabanes, B. Matei et Y. Bennani. Online Semi-supervised Growing Neural Gas for Multi-label Data Classification. *The 2018 International Joint Conference on Neural Networks (IJCNN 2018), Rio de Janeiro, Brazil, July 8-13, 2018*, pages 1–8, 2018.

[CI-42] I. Brahmi, G. Cabanes, Y. Bennani et B. Matei. Learning Useful Representations Through Stacked Self-Organizing Maps. *The 2018 International Joint Conference on Neural Networks (IJCNN 2018), Rio de Janeiro, Brazil, July 8-13, 2018*, pages 1–8, 2018.

[CI-41] C. Cérin, T. Menouer et M. Lebbah. Accelerating the Computation of Multi-Objectives Scheduling Solutions for Cloud Computing. *The 8th IEEE International Symposium on Cloud and Service Computing (SC2 2018), Paris, France, November 18-21, 2018*, pages 49–56, 2018.

[CI-40] Z. C. Dagdia, C. Zarges, G. Beck, H. Azzag et M. Lebbah. A Distributed Rough Set Theory Algorithm based on Locality Sensitive Hashing for an Efficient Big Data Preprocessing. *The IEEE International Conference on Big Data (Big Data 2018), Seattle, WA, USA, December 10-13, 2018*, pages 2597–2606, 2018.

[CI-39] N.-Q. Doan, H. Azzag et M. Lebbah. Hierarchical Laplacian Score for unsupervised feature selection. *The 2018 International Joint Conference on Neural Networks (IJCNN 2018), Rio de Janeiro, Brazil, July 8-13, 2018*, pages 1–7, 2018.

[**CI-38**] I. Falih, N. Grozavu, R. Kanawati et Y. Bennani. Community detection in Attributed Network. *The 2018 Word Wide Web Conference (WWW 2018), Lyon, France, April 23-27, 2018*, pages 1299–1306, 2018.

[CI-37] I. Falih, N. Grozavu, R. Kanawati et Y. Bennani. Topological Multi-View Clustering for Collaborative Filtering. *The 3rd INNS Conference on Big Data and Deep Learning 2018 (INNS BDDL 2018), Sanur, Bali, Indonesia, 17-19 April 2018*, volume 144 de *Procedia Computer Science*, pages 306–312, 2018.

[CI-36] I. Falih, N. Grozavu, R. Kanawati, Y. Bennani et B. Matei. Collaborative Multi-View Attributed Networks Mining. *The 2018 International Joint Conference on Neural Networks (IJCNN 2018), Rio de Janeiro, Brazil, July 8-13, 2018*, pages 1–8, 2018.

[CI-35] F. Forest, J. Lacaille, M. Lebbah et H. Azzag. A Generic and Scalable Pipeline for Large-Scale Analytics of Continuous Aircraft Engine Data. *The 2018 IEEE International Conference on Big Data (Big Data 2018), Seattle, WA, USA, December 10-13, 2018*, pages 1918–1924, 2018.

[CI-34] M. Ghesmoune, M. Lebbah, H. Azzag, S. Benbernou, M. Ouziri et T. Duong. A Complete Data Science Work-flow For Insurance Field. *The 2018 EEE International Conference on Big Data (Big Data 2018), Seattle, WA, USA, December 10-13, 2018*, pages 1925–1930, 2018.

[CI-33] E. Kalinicheva, J. Sublime et M. Trocan. Neural Network Autoencoder for Change Detection in Satellite Image Time Series. *The 25th IEEE International Conference on Electronics, Circuits and Systems (ICECS 2018), Bordeaux, France, December 9-12, 2018*, pages 641–642, 2018.

[CI-32] S. Legras, C. Rouveirol et V. Ventos. The Game of Bridge : A Challenge for ILP. *The 28th International Conference on Inductive Logic Programming (ILP 2018), Ferrara, Italy, September 2-4, 2018*, volume 11105 de *Lecture Notes in Artificial Intelligence*, pages 72–87, 2018.

[**CI-31**] P.-A. Murena, J. Sublime, B. Matei et A. Cornuéjols. An Information Theory based Approach to Multisource Clustering. *The 27th International Joint Conference on Artificial Intelligence (IJCAI 2018), July 13-19, 2018, Stockholm, Sweden*, pages 2581–2587, 2018.

[CI-30] L. H. Ngo, M. Luong, N. Metodiev, T. Le-Tien, S. Guérif et E. Viennet. Sparse Representation Wavelet Based Classification. *The 2018 IEEE International Conference on Image Processing (ICIP 2018), Athens, Greece, October 7-10, 2018*, pages 2974–2978, 2018.

[CI-29] A. Osmani et M. Hamidi. Hybrid and convolutional neural networks for locomotion recognition. *The 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers (UbiComp '18), Singapore, October 9-11, 2018*, pages 1531–1540, 2018.

[CI-28] P. Rastin, G. Cabanes, B. Matei et J.-M. Marty. Change Detection in Individual Users' Behavior. *The 27th International Conference on Artificial Neural Networks and Machine Learning (ICANN 2018) , Rhodes , Greece, October 4-7, 2018*, pages 501–510, 2018.

[CI-27] H. Soldano, S. Bary, G. Santini et D. Bouthinon. Core Stratification of Two-Mode Networks. *The 7th International Conference on Complex Networks and their Applications (COMPLEX NETWORKS VII), Cambridge, UK, December 11-13, 2018*, volume 812 de *Studies in Computational Intelligence*, pages 130–142, 2018.

[**CI-26**] H. Soldano, G. Santini, D. Bouthinon, S. Bary et E. Lazega. Bi-Pattern Mining of Two Mode and Directed Networks. *The 2018 World Wide Web Conference 2018 (WWW'18), Lyon, France, April 23-27, 2018*, pages 1287–1294, 2018.

[CI-25] J. Sublime et S. Lefebvre. Collaborative Clustering through Constrained Networks using Bandit Optimization. *The 2018 International Joint Conference on Neural Networks (IJCNN 2018), Rio de Janeiro, Brazil, July 8-13, 2018*, pages 1–8, 2018.

[CI-24] J. Sublime, D. Maurel, N. Grozavu, B. Matei et Y. Bennani. Optimizing exchange confidence during collaborative clustering. *The 2018 International Joint Conference on Neural Networks (IJCNN 2018), Rio de Janeiro, Brazil, July 8-13, 2018*, pages 1–8, 2018.

[CI-23] L.-M. Veillon, G. Bourgne et H. Soldano. Better Collective Learning with Consistency Guarantees. *The 21th International Conference on Principles and Practice of Multi-Agent Systems (PRIMA 2018), Tokyo, Japan, October 29-November 2, 2018*, volume 11224 de *Lecture Notes in Artificial Intelligence*, pages 671–679, 2018.

[CI-22] S. Zouinina, N. Grozavu, Y. Bennani, A. Lyhyaoui et N. Rogovschi. A Topological k-Anonymity Model Based on Collaborative Multi-view Clustering. *The 27th International Conference on Artificial Neural Networks (ICANN 2018), Rhodes, Greece, October 4-7, 2018*, volume 11141 de *Lecture Notes in Computer Science*, pages 817–827, 2018.

[CI-21] S. Zouinina, N. Grozavu, Y. Bennani, A. Lyhyaoui et N. Rogovschi. Efficient k-Anonymization through Constrained Collaborative Clustering. *The 2018 IEEE Symposium Series on Computational Intelligence (SSCI 2018), Bangalore, India, November 18-21, 2018*, pages 405–411, 2018.

2017 [**CI-20**] S. Bouveret, Y. Chevaleyre, F. Durand et J. Lang. Voting by sequential elimination with few voters. *The 26h International Joint Conference on Artificial Intelligence (IJCAI 2017), Melbourne, Australia, August 19-25, 2017*, pages 128–134, 2017.

[CI-19] C. Cérin, J.-L. Gaudiot, M. Lebbah et F. Yuehgoh. Return of experience on the mean-shift clustering for heterogeneous architecture use case. *The 2017 IEEE International Conference on Big Data (BigData 2017), Boston, MA, USA, December 11-14, 2017*, pages 3499–3507, 2017.

[CI-18] Z. C. Dagdia, C. Zarges, G. Beck et M. Lebbah. A Distributed Rough Set Theory based Algorithm for an Efficient Big Data Pre-processing under the Spark Framework. *The 2017 IEEE International Conference on Big Data (Big Data 2017), Boston, MA, USA, December 11-14, 2017*, pages 911–916, 2017.

[CI-17] I. Falih, N. Grozavu, R. Kanawati et Y. Bennani. ANCA : Attributed Network Clustering Algorithm. *The 6th International Conference on Complex Networks and Their Applications (COMPLEX NETWORKS VI), Lyon, France, November 29 - December 1, 2017*, volume 689 de *Studies in Computational Intelligence*, pages 241–252, 2017.

[CI-16] M. Hmimida et R. Kanawati. A graph-based meta-approach for tag recommendation. H. Cherifi, S. Gaito, W. Quattrociocchi et A. Sala, éditeurs, *The 5th International Conference on Complex Networks and Their Applications (COMPLEX NETWORKS 2016), Milano, Italy, Novembre 21-15, 2016*, volume 693 de *Studies in Computational Intelligence*, pages 309–320. Springer, 2017.

[CI-15] A. Lachaud, N. Grozavu, B. Matei et Y. Bennani. Collaborative clustering between different topological partitions. *The 2017 International Joint Conference on Neural Networks (IJCNN 2017), Anchorage, AK, USA, May 14-19, 2017*, pages 4111–4117. IEEE, 2017.

[**CI-14**] C. Laclau, I. Redko, B. Matei, Y. Bennani et V. Brault. Co-clustering through Optimal Transport. *The 34th International Conference on Machine Learning (ICML 2017), Sydney, NSW, Australia, 6-11 August 2017*, volume 70 de *Proceedings of Machine Learning Research*, pages 1955–1964, 2017.

[CI-13] H. Léger, D. Bouthinon, M. Lebbah et H. Azzag. An Instance Based Model for Scalable θ-Subsumption. *The 29th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2017), Boston, MA, USA, November 6-8, 2017*, pages 846–852, 2017.

[CI-12] F. Lezzar, D. Benmerzoug, I. Kitouni et A. Osmani. A GRC-Centric Approach for Enhancing Management Process of IoT-Based Health Institution. *The 3rd International Conference on Mobile, Secure, and Programmable Networking (MSPN 2017), Paris, France, June 29-30, 2017, Revised Selected Papers*, volume 10566 de *Lecture Notes in Computer Science*, pages 207–221, 2017.

[CI-11] D. Maurel, J. Sublime et S. Lefebvre. Incremental Self-Organizing Maps for Collaborative Clustering. *The 24th International Conference on Neural Information Processing (ICONIP 2017), Guangzhou, China, November 14-18, 2017*, volume 10634 de *Lecture Notes in Computer Science*, pages 497–504, 2017.

[CI-10] A. Osmani, M. Hamidi et A. Chibani. Machine Learning Approach for Infant Cry Interpretation. *The 29th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2017), Boston, MA, USA, November 6-8, 2017*, pages 182–186, 2017.

[CI-9] A. Osmani, M. Hamidi et A. Chibani. Platform for Assessment and Monitoring of Infant Comfort. *The 2017 AAAI Fall Symposia, Arlington, Virginia, USA, November 9-11, 2017*, pages 36–44, 2017.

[CI-8] P. Rastin, B. Matei et G. Cabanes. Signal-Based Autonomous Clustering for Relational Data. *The 2017 International Joint Conference on Neural Networks (IJCNN 2017), Anchorage, Alaska, USA, May 14-19, 2017*, pages 4103–4110, 2017.

[CI-7] M. Rifi, M. Hibti et R. Kanawati. Exploring network metrics for accident scenarios : a Case of study of the Uncontrolled Level Drop. *The 2017 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2017), Pittsburgh, Pennsylvanie, USA, September 24–28, 2017*, Proceedings.com, pages 1177–1184, 2017.

[CI-6] N. Rogovschi, J. Kitazono, N. Grozavu, T. Omori et S. Ozawa. t-Distributed stochastic neighbor embedding spectral clustering. *The 2017 International Joint Conference on Neural Networks (IJCNN 2017), Anchorage, Alaska, USA, May 14-19, 2017*, pages 1628–1632, 2017.

[CI-5] N. Sokolovska, Y. Chevaleyre et J.-D. Zucker. The Fused Lasso Penalty for Learning Interpretable Medical Scoring Systems. *The 2017 International Joint Conference on Neural Networks (IJCNN 2017), Anchorage, AK, USA, May 14-19, 2017*, pages 4504–4511, 2017.

[CI-4] H. Soldano, G. Santini, D. Bouthinon et E. Lazega. Hub-Authority Cores and Attributed Directed Network Mining. *The 29th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2017), Boston, MA, USA, November 6-8, 2017*, pages 1120–1127, 2017.

[CI-3] J. Sublime, B. Matei et P.-A. Murena. Analysis of the influence of diversity in collaborative and multi-view clustering. *The 2017 International Joint Conference on Neural Networks (IJCNN 2017), Anchorage, AK, USA, May 14-19, 2017*, pages 4126–4133, 2017.

[CI-2] L.-M. Veillon, G. Bourgne et H. Soldano. Effect of Network Topology on Neighbourhood-Aided Collective Learning. *The 9th International Conference on Computational Collective Intelligence (ICCCI), Nicosia, Cyprus, September 27-29, 2017*, volume 10448 de *Lecture Notes in Artificial Intelligence*, pages 202–211, 2017.

[CI-1] L.-M. Veillon, G. Bourgne et H. Soldano. Waves : a Model of Collective Learning. *2017 IEEE/WIC/ACM International Conference on Web Intelligence (WI'17), Leipzig, Germany, August 23-26, 2017*, pages 314–321, 2017.

Autres articles dans des conférences ou workshops à comité de lecture

2023 [CO-32] R. Khoufache, M. D. Dilmi, H. Azzag, M. Lebbah et E. Goffinet. Propriétés émergentes du multi-clustering bayésien non paramétrique : Application aux données d'images multivues. *20ème Journées Francophones Extraction et Gestion des Connaissances (EGC 2023), Lyon, France, 16-20 Janvier 2023*, volume E-39 de *Revue des Nouvelles Technologies Informatiques*, pages 639–640, 2023.

2022 [CO-31] M. Atzmueller et R. Kanawati. Explainability in Cyber-Security using Complex Network Analysis : A Brief Methodological Overview. *The European Interdisciplinary Cybersecurity Conference (EICC 2022), Barcelona, Spain, June 15 - 16, 2022*, pages 49–52, 2022.

[CO-30] R. Khoufache, M. D. Dilmi, H. Azzag, E. Goffinet et M. Lebbah. Emerging properties from Bayesian Non-Parametric for multiple clustering : Application for multi-view image dataset. *2022 IEEE International Conference on Data Mining Workshops (ICDM Workshops 2022), Orlando, FL, USA, November 28-December 1, 2022*, pages 31–38, 2022.

[CO-29] E. Leclercq, J. Rivalan, F. Roupin et C. Rouveirol. Allocation de ressources par une méthode hybride machine learning - optimisation en contexte de conteneurs. *Extraction et Gestion des Connaissances (EGC 2022), Blois, France, 24-28 janvier 2022*, volume E-38 de *Revue des Nouvelles Technologies de l'Information*, pages 499–500, 2022.

[CO-28] J.-V. Loddo et R. Kanawati. Mariotel : A virtual remote computer science lab. *The 1st International Conference on Online-Labs in Education, Stuttgart, Germany, March 10-12, 2022*, pages 485–494, 2022.

2021 [CO-27] M. Chevallier, N. Rogovschi, F. Boufarès et N. Grozavu. Techniques de génération de population initiale d'algorithmes génétiques pour la sélection de caractéristiques. *9ème Conférence Internationale Francophone sur la Science des Données (CIFSD 2021), Marseille, France, 9-11 Juin 2021*, pages 1–12, 2021.

[CO-26] E. Goffinet, M. Lebbah, H. Azzag, L. Giraldi et A. Coutant. Multiple Co-clustering de séries temporelles. *52èmes Journées de Statistique de la Société Française de Statistique (SFdS 2021), Nice, France, 7-11 Juin 2021*, 2021.

[CO-25] M. Hamidi et A. Osmani. Modélisation du processus de génération de données pour la reconnaissance d'activités. *Extraction et Gestion des Connaissances (EGC 2021), Montpellier, France, 25-29 janvier 2021*, volume E-37 de *Revue des Nouvelles Technologies de l'Information*, pages 405–412, 2021.

[CO-24] T. Mecharnia, L. Chibout Khelifa, F. Hamdi, N. Pernelle et C. Rouveirol. Découverte de règles contextuelles pour prédire la présence d'amiante dans les bâtiments. *32èmes Journées francophones d'Ingénierie des Connaissances (IC 2021), Bordeaux, France, June 30 - July 2, 2021*, pages 81–88, 2021.

[CO-23] J. Schoenke, N. Aschenbruck, R. Interdonato, R. Kanawati, A.-C. Meisener, F. Thierart, G. Vial et M. Atzmueller. Gaia-AgStream : An Explainable AI Platform for Mining Complex Data Streams in Agriculture. *The First International Conference on Smart and Sustainable Agriculture (SSA 2021), Virtual Event, June 20-21, 2021*, volume 1470 de *Communications in Computer and Information Science*, pages 71–83, 2021.

2020 [CO-22] M. O. Attaoui, H. Azzag, M. Lebbah et N. Keskes. Multi-Objective Data Stream Clustering. *The 2020 Genetic and Evolutionary Computation Conference (GECCO'20), Cancún, Mexico, July 8-12, 2020*, pages 113–114, 2020. (Poster).

[CO-21] M. O. Attaoui, H. Azzag, M. Lebbah, N. Keskes et M. Ghesmoune. Soft Subspace Growing Neural Gas pour le Clustering de Flux de Données. *Extraction et Gestion des Connaissances (EGC 2020), Brussels, Belgium, January 27-31, 2020*, volume E-36 de *Revue des Nouvelles Technologies de l'Information*, pages 441–448, 2020.

[CO-20] H. N. Benkhaled, D. Berrabah et F. Boufarès. Block Sizes Control For an Efficient Real Time Record Linkage. *The 5th IEEE International Conference on Cloud Computing and Artificial Intelligence : Technologies and Applications (CloudTech 2020), Marrakesh, Morocco, November 24-26, 2020*, pages 1–6, 2020.

[CO-19] C. Deheeger, J. Pierre, J. Baptiste, S. Legras, A. Rimbaud, C. Rouveirol, H. Soldano et V. Ventos. Une expérience d'élicitation de connaissances expertes dans le domaine du Bridge. A. Cornuéjols et E. Cuvelier, éditeurs, *Extraction et Gestion des Connaissances (EGC 2020), Brussels, Belgium, January 27-31, 2020*, volume E-36, de *Revue des Nouvelles Technologies de l'Information*, pages 97–108. Éditions RNTI, 2020.

[CO-18] F. Forest, M. Lebbah, H. Azzag et J. Lacaille. Carte SOM profonde : Apprentissage joint de représentations et auto-organisation. *Conférence d'Apprentissage (CAp 2020), Vannes, France, June 23-26, 2020*, 2020.

[CO-17] E. Goffinet, M. Lebbah, H. Azzag et L. Giraldi. Clustering de séries temporelles par construction de dictionnaire. *20ème Edition des Journées Francophones Extraction et Gestion des Connaissances (EGC 2020), Bruxelles, Belgique, 27-31 Janvier 2021*, volume E-36 de *Revue des Nouvelles Technologies*, pages 181–192, 2020.

[CO-16] D. Maurel, S. Lefebvre et J. Sublime. Deep Cooperative Reconstruction with Security Constraints in Multi-View Environments. *The 20th International Conference on Data Mining Workshops (ICDM Workshops 2020), Sorrento, Italy, November 17-20, 2020*, pages 579–588, 2020.

[CO-15] S. Zevio, G. Santini, H. Soldano, H. Zargayouna et T. Charnois. A Combination of Semantic Annotation and Graph Mining for Expert Finding in Scholarly Data. *The 2020 Workshop on Graph Embedding and Mining Worshop (GEM 2020) co-located with European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2020), Virtual Event, September 18, 2020*, page 16, 2020.

2019

[CO-14] P. Alizadeh, A. Osmani et E. Traversi. Calcul d'une politique déterministe dans un MDP avec récompenses imprécises. *Extraction et Gestion des connaissances (EGC 2019), Metz, France, January 21-25, 2019*, volume E-35 de *Revue des Nouvelles technologies*, pages 45–56, 2019.

[CO-13] H. N. Benkhaled, D. Berrabah et F. Boufarès. A novel approach to improve the Record Linkage process. *The 6th IEEE International Conference on Control, Decision and Information Technologies (CoDIT 2019), Paris, France, April 23-26, 2019*, pages 1504–1509, 2019.

[CO-12] M. Hamidi et A. Osmani. HAR Data Generation Model to Highly Reduce Learning Space. *Knowledge Representation & Reasoning Meets Machine Learning workshop (K2RML@NeurIPS 2019), Vancouver, Canada, September 19, 2019*, pages 1–11, 2019.

[CO-11] M. Hamidi et A. Osmani. Improving Human Activity Recognition with Data Sources Integration. *Deep Learning for Human Activity Recognition workshop @IJCAI 2019, Macao, China, August 10-16, 2019*, 2019.

[CO-10] P. Rastin, B. Matei et G. Cabanes. Apprentissage non-supervisé relationnel dans l'espace des coordonnées barycentriques. *Extraction et Gestion des Connaissances (EGC 2019), Metz, France, January 21-25, 2019*, volume E-35 de *Revue des Nouvelles Technologies de l'Information*, pages 413–418, 2019.

[CO-9] P. Rastin, B. Matei et G. Cabanes. Détection de changement dans les profils en ligne d'utilisateurs. *Extraction et Gestion des Connaissances (EGC 2019), Metz, France, January 21-25, 2019*, volume E-35 de *Revue des Nouvelles Technologies de l'Information*, pages 419–424, 2019.

[CO-8] S. Zevio, G. Santini, H. Zargayouna et T. Charnois. Fouille de texte et fouille de graphe appliquées à la recherche d'experts. *30èmes journées francophones d'Ingénierie des Connaissances (IC 2019), Toulouse, France, 2-4 juillet 2019*, pages 222–223, 2019.

2018

[CO-7] M. Atzmueller, H. Soldano, G. Santini et D. Bouthinon. MinerLSD : Efficient Local Pattern Mining on Attributed Graphs. *The 2018 IEEE International Conference on Data Mining Workshops (ICDMW 2018), Singapore, November 17- 20, 2018*, pages 219–228, 2018.

[CO-6] D. Maurel, J. Sublime et S. Lefebvre. Cartes Auto-Organisatrices Incrémentales appliquées au Clustering Collaboratif. *Extraction et Gestion des Connaissances (EGC 2018), Paris, France, January 23-26, 2018*, volume E-34 de *Revue des Nouvelles Technologies de l'Information*, pages 403–408, 2018.

[CO-5] M. Rifi, M. Hbiti et R. Kanawati. A Complex Network Analysis Approach for Risk Increase Factor Prediction in Nuclear Power Plants. *The 3rd International Conference on Complexity, Future Information Systems and Risk (COMPLEXIS 2018), Funchal, Madeira, Portugal, March 20-21, 2018*, 2018.

[CO-4] S. Zevio, H. Zargayouna, G. Santini et T. Charnois. Vers une cartographie automatique des thématiques et profils d'experts associés à une conférence scientifique : 9 ans d'ateliers recherche d'information semantique. *10ème Atelier Recherche d'Information Sémantique (RISE@CORIA/TALN 2018), Rennes, France, 15 Mai 2018*, rennes, france, 2018.

2017

[CO-3] G. Beck, H. Azzag, M. Lebbah et T. Duong. Mean-shift : Clustering scalable avec les plus proches voisins approximés. *Extraction et Gestion des Connaissances (EGC 2017), Grenoble, France, 24-27 Janvier 2017*, volume E-34 de *Revue des Nouvelles Technologies de l'Information*, pages 415–425, 2017.

[CO-2] A. Coutant et C. Rouveirol. Network Inference of Dynamic Models by the Combination of Spanning Arborescences. *Journées Ouvertes en Biologie, Informatique et Mathématiques (JOBIM 2017), Lille, France, 3-6 juillet 2017*, pages 45–52, 2017.

[CO-1] H. Léger, D. Bouthinon, M. Lebbah et H. Azzag. Nouveau modèle pour un passage à l'échelle de la 0-subsomption. *17ème Journées Francophones Extraction et Gestion des Connaissances (EGC 2017), 24-27 Janvier 2017, Grenoble, France*, volume E-33 de *Revue des Nouvelles Technologies de l'Information*, pages 339–344, 2017.

Direction d'ouvrages (DO)

Livres

2019 [**LI-1**] I. Redko, A. Habrard, E. Morvant, M. Sebban et Y. Bennani. *Advances in Domain Adaptation Theory*. Elsevier, 2019.

Édition d'ouvrages collectifs

[ED-2] E. Banissi, A. Ursyn, M. W. M. Bannatyne, N. Datia, R. Francese, M. Sarfraz, T. G. Wyeld, F. Bouali, G. Venturini, H. Azzag, M. Lebbah, M. Trutschl, U. Cvek, H. Müller, M. Nakayama, S. Kernbach, L. Caruccio, M. Risi, U. Erra, A. Vitiello et V. Rossano. *The 23rd International Conference on Information Visualisation (IV 2019), Paris, France, July 2-5, 2019, Part I*. IEEE, 2019.

2018 [ED-1] M. Lebbah, C. Largeron et H. Azzag. *18ème Conférence Internationale sur Extraction et Gestion des Connaissances (EGC 2018), Paris, France, January 22-26, 2018*, volume E-34 de *Revue des Nouvelles Technologies de l'Information*. Hermann-Éditions, 2018.

Logiciels et Ressources (LO)

2022 [LO-13] M. Hamidi et A. Osmani. Reduction of the Position Bias via Multi-Level Learning for Activity Recognition. https://github.com/alphaequivalence/positionAbstraction, 2022.

[LO-12] S. M. R. Modaresi, A. Osmani, M. Razzazi et A. Chibani. EvalSeg : Multi Modal Evaluation For Medical Image Segmentation, 2022.

[LO-11] S. M. R. Modaresi, A. Osmani, M. Razzazi et A. Chibani. UnifiedAR : A Modular Platform for Activity Recognition from Sensor Data, 2022.

2021 [LO-10] F. Forest, M. Lebbah, H. Azzag et J. Lacaille. DESOM : Deep Embedded Self-Organizing Map model. https://github.com/FlorentF9/DESOM, 2021.

[LO-9] E. Goffinet, M. Lebbah et H. Azzag. MultiCoclustering : 3 méthodes de CoClustering. https://github.com/EtienneGof/MultiCoclustering, 2021.

[LO-8] M. Hamidi et A. Osmani. Augmented Experiments in Material Engineering Using Machine Learning. https://github.com/alphaequivalence/nx-project, 2021.

[LO-7] M. Hamidi et A. Osmani. Hierarchical Learning of Dependent Concepts for Human Activity Recognition. https://github.com/alphaequivalence/hierarchicalSHL, 2021.

[LO-6] S. M. R. Modaresi, A. Osmani, M. Razzazi et A. Chibani. Auto Profiler : Automatic Interactive Tree-based Profiling of Python Scripts in Jupyter Notebook. https://github.com/modaresimr/auto_profiler, 2021.

[LO-5] S. M. R. Modaresi, A. Osmani, M. Razzazi et A. Chibani. ConfJournalRank : Integrated Multi SourcesConference and Journal Information". https://rank.modaresi.eu.org, 2021.

[LO-4] S. M. R. Modaresi, A. Osmani, M. Razzazi et A. Chibani. JCV2 : Enabling Compatibility of OpenCV UI Components with Jupyter Notebooks. https://github.com/modaresimr/opencv_jupyter_ui, 2021.

2020 [LO-3] A. Coutant, E. Goffinet, M. Lebbah et H. Azzag. spark-tss : une API pour le prétraitement de séries temporelles multi-variables. https://github.com/spark-tss/spark-tss, 2020.

[LO-2] M. Hamidi et A. Osmani. Learning Mobility-Related Human Activities from Sensor-Rich Environments. https://github.com/HamidiMassinissa/shl-nas, 2020.

2018 [LO-1] G. Beck, M. Lebbah, H. Azzag et A. Coutant. Clustering4Ever : une API pour le clustering. https://github.com/Clustering4Ever, 2018.

Thèses et Habilitations (TH)

Thèses

2022 [TH-18] M. Chevallier. L'Apprentissage Artificiel au service de la qualité de données : détection de contraintes. Thèse, Université Sorbonne Paris Nord, 2022.

[TH-17] M. El Hamri. Structural Optimal Transport for Domain Adaptation with Theoretical Guarantees. Thèse, Université Sorbonne Paris Nord, 2022.

[TH-16] Y. Foucade. Learning from Data and Learners. Thèse, Université Sorbonne Paris Nord, 2022.

[TH-15] M. Hamidi. Metaapprentissage guidé par les connaissances du domaine. Thèse, Université Sorbonne Paris Nord, 2022. Sous la direction de Aomar OSMANI.

[TH-14] A. Zaiou. Quantum Machine Learning Approaches for Graphs and Sequences : Application to Nuclear Safety Assessment. Thèse, Université Sorbonne Paris Nord, 2022.

2021 [TH-13] M. O. Attaoui. Vers de Nouvelles Méthodes de Clustering de Flux de Données. Thèse, Université Sorbonne Paris Nord, 2021.

[TH-12] F. Forest. Apprentissage non supervisé de représentations de données et structures departitionnement : applications à la surveillance à grande échelle de turbofans. Thèse, Université Sorbonne Paris Nord, 2021.

[TH-11] E. Goffinet. Clustering Multi-Blocs et Visualisation Analytique de Données Séquentielles Massives Issues de Simulation du Véhicule Autonome. Thèse, Université Sorbonne Paris Nord, 2021.

2020 [TH-10] F. E. Ben Bouazza. Multi-models clustering through optimal transport theory. Thèse, Université Sorbonne Paris Nord, 2020.

[TH-9] K. Benlamine. Apprentissage non supervisé profond quantique de représentations de données évolutives. Thèse, Apprentissage non supervisé profond quantique de représentations de données évolutives, 2020.

[TH-8] S. Zouinina. Data Anonymization through Unsupervised Learning. Thèse, Université Sorbonne Paris Nord, 2020.

2019 [TH-7] G. Beck. Scalable Clustering Applying Local Accretions. Thèse, Université Sorbonne Paris Nord, 2019.

[TH-6] H. Léger. Passage à l'échelle de la theta-subsomption. Thèse, Université Sorbonne Paris Nord, 2019.

[TH-5] M. Rifi. Modélisation et Analyse des Réseaux Complexes : application à la sûreté nucléaire. Thèse, Université Sorbonne Paris Nord, 2019.

2018 [TH-4] I. Falih. Attributed Network Clustering : Application to recommender systems. Thèse, Université Sorbonne Paris Cité, 2018.

[TH-3] P. Rastin. Apprentissage non-supervisé à partir de données évolutives. Thèse, Université Sorbonne Paris Cité, 2018.

[TH-2] T. Sarazin. Apprentissage distribué de grandes masses de données. Thèse, Université Sorbonne Paris Cité, 2018.

[TH-1] L.-M. Veillon. Apprentissage artificiel collectif, aspects dynamiques et structurels. Thèse, Université Sorbonne Paris Cité, 2018.

Habilitations

2021 [HR-3] N. Rogovschi. Contributions to data dimensionality reduction and co-clustering. HdR, Université Sorbonne Paris Nord, 2021.

[HR-2] J. Sublime. Contributions to modern unsupervised learning : Case studies of multi-view clustering and unsupervised Deep Learning (Contributions à l'apprentissage non-supervisé moderne : Applications aux cas du clustering multi-vue et de l'apprentissage profond non-supervisé). HdR, Université Sorbonne Paris Nord, 2021.

2020 [HR-1] N. Grozavu. Contributions on Collaborative Unsupervised. HdR, Université Sorbonne Paris Nord, 2020.

Publications antérieures à l'arrivée au LIPN

Les références comportent « ↑ » et correspondent à des publications antérieures à l'arrivée au LIPN, non référencées par ailleurs.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

3.3 Domaine 4. Inscription des activités de recherche dans la société

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique.

L'équipe a plusieurs collaborations industrielles qui permettent à ces membres de financer les doctorants et les diverses actions pour la recherche. Durant la période entre 2017 et 2022, l'équipe a eu 19 projets avec des acteurs socio-économiques variés (SAFRAN, Renault, EDF, FaceBook, Nukkai, etc.). L'ensemble de ces contrats représente plus de 2 millions d'euros (voir section 3.1.2).

3.3.2 Référence 2. L'équipe développe des produits à destination du monde culturel, économique et social.

Le laboratoire met tout en œuvre pour accompagner les initiatives de valorisation et de transfert comme il a pu le prouver avec l'accompagnement actif de la startup HephIA.

Suite au succès de plusieurs collaborations industrielles et à la participation à plusieurs projets investissement d'avenir (PIA), ANR et CIFRE, fin 2020, un membre de l'équipe A³ ainsi qu'un ancien doctorant et un ancien postdoctorant de l'équipe ont porté un projet de création de startup devant la Société d'Accélération du Transfert de Technologies (SATT) Erganeo. L'obtention d'un financement de maturation (470k€) a permis de concrétiser cet objectif. L'objectif de HephIA est de cartographier les données massives de façon automatisée en détectant les informations les plus pertinentes à extraire. A la fin 2022, la taille de la startup HelphIA hébergée par l'équipe A^3 a atteint 14 membres.

Chapitre 4
Trajectoire de l'équipe A3

Sommaire

4.1 Organisation et ambitions de l'équipe 107
4.2 Analyse AFOM/SWOT 108
4.3 Perspectives et projet de recherche 109
4.3.1 Axe ADA : Apprentissage à partir de Données et d'Apprenants . . . 109
4.3.2 Axe MAR : Métaapprentissage et Apprentissage de Structures . . . 110
4.3.3 Axe ARG : Apprentissage Relationnel et Graphes 113

4.1 Organisation et ambitions de l'équipe

La structuration de l'équipe, sous sa forme actuelle, est récente. Les objectifs attendus sont une meilleure coordination des membres pour créer des synergies autour des nouveaux axes structurants de l'équipe. Cette coordination doit se centrer sur l'amélioration de la visibilité de nos activités, des recrutements centrés sur ces ambitions et la recherche de collaborations, notamment, internationales.
Des collaborations locales doivent aussi être renforcées. L'intelligence artificielle est devenue incontournable pour beaucoup d'activités de recherche, l'équipe renforcera

naturellement ses collaborations avec équipes RCLN, AOC et dans une moindre mesure avec les équipes LOVE et CALIN. L'équipe collabore aussi avec plusieurs équipes au sein de la fédération Mathstic à travers des coencadrement de doctorants et la participation à plusieurs projets de recherche. L'équipe souhaite, également, développer un projet interdisciplinaire d'intérêt général au niveau de l'université USPN. Nous discutons actuellement sur un projet autour de l'IA, l'apprentissage automatique et l'éthique. Ce type de projet nécessite des disponibilités de quelques membres de l'équipe.

Trajectoire de l'équipe A3

L'équipe a redéfini trois axes de recherche et souhaite limiter la diversification des thématiques afin de contribuer significativement sur ces thèmes et renforcer son expertise théorique. Dans chacun des axes, des problèmes fondamentaux sont étudiés en lien avec des domaines d'applications qui répondent aux besoins de l'environnement socio-économique essentiellement à travers la formation par la recherche de doctorants en CIFRE.
Malgré l'ambition affichée, le risque principal est la poursuite de l'affaiblissement de l'équipe avec des départs non compensés, notamment celui d'un rang A en 2017.

4.2 Analyse AFOM/SWOT

- **Forces et Atouts**
  - *Positionnement original et large spectre scientifique de l'équipe*, lui assurant une bonne visibilité extérieure sur ses thèmes de recherches ;
  - *Attractivité du thème et de l'équipe dans le monde académique et industriel* : nombreux doctorants, bon taux d'insertion académique et industriel à l'issue de la thèse, nombreux partenariats industriels (ANR) et industriels (FUI, PIA, CIFRE) ;
  - *Formation à la recherche* : adossement à un master très visible (EID2) ;
  - *Valorisation de la recherche* : transfert technologique par la création de start-ups :
  - *Nombreuses initiatives d'animation de la recherche* : organisation de conférences françaises, d'écoles thématiques, de tutoriels, de sessions spéciales dans des conférences internationales.

- **Faiblesses**
  - *Absence de chercheurs permanents* : la politique de recrutement de chercheurs CNRS doit être dynamisée, notamment à travers le recrutement de postdocs et un recours plus fréquent à des mois de chercheurs invités ;
  - *Relations internationales à intensifier*, notamment au niveau des contrats internationaux ;
  - *Diffusion de logiciels à renforcer* : des logiciels sont développés dans l'équipe, mais leur diffusion est encore trop restreinte, faute de personnel dédié au développement ;
  - *Concurrence* : la concurrence avec les unités parisiennes et les centres de recherche privés fragilise l'équipe (difficulté à recruter doctorants et permanents), malgré un positionnement fort et original ;
  - *arité de genre* : L'équipe compte peu de femmes.

- **Opportunités**
  - *Visibilité accrue de l'apprentissage (et de l'Intelligence Artificielle en général)* dans le grand public, ce qui peut susciter des vocations d'étudiants ;
  - *Renforcement de notre position à l'international* : les liens avec d'autres équipes et chercheurs à l'international doivent s'intensifier ;
  - *Large couverture des thèmes de l'apprentissage avec des positionnements originaux ;*

— *Développement de recherches transverse en apprentissage à travers des projets interdisciplinaires notamment avec les laboratoires MathSTIC et L2TI.*

- **Menaces**
  — *Malgré l'adossement à un M2*, les entreprises attirent de plus en plus de bons étudiants qui se détournent des carrières universitaires sur le domaine de l'apprentissage artificiel et de la fouille de données ;
  — *Les départs d'HDR déstabilisent certains thèmes de recherche qui risquent de disparaître ;*
  — *L'affaiblissement de l'équipe du fait du non-remplacement des départs.*

4.3 Perspectives et projet de recherche

Le projet de l'équipe sur la prochaine période poursuivra les travaux en cours en allant vers une meilleure intégration des activités et un renforcement de l'expertise en fondements théoriques de l'apprentissage. La réalisation de ce projet dépendra de l'évolution de l'équipe, notamment en termes de recrutement d'enseignants chercheurs permanents.

4.3.1 Axe ADA : Apprentissage à partir de Données et d'Apprenants

Transport Optimal et Analyse Toplogique pour l'Apprentissage Multi-modale : La théorie du transport optimal offre un cadre formel très puissant pour l'apprentissage artificiel. Trois thèses de doctorat sur cette thématique ont été soutenues. Le départ de ces diplômés ainsi que certains membres permanents de l'axe ADA, vers d'autres laboratoires académiques va réduire les activités autour de ce thème pourtant pertinent et présentant de nombreuses perspectives. Par ailleurs, suite au manque de forces locales, le thème portant sur l'apprentissage quantique ne sera plus poursuivi dans l'axe ADA. Néanmoins, nos contributions sur cette thématique très prometteuse a permis de proposer des approches très pertinentes et innovantes dans le domaine, à travers notamment deux thèses de doctorat soutenues au $LIPN$, dont une dans le cadre d'une $CIFRE$ avec EDF. Par conséquent, et suite à l'évolution des questions scientifiques autour de l'apprentissage, le projet que nous présentons pour l'axe ADA sera resserré uniquement autour d'un seul thème, ce qui permettra d'animer plus efficacement l'axe. Les données sont souvent présentées dans un format qui couvre plusieurs modalités, comme le matériel vidéo, qui comprend normalement au moins une image et du son (voix, musique), ainsi que du texte sous forme de sous-titres, qui sont souvent dans une langue autre que les données parlées. D'autres situations, comme la détection humaine, fournissent des données dans plusieurs modalités, par exemple des données sur l'expression du visage sous forme de photographies mélangées à des données auditives (voix, son), haptiques (toucher) ou autres données sensorielles. Le défi de la recherche sur l'analyse des données multimodales est donc axé sur l'intégration et l'interprétation des données dans et entre les modalités, ainsi que les connaissances et idées qui en découlent. Les résultats de cette recherche permettront d'améliorer la compréhension et la modélisation de sources de données riches dans des domaines tels que les affaires, la santé, l'environnement et l'éducation. Le problème fondamental est d'apprendre à décrire et à résumer les données multimodales d'une manière qui exploite la complémentarité et la redondance des différentes modalités. Pour ce faire, des approches fondées sur le transport optimal et l'analyse toplogique des données (TDA : Topological Data Analysis) seront dévelopées.

4.3.2 Axe MAR : Métaapprentissage et Apprentissage de Structures

Trajectoire de l'équipe A3

Modèles génératifs et Multi-Coclustering : Nous souhaitons poursuivre l'étude des modèles génératifs pour les données temporelles multivariées [CO-26, CO-18]. Pour résoudre le problème de données manquantes, nous proposons d'utiliser des modèles génératifs (GAN) ceci dans le contexte d'une thèse $CIFRE$ en cours avec $Safran$. Des travaux récents proposent des TTS-GAN basés sur les transformers et constituent une autre méthode permettant de résoudre en grande partie le problème de disparition du gradient qui entrave les RNNs en utilisant des neurones basés sur l'attention . Dans le même sens, dans le cadre du pôle $MathStic$, en collaboration avec le laboratoire L2ti (laboratoire de Traitement et Transport de l'Information de Sorbonne Paris Nord) renforcée par le LabCom $IRISER$, nous étudions l'apport des mécanismes d'attention dans le traitement des données multimodales (images et textes). Une des techniques permettant de prendre en compte la multimodalité est d'agir au niveau du processus de normalisation. Ces travaux seront également renforcés par un projet IHU qui a été accepté récemment (courant Mai 2023) dont l'objectif est de créer un écosystème multidisciplinaire alliant des forces de recherche fondamentale et clinique pour analyser et réduire l'impact (décès et séquelles) de la maladie du sepsis.

En ce qui concerne le *Multi-Coclustering*, pour tenir compte des données non euclidiennes, nous souhaitons utiliser une approche à bloc latent bayésien non paramétrique. Nous envisageons d'étendre nos activités dans le domaine du *Multi-Coclustering* [CI-87] pour prendre en compte des données multivues et multimodales provenant de différentes sources (image, texte, etc). Nous proposons également de prendre en compte le principe de l'adaptation des modèles topologiques aux données multivues, nous proposons de travailler sur de nouvelles approches topologiques du *Multi-CoClustering* (*MCC Topologique*) qui regroupe les variables partageant la même partition d'observations et la même vue [18]. Un de nos principaux objectifs sera d'étudier l'explicitation des résultats du *Multi-CoClustering Topologique*, c'est-à-dire permettre un retour visuel le plus efficace du *clustering* [TH-7]. Ces travaux seront soutenues par une thèse financée par le programme $DIM\ AI4IDF$ en collaboration avec l'UVSQ Paris-Saclay qui a débutée en Janvier 2023.

Par ailleurs, les applications basées sur l'apprentissage profond ont connu un fort succès ces dernières années. Poursuivre nos activités en apprentissage profond et mécanismes d'attention est une des perspectives à court termes de l'équipe. Notamment, pour encoder/vectoriser les données brutes non formatées (séries temporelles, images, etc.). Les données disponibles pour le clustering sont issues de sources hétérogènes et déséquilibrées que nous souhaitons explorer pour la détection de fraude dans les données bancaires dans le cadre d'une collaboration $CIFRE$ avec la $BPCE$.

Métaaprentissage : En métaapprentissage, la minimisation du nombre d'exemples pour apprendre, la quantification de tous les biais d'apprentissage et la certification ou la quantification de la qualité de l'apprentissage font partie des objectifs des activités de l'équipe. L'intégration des connaissances du domaine passent par la recherche de prédicats génériques et d'invariants (plus généralement régularités) génériques indépendants du domaine [**CI-97**]. Il peut s'agir de modèles équationnels, de topologies ou de toute forme de régularités sur les processus ou les structures qui minimisent l'ensemble des fonctions admissibles [21]. On utilise des applications à base de capteurs, car elles présentent un cadre complet pour l'étude des problèmes du métaapprentissage. Ces problèmes, comme nous l'avons montré dans le papier qui synthétise les biais en apprentissage [RI-18] sont relativement complets. Ils portent sur diverses multimodalités (images, son, accélération, ...), la relativité des données [**CI-98**], les combinaisons de capteurs [**CI-52**], le transfert des données (traitement séquentiel, parallèle des données ou synthèse du transfert de l'apprentissage), la coexistence de multiples théories d'apparence contradictoires (correspondant aux points de vue différents) ou

aux mécanismes de conciliation [CI-97], la structuration et le traitement des données, le temps réel et l'intégration des questions liées à la consommation d'énergie, au débit et aux capacités de traitement locales, à l'écologie, à la protection des données et à l'éthique qui doivent aussi être traités par les approches du métaapprentissage. Deux véhicules, conçus par deux constructeurs différents, peuvent, par exemple, utiliser les mêmes données pour aboutir à des théories inconsistantes, mais qui coexistent. La collaboration dans ce cas n'est pas centrée sur la théorie générale à apprendre mais sur d'autres objectifs externes : ici, éviter les accidents et fluidifier la circulation, par exemple. L'expressivité du modèle de métaapprentissage doit être capable d'englober la formalisation de ce type de problèmes.

Il en est de même pour des questions qui portent, par exemple, sur l'interprétation, l'explication de modèles d'apprentissages ou par la nécessité qu'ils respectent des normes(loi, éthique, etc.). Il s'agit de concepts à traiter par l'apprentissage, mais à d'autres niveaux d'abstraction, comme c'est le cas, par exemple, des modèles MOF en métamodélisation. La modélisation du métaaprentissage doit gérer les hyperparamètres qui les définissent et gèrent leurs liens avec les algorithmes d'apprentissage. Prenons l'exemple de l'explication : au niveau du modèle abstrait du métapparentissage, l'explication peut se définir comme un problème d'optimisation et. d'apprentissage de représentations entre trois espaces : la représentation de la théorie apprise, la représentation de l'acteur à qui l'explication est destinée et la représentation d'un espace intermédiaire qui est un simple changement de représentation de la théorie apprise avec le moins de perte possible tout en s'approchant le plus possible de la représentation de l'acteur. Du point de vue de la spécification concrète, il s'agit d'un jeu d'optimisation de métriques entre ces trois espaces de représentations. Le métaapprentissage doit donner un cadre pour expliciter ces objets et les leurs interrelations. Il en est de même pour l'éthique. Dans le cas de la recherche de solutions d'apprentissage automatique qui respectent l'éthique, la décomposition du problème par le métaapprentissage le rend plus facile à résoudre. Initialement, il convient de respecter un ensemble de principes (implémentés par des règles) décrit dans un langage de représentation donné pour lesquels il convient de trouver une bonne représentation qui minimise la perte par rapport au langage initial et qui permet d'être traduite dans la représentation de la théorie apprise, par les modèles d'apprentissage, pour contraindre les réponses à celles admises par l'éthique. Du point de vue opérationnel, il existe plusieurs manières de faire. L'une d'elle consiste à incorporer ces règles lors de la conception du système. Soit via des coefficients de régularisation (solution souvent utilisée par les réseaux de neurones profonds), soit via des modèles proches de ceux sur lesquels nous travaillons : recherche de la modularité et d'invariants dans les frontières de décisions de sorte à définir des composants autonomes évaluables et composables. Ce procédé permet d'assurer une transparence des boites noires en contrôlant les entrées et sorties des composants élémentaires les constituant, et en intégrant les contraintes (dont celles de l'éthique, dans ce cas) par le contrôle des échanges entre ces composants/Services élémentaires. La deuxième manière de faire est totalement exogène, elle consiste à considérer la théorie apprise comme une boite noire et à développer des outils d'audit externes et indépendants permettant de garantir le test et la validation des règles que l'on souhaite faire imposer au système.
Ces types de structures et de contrôles nécessitent de construire un modèle de métaaprentissage global permettant de formaliser, décrire, instancier et exécuter ce type de scénarios. Construire une telle architecture est complexe. Elle impose la recherche d'une simplicité de la métamodélisation qui doit comporter très peu de prédicats. Cette recherche de la simplification passe par un élément fondamental, la définition du système d'évaluation qui structure la définition du pipeline de métaapprentissage de bout en bout. Dans cette définition du métaapprentissage, l'algorithme (ou les algorithmes) n'est qu'un hyperparamètre parmi d'autres. La recherche d'une "bonne" architecture nécessite de se poser les questions non limitatives à propos des problèmes suivants (1) les données qui sont rarement des faits, (2) les propriétés intrinsèques du problème et à

la structure des solutions disponibles pour le résoudre, (3) la modélisation explicite de tous les biais et leur quantification et (4) la structure de l'évaluation qui impose le plus de biais implicites. En plus de ces objectifs à court terme, l'équipe souhaite traiter des questions plus théoriques liées au métaapprentissage telles que :

Trajectoire de l'équipe A3

Caractérisation du paysage d'optimisation Cette perspective regroupe les travaux sur l'étude des régularisations dans l'espace des biais qui permettent de structurer les solutions à la fois pour faciliter le transfert, permettre une modularité (ou un apprentissage par composants) facilitant ainsi la transparence des résultats et ouvrant une voie pour l'explication, l'interprétation et le contrôle explicite des résultats (imposer le respect de la réglementation, de l'éthique, de standards, etc). En particulier,

- La connaissance du domaine a un impact important sur le paysage de l'optimisation pour faciliter l'exploration. C'était le cas dans notre travail sur l'utilisation des invariants du domaine sous forme de symétries entre les différentes vues fournies par les capteurs ayant une distribution dynamique, mais structurée et régulière. Comme le rappelle [23], *Les symétries des modèles d'observation deviennent des symétries du problème d'optimisation*. De plus, comme nous l'avons montré [**CI-97**, **RI-26**], la manière dans le problème d'apprentissage est structurée en sous-problèmes en interaction via le transfert, le partage et la réutilisation, se traduit également dans le paysage d'optimisation par des propriétés qui semblent intéressantes qu'il convient d'explorer. Une voie possible est d'explorer le lien qui existe entre les limites de décisions du paysage d'optimisation et les invariants explicites du domaine.
- Les connaissances a priori ajoutées à une représentation explicite du paysage d'optimisation peuvent-elles conduire à des sauts inductifs substantiels dans l'espace d'optimisation ? Peut-on utiliser les approches d'optimisation globales en présence de connaissances structurelles à la place des méthodes locales ? Ceci permettrait une meilleure structuration des biais des algorithmes et faciliterait le transfert d'hypothèses,
- Les travaux de Vapnik sur les prédicats à travers les invariants suggèrent qu'il existerait un petit nombre de prédicats qui structureraient les connaissances du domaine. C'est le cas notamment en utilisant les groupes de symétries comme les groupes de Lie que nous comptons poursuivre pour encoder des connaissances du domaine [19]. Des travaux récents généralisent ce problème d'apprentissage au calcul variationnel nécessitant une optimisation des paramètres dans l'espace courbe [19]. Une voie prometteuse est l'optimisation sur des variétés lisses ou Riemanniennes, ce qui revient à faire l'optimisation sur une structure de variété connue [16]. Ceci pourrait avoir comme conséquence d'améliorer le processus d'optimisation car il existe une vaste littérature sur les méthodes d'optimisation plus adaptées à ces variétés [1, 2, 3, 22, 13, 12] qui étendent les algorithmes de convergence classiques du cadre Euclidien au cadre Riemanien.

Recherche de correspondances entre connaissances du domaine et les paramètres des modèles Dans le cas des modèles symboliques, cette correspondance est plus simple à établir. Elle est plus difficile dans le cas des approches connexionnistes. Le choix de certaines architectures imposent des contraintes aux modèles et les formes à apprendre dans des sous-espaces contraints comme nous l'avons fait dans nos travaux [**CI-98**]. Dans les derniers travaux de V. Vapnik [21], il résume cette tâche à la recherche de prédicats encodant des régularités (invariants, etc.), à la recherche de fonctions admissibles et à la convergence faible. Les méthodes qui permettent de capturer naturellement des aspects précis du domaine permettront de construire des apprenants robustes et évolutifs. Par exemple, existe-t-il une correspondance entre les modèles d'apprentissage et leurs capacités de calcul et la connaissance du domaine exprimée dans un langage spécifique, à l'image de la correspondance Curry et Howard [10] entre les modèles de calcul et les systèmes de preuves ? Peut-on concevoir des mo-

dèles qui encodent, dans leurs mécanismes et composants internes, des biais inductifs correspondant aux connaissances du domaine ? Il existe plusieurs travaux dans ce sens pour tenter de combler l'aspect boite noire des modèles neuronaux [6, 7, 9]. Ces systèmes suggèrent, comme c'est le cas par exemple, dans les architectures logicielles, que chaque composant neuronal soit responsable d'une connaissance spécifique, cette modularité, comme nous avons pu le faire dans [**CI-86**, **CI-97**] et [**CI-98**] est une voie possible pour la transparence et l'explication.

4.3.3 Axe ARG : Apprentissage Relationnel et Graphes

Fouille de graphes attribués, caractérisation des structures au-delà des énumérations Nos travaux portant sur la fouille de graphes attribués nous ont permis de proposer i) une approche d'énumération réduisant l'espace d'exploration aux seuls motifs clos supportés par un ensemble de sommets dont le sous-graphe induit vérifie une propriété de coeur et ii) une approche de sélection réduisant le nombre de motifs en assurant que les motifs présentés soient suffisamment différents. Si ces restrictions sont intéressantes du point de vue de l'exploitation humaine des résultats en limitant le nombre de motifs à considérer, elles induisent nécessairement une perte d'information par rapport au graphe attribué. Dans un premier temps, nous chercherons à quantifier et à limiter cette perte d'information en proposant des représentations augmentées des motifs actuels qui puissent par exemple prendre en compte des informations statistiques sur les motifs exclus lors des sélections. Nous chercherons ensuite à mettre en place des méthodes d'exploitation des motifs énumérés, par exemple, pour des systèmes de recommandation en considérant les distances en intention et en extension des points exclus de l'ensemble support des motifs retenus. Nous chercherons à valider ces approches dans le cadre de collaborations en cours de constitution en les appliquant à des graphes de terrain élaborés à partir de données d'enquêtes de type ethnographiques.

Explications pour la prise de décision séquentielle et en monde partiellement observable Nous nous sommes intéressés dans des travaux récemment acceptés à la notion d'explication dans des langages expressifs (restrictions de la logique des prédicats) pour problèmes de décision séquentiels modélisés par des processus de décision de Markov (MDP)[17] . Nous avons proposé une formalisation et des algorithmes de Programmation Logique Inductive [5] pour réinterpréter la notion d'explication *abductive* de P. Marquis [14] et proposer des explications communes minimales pour un groupe d'exemples partageant une même classe. Ces travaux peuvent être étendus dans de nombreuses directions, une première étant de relever les défis liés aux différents aspects à la complexité de calculer de telles explications en logique d'ordre un (borner le nombre d'explications abductives/ la complexité du calcul de ces explications). Une autre extension que nous envisagerons est de produire des explications de trajectoires dans un cadre partiellement observable. Cette extension ouvre de nombreuses perspectives d'applications, car dans de nombreuses situations réelles, la prise de décision séquentielle se fait dans un environnement partiellement observable (POMDP)[17] , c'est-à-dire que tout ou partie des états sur les trajectoires étudiées – dont des éléments éventuellement essentiels pour la prise de décision – ne sont pas observés. Par exemple, dans un jeu de cartes, les mains des adversaires ne sont pas observées. Elle pose également de nombreux défis théoriques et algorithmiques.

Dans le cadre qui nous intéresse, une théorie du domaine (modèle partiel de l'environnement) nous permettra de raisonner et de calculer un *état de croyance*, i.e. une distribution de probabilité sur les états possibles satisfaisant les contraintes de la théorie du domaine. De nombreux formalismes et outils ont été conçus et développés ces dernières années pour faire du raisonnement et de l'apprentissage dans un cadre incertain (mentionnons [15] et plus récemment [20]) , nous allons pouvoir nous appuyer sur ces formalisations et outils développés en logique des prédicats probabiliste pour étendre notre cadre logique d'explications en logique des prédicats pour des POMDPs.

Trajectoire de l'équipe A3

Au-delà des aspects techniques, construire une explication dans un cadre POMDP est un problème ouvert. On peut l'exprimer comme un problème apprentissage d'une politique explicite dans un cadre POMDP et tenter de généraliser des travaux sur l'élicitation de politiques dans un cadre MDP [8, 4] à un cadre partiellement observable. Un article de J. Lang et B. Zanuttini [11] propose de représenter des politiques dans un cadre partiellement observable avec des *Knowledge Based Probabilistic programs*, qui référencent explicitement à un état de croyance de l'agent "décideur". De tels *Knowledge Based Probabilistic programs* exprimés sous la forme de programme logiques probabilistes peuvent servir à exprimer des politiques à suffisamment haut niveau pour être compréhensibles et transférables. Apprendre de tels programmes est donc une voie prometteuse pour construire des explications que nous souhaitons. Nous évaluerons les approches proposées notamment dans le domaine des jeux de cartes et du Bridge, en collaboration avec la société Nukkai, avec laquelle ces approches ont été initiées.

Références complémentaires

[1] P.-A. Absil, R. Mahony, and R. Sepulchre. Optimization algorithms on matrix manifolds. In *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2009.

[2] S. Bonnabel. Stochastic gradient descent on riemannian manifolds. *IEEE Trans. on Automatic Control*, 58(9), 2013.

[3] N. Boumal, P.-A. Absil, and C. Cartis. Global rates of convergence for nonconvex optimization on manifolds. *IMA Journal of Numerical Analysis*, 39(1) :1–33, 2019.

[4] Y. Coppens, D. Steckelmacher, C. M. Jonker, and A. Nowé. Synthesising reinforcement learning policies through set-valued inductive rule learning. *CoRR*, abs/2106.06009, 2021.

[5] A. Cropper and S. Dumancic. Inductive logic programming at 30 : A new introduction. *J. Artif. Intell. Res.*, 74 :765–850, 2022.

[6] L. De Raedt, A. Kimmig, and H. Toivonen. Problog : A probabilistic prolog and its application in link discovery. In *IJCAI*, volume 7, pages 2462–2467. Hyderabad, 2007.

[7] L. De Raedt, R. Manhaeve, S. Dumancic, T. Demeester, and A. Kimmig. Neuro-symbolic= neural+ logical+ probabilistic. In *NeSy'19@ IJCAI, the 14th International Workshop on Neural-Symbolic Learning and Reasoning*, 2019.

[8] S. Džeroski, L. De Raedt, and K. Driessens. Relational reinforcement learning. *Machine Learning*, 43(1) :7–52, Apr 2001.

[9] P. Hitzler, A. Eberhart, M. Ebrahimi, M. K. Sarker, and L. Zhou. Neuro-symbolic approaches in artificial intelligence. *National Science Review*, 9(6) :nwac035, 2022.

[10] W. A. Howard. The formulae-as-types notion of construction. *To HB Curry : essays on combinatory logic, lambda calculus and formalism*, 44 :479–490, 1980.

[11] J. Lang and B. Zanuttini. Probabilistic knowledge-based programs. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pages 1594–1600, 2015.

[12] M. Lezcano-Casado and D. Martinez-Rubio. Cheap orthogonal constraints in neural networks : A simple parametrization of the orthogonal and unitary group. In *International Conference on Machine Learning*, pages 3794–3803. PMLR, 2019.

[13] Y. Liu, F. Shang, J. Cheng, H. Cheng, and L. Jiao. Accelerated first-order methods for geodesically convex optimization on riemannian manifolds. *Advances in Neural Information Processing Systems*, 30, 2017.

[14] P. Marquis. Extending abduction from propositional to first-order logic. In *FAIR*, volume 535, pages 141–155. Springer, 1991.

[15] L. D. Raedt, K. Kersting, S. Natarajan, and D. Poole. *Statistical Relational Artificial Intelligence : Logic, Probability, and Computation*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2016.

[16] W. Ring and B. Wirth. Optimization methods on riemannian manifolds and their application to shape space. *SIAM Journal on Optimization*, 22(2) :596–627, 2012.

[17] R. S. Sutton and A. G. Barto. *Reinforcement learning - an introduction, Second Edition*. MIT Press, 2018.

[18] T.-J.-Y.-S.-G. O.-M. Takamura-Y-Okamoto-S-Yamawaki-and-K-Doya. Multiple co-clustering based on nonparametric mixture models with heterogeneous marginal distributions. In *PLOS ONE 12(10)*, oct-2017.

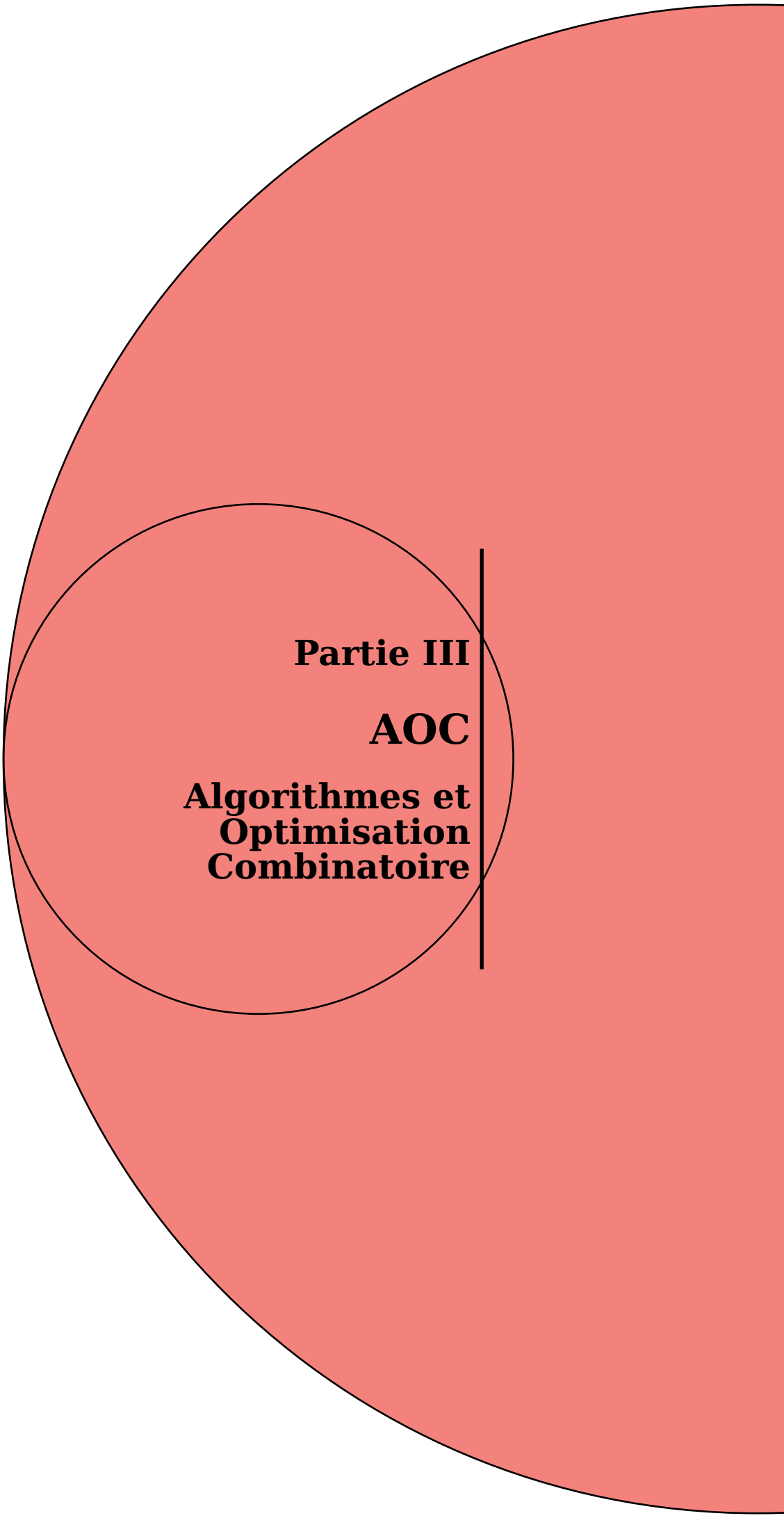
[19] M. Tao and T. Ohsawa. Variational optimization on lie groups, with examples of leading (generalized) eigenvalue problems. In *International Conference on Artificial Intelligence and Statistics*, pages 4269–4280. PMLR, 2020.

[20] G. Van den Broeck, K. Kersting, S. Natarajan, and D. Poole. *An Introduction to Lifted Probabilistic Inference*. MIT Press, aug 2021.

[21] V. Vapnik and R. Izmailov. Complete statistical theory of learning : learning using statistical invariants. In *Conformal and Probabilistic Prediction and Applications*, pages 4–40. PMLR, 2020.

[22] H. Zhang, S. J Reddi, and S. Sra. Riemannian svrg : Fast stochastic optimization on riemannian manifolds. *Advances in Neural Information Processing Systems*, 29, 2016.

[23] Y. Zhang, Q. Qu, and J. Wright. From symmetry to geometry : Tractable nonconvex problems. *arXiv preprint arXiv :2007.06753*, 2020.



Partie III

AOC

Algorithmes et Optimisation Combinatoire

Sommaire

Informations générales sur l'équipe AOC **123**
1.1 Membres de l'équipe AOC . . . 124
1.2 Les thématiques scientifiques et leurs enjeux . . . 126
1.3 Profil d'activités liées à la recherche . . . 134
1.4 Prise en compte des recommandations du précédent rapport . . . 134

Introduction du portfolio **137**

Auto-évaluation de l'équipe AOC **139**
3.1 Domaine 2. Attractivité . . . 140
3.2 Domaine 3. Production scientifique . . . 146
3.3 Domaine 4. Inscription des activités de recherche dans la société . . . 160

Trajectoire de l'équipe AOC **161**
4.1 SWOT . . . 162
4.2 Perspectives sur les prochaines années . . . 163

Overview

The AOC team was issued in 2010 from two research domains of the former OCAD team: Combinatorial Optimization, and Parallel and Distributed Computing. The team is organized along three axes: *Graphs and Polyhedra*, *Mathematical Programming* and *Parallel and Distributed Computing*. These three research axes are strongly linked together. The first one mainly deals with optimization problems in graphs with a focus on polyedral theory, approximation and complexity. One main topic of the axis is the use of geometric and algebraic tools from polyhedral theory to derive new insights towards the essential properties of the underlying combinatorial problem. Such problems can rarely be solved by enumeration, and their study often requires unveiling structural properties of the sets of solutions. The second axis deals with the design and analysis (both theoretical and empirical) of mathematical programming approaches and algorithms. Both exact and heuristic approaches for solving hard combinatorial optimization problems are developed. Moreover, both mixed integer linear programming and mixed integer non linear programming are studied. The third axis deals with problems intersecting three strongly linked domains: *distributed middleware and architecture*, *distributed programming* and *distributed algorithms*.

An interesting characteristic of the AOC team is that its expertise covers a broad spectrum of topics either horizontally (from graph theory to heuristics) or vertically (from algorithm design to detailed implementation). It is unusual in Europe and all over the world to find an equivalent spectrum. The strengths of the team are the quantity and quality of publications together with open source software development. Researchers of the team have published in high-level international journals and conferences as, for example, *Mathematical Programming*, *Operations Research*, *Mathematical Programming Computations*, *Journal of Supercomputing*, and *EURO-Par* and *IPCO*, and are editors or guest editors of several international journals on these topics. As for the software and/or instances collection implementation, we developed, for example, *BigCrunch* (i.e., software) and *QPLib* (i.e., instances). The team did actively participate in national projects (in some cases, as principal coordinator) and has several international collaborations. The international expertise of the team is also assessed by its ability to attract international researchers as visiting scholars or to cover temporary positions (PhDs and postdocs).

The research topics of the team for the next five years aim at investigating promising areas where the team may emerge as a leader. At the same time our researchers can give fruitful contributions in these areas thanks to their know-how and the cross fertilization of different axes. The topics are listed below, but we shall also continue our research in our major areas of expertise like: *polyedral, exact and heuristic approaches for hard combinatorial optimization problem*, *cloud and grid*, *high performance computing*, *quadratic programming*. These promising research topics are the following:

— Polyhedra behind min-max relations;
— Machine learning and optimization;
— Robustness;
— Sustainable digital and Edge computing

The researchers of the team are involved in teaching at the undergraduate and graduate level, in Sorbonne Paris Nord University and in other institutions.

Chapitre 1

Informations générales sur l'équipe AOC

Sommaire

1.1 Membres de l'équipe AOC 124
1.1.1 Membres permanents 124
1.1.2 Membres associés 125
1.1.3 Doctorants au 1er oct. 2022 125
1.1.4 Doctorants associés au 1er oct. 2022 125
1.1.5 Post-Doctorants au 1er oct. 2022 125
1.1.6 Personnel temporaire du 01/09/2017 au 01/10/2022 125
1.1.7 Évolution de l'équipe (permanents) 126
1.2 Les thématiques scientifiques et leurs enjeux 126
1.2.1 Axe 1: Graphes et polyèdres 127
1.2.2 Axe 2: programmation mathématique 129
1.2.3 Axe 3 : algorithmes, logiciel et architectures distribués 131
1.3 Profil d'activités liées à la recherche 134
1.4 Prise en compte des recommandations du précédent rapport . . . 134

Nom de l'équipe : Algorithmes et Optimisation Combinatoire (AOC)

Responsable de l'équipe : Roberto WOLFLER CALVO

1.1 Membres de l'équipe AOC

Corps	2017	2018	2019	2020	2021	2022	(2023)
PU	3	3	3	4	4	5	(5)
MCF	10	10	10	9	9	8	(8)
Post-docs et ATER	2	1	1	0	0	1	(1)
Doctorants	4	5	4	6	7	10	(9)
Prof ext.	0	0	1	1	0	0	0
Total	19	19	19	20	20	24	(23)

Informations générales sur l'équipe AOC

1.1.1 Membres permanents

Membres ou, *en italique, anciens membres*, et, **en gras, responsable de l'équipe** au 01/10/2022

	Nom	*Prénom*	*Situation*	*Institution*
	Borne	Sylvie	MCF	IG
	Butelle	Franck	MCF	IUTV
	Cérin	Christophe	PU	IUTV
	Coti[1]	Camille	MCF	IUTV
	Dubacq	Jean-Christophe	MCF	IUTV
	Fouilhoux	Pierre	PU	IG
	Grappe	Roland	MCF	IG
	Lacroix	Mathieu	MCF	IUTV
	Létocart	Lucas	PU	IG
	Pekergin	*Ferhan*	*MCF*	*IUTV*
	Roupin	Frédéric	PU	IG
	Toulouse	Sophie	MCF	IG
	Traversi	Emilanio	MCF	IUT
	Wolfler Calvo	**Roberto**	PU	IG

1. En détachement, poste au Canada

1.1.2 Membres associés

	Nom	*Prénom*	*Situation*	*Institution*
	Abbes	Heithem	MCF	**FST!**
	Alfandari	Laurent	PR	ESSEC
	Benallouche	Yacine	PAST	IUTV
	Finta	Lucian	MCF	IUTV
	Saad	Walid	MCF	**ENSIT!**
	Plateau	Gérard	membre d'honneur	

Informations générales sur l'équipe AOC

1.1.3 Doctorants au 1er oct. 2022

Nom	*Prénom*	*Situation*	*Financement*
Demelas	Francesco	Allocataire	Université SPN
Dupont-Bouillard	Alexandre	Allocataire	Université SPN
Grenèche	Nicolas		DSI-Université SPN
Pisanu	Francesco	Allocataire	Université SPN
Schneider	Alexis	CIFRE	Orange
Schulz	Alexandre	Fond propre	Université SPN
Vallée	Mathieu	Allocataire	Université SPN
Zhang	Yue	Allocataire	Université SPN
Zigrand	Louis	CIFRE	Padam

1.1.4 Doctorants associés au 1er oct. 2022

(*Les doctorants associés sont encadrés par des membres du LIPN et dépendent d'une école doctorale extérieure à l'Université Paris 13.*)

Nom	*Prénom*	*Situation*	*Institution*
Concas	Emanuele	Doctorant	ENPC

1.1.5 Post-Doctorants au 1er oct. 2022

Nom	*Prénom*	*Situation*	*Début*	*Fin*	*Financement*
Bettiol	Enrico	Post-Doc	2022	2023	Huawei

1.1.6 Personnel temporaire du 01/09/2017 au 01/10/2022

Nom	**Prénom**	**Situation**	**Début**	**Fin**	**Financement**
De Melo Da Silva	Marcos	Doctorant	2013	2017	Ministère
Pan	Stefania	Doctorante	2015	2018	Bourse CIFRE
Bettiol	Enrico	Doctorant	2016	2019	Mnistère
Lancini	Emiliano	Doctorant	2016	2019	Ministère

Leclercq	Etienne	Doctorant	2017	2022	CIFRE
Menouer	Tarek	Postdoc	2017	2018	sur projet UMANIS
Chahbar	Mohammed	Doctorant	2018	2021	Ministère
Sliwak	Julie	Doctorante	2018	2021	Bourse CIFRE
Torres Figueroa	Juan José	Doctorant	2018	2022	Ministère
Torres Gonzales	Daniel	Doctorant	2018	2021	Ministère
Pass-Lanneau	Adèle	Doctorante	2018	2021	CIFRE
Delle Donne	Diego	Postdoc	2018	2019	Bourse MathStic
Rios Solis	Yasmine	Prof. étrangère	2019	2020	en congé sabatique

1.1.7 Évolution de l'équipe (permanents)

Informations générales sur l'équipe AOC

L'équipe AOC a été créée en 2010, lors de la restructuration du laboratoire en 5 équipes, par une partie des membres de l'équipe OCAD, regroupant les compétences en « Optimisation Combinatoire » et en « Algorithmes, logiciels et architectures distribués ». Pendant la période précédente (2017–2022), elle n'a été renforcée que par l'arrivée de P. Fouilhoux et la promotion de L. Létocart. Dans la même période, G. Plateau est devenu membre honoraire après la fin de son émérita, F. Pekergin est parti en retraite, C. Coti en disponibilité et C. Cérin en délégation.

Les axes de recherche de l'équipe (Graphes et polyèdres, Programmation mathématique et Algorithmes, logiciels et architectures distribués) sont maintenus et ont progressé dans leurs développements.

On peut aussi noter les soutenances d'Habilitation à Diriger des Recherches de E. Traversi (01/2023) et R. Grappe (11/2021).

Promotions

2022 L. Létocart, Professeur

Départs

2022 F. Pekergin, MCF IUTV (retraite)

Arrivées

2020 P. Fouilhoux, PU IG (Paris 6)

1.2 Les thématiques scientifiques et leurs enjeux

L'équipe a une taille moyenne dans le panorama national et international, et il est intéressant de remarquer le large spectre de ses thèmes de recherche. Au niveau national, on retrouve généralement soit des équipes de recherche fortement orientées vers le transfert de technologie et le rapport avec le monde industriel, soit des équipes très théoriques, plus orientées graphes et polyèdres, ou encore des équipes se focalisant principalement sur la programmation mathématique. Nous avons l'ensemble des trois aspects de manière bien équilibrée et bien représentée. De plus l'équipe a une forte composante dans l'axe calcul parallèle et distribué très reconnue localement, nationalement et internationalement.

Les titres des axes définissent bien les communautés auxquelles appartient l'équipe. De plus, pour chaque axe, nous avons au moins un thème de recherche dans lequel nous sommes très reconnus et très bien integrés dans ces communautés nationales et internationales (aussi bien en tant qu'organisateurs d'évènements, qu'invités dans des conférences renommées car nous sommes des animateurs de ces communautés scientifiques). Citons, par exemple, de façon non exhaustive, polyédre Box-TDI, méthodes de décomposition, optimisation non-lineaire, *c*loud et environnement d'exécution distribuée. Au niveau de l'équipe, nous pouvons remarquer une plus grande quantité de projets et travaux co-signés par rapport à la période précédente. Nous avons aussi amélioré notre rayonnement et visibilité comme équipe. La remarquable diversité des thèmes abordés et des compétences mises en œuvre est accompagnée de deux autres facteurs : la croissance en nombre et en qualité des publications ainsi que la quantité de thèses CIFRE et de contrats développés dans la période concernée.

La vie scientifique de l'équipe est organisée à travers des rencontres hebdomadaires, à savoir : un séminaire fait par un invité international (parfois par visio depuis le covid), un séminaire fait par les membres juniors de l'équipe (doctorants ou post-doctorants), une réunion d'équipe et une matinée scientifique. Cette dernière est l'occasion de partager des réflexions sur des sujets brûlants et de stimuler les collaborations entre membres. Plus en details on a donné le nom de matinées scientifiques à des rencontres informelles de type *brainstorming* sur un sujet de recherche qu'un membre de l'équipe souhaite aborder. Ces matinées scientifiques se transforment parfois en de vrais groupes de travail hebdomadaire quand le sujet est vaste et intéresse plusieurs membres de l'équipe. On peut remarquer que, même pendant la période du COVID, nous avons organisé et mis en place des rencontres hebdomadaires. Ces rencontres ont aussi été l'occasion d'inviter des conférenciers étrangers.

D'une manière générale, l'équipe AOC a bien avancé pendant ces six ans sur plusieurs thèmes de recherche qui font l'objet des prochains paragraphes.

Le lecteur intéressé par les détails des différents axes pourra remarquer des résultats innovants et excellents qui ont relevé des défis importants.

Informations générales sur l'équipe AOC

Enfin, il faut signaler que les membres de l'équipe ont développé d'autres méthodes que celles mise en évidence ici et qui ont donné lieu à d'autres publications (comme par exemple, mais la liste et loin d'être exhaustive : [**RI-48**, RI-50, RI-49, RI-45, RI-39]).

1.2.1 Axe 1 : Graphes et polyèdres

Les principaux résultats obtenus dans cet axe portent sur la caractérisation des systèmes et polyèdres box-TDI et l'étude du polyèdre induit par l'ensemble des solutions réalisables d'un problème. Les paragraphes suivants détaillent le contexte et certains de ces résultats.

- **Polyèdres box-TDI**

En optimisation combinatoire, de nombreux résultats importants découlent de relations min-max, tel que le théorème FlotMax-CoupeMin de Ford et Fulkerson [4] qui stipule que, dans un graphe orienté ayant une source s, un puits t et des capacités sur ses arcs, la quantité maximale de flot de la source au puits respectant les capacités est égale à la capacité minimale d'un ensemble d'arcs à supprimer pour déconnecter s de t — une *st-coupe*. Ce type de relation est essentielle d'un point de vue théorique, caractérisant les optimalités de diverses familles d'objets combinatoires, comme les flots et les st-coupes ci-dessus. Ceci a aussi de fortes conséquences algorithmiques : une relation min-max fournit un critère d'arrêt particulièrement utile lors de la recherche de solutions optimales. Bien souvent, de telles relations découlent de propriétés d'intégralité de certains systèmes linéaires. Les polyèdres box-TDI jouent un rôle important dans ce contexte car ils sont associés à des systèmes linéaires qui ont de fortes propriétés d'intégralité. Les systèmes linéaires ayant la propriété qu'il existe toujours une solution optimale entière dans leur dual sont appelés des systèmes *TDI* (totally dual integral). Parmi les systèmes TDI, il en existe qui produisent des relations min-max plus fortes : les systèmes *box-TDI*. Ces systèmes sont ceux qui restent TDI lorsque des bornes inférieures et supérieures sont imposées sur les variables. Dans le dual, l'effet d'être box-TDI est que l'on peut modifier la fonction objective du problème primal par une quantité entière et à un certain coût avant de résoudre le problème dual résultant. Il se trouve que le système linéaire derrière le théorème de MaxFlow-MinCut est box-TDI. Les polyèdres décrits par de tels systèmes sont appelés des polyèdres *box-TDI*. Ces systèmes et polyèdres ont été introduits autour des années 80, mais on en savait peu à leur sujet avant les vingt dernières années. Les polyèdres box-TDI sont fondamentaux en optimisation combinatoire car ils sous-tendent de forts théorèmes min-max, et ils sont l'un des principaux sujets des recherches de l'axe 1 depuis 2016. Dans [**RI-46**], avec Chervet et Robert, nous

fournissons plusieurs caractérisations des polyèdres box-TDI, essentiellement géométriques et matricielles. Ces caractérisations ouvrent une variété de nouvelles questions, et comblent “le manque d’un outil approprié pour établir le caractère box-TDI” — pour citer Ding, Tan, and Zang [2]. Entre autres, elles nous ont permis de réfuter la conjecture de Ding, Zang, and Zhao [3] sur les graphes box-parfaits. Ces caractérisations nous ont permis ultérieurement d’étudier le caractère box-TDI d’un certain nombre de polyèdres classiques en optimisation combinatoire. Dans [RI-62], nous fournissons le système de Schrijver du cône des flots d’un graphe série-parallèle. Dans [CI-44], nous étudions le polyèdre lié aux sous-graphes couvrants k-arête-connexes d’un graphe donné : ce sont les sous-graphes qui restent connexes après la suppression de n’importe quel ensemble de $k-1$ arêtes. Notre caractérisation de son aspect box-TDI renforce un certain nombre de résultats bien connus sur ce polyèdre. Avec Cornaz et Lacroix dans [RI-24], nous fournissons un système box-TDI décrivant le cône des coupes d’un graphe série-parallèle, ce qui renforce le résultat TDI de Schrijver [6, Corollary 29.9c]. Ces articles répondent à deux défis conjoints dans ce thème : le défi théorique de caractériser les propriétés géométriques et matricielles de ces familles de polyèdres d’abord, et dans un second temps celui de traduire ces propriétés en un théorème min-max, utile par exemple au développement d’algorithmes combinatoires résolvant les problèmes sous-jacents.

Informations générales sur l’équipe AOC

- **Polyèdres et inégalités**

On appelle souvent *polyèdres combinatoires* les polyèdres définis comme l’enveloppe convexe des vecteurs solutions d’un problème d’optimisation combinatoire. Étudier la structure de tels polyèdres permet à la fois de déduire des propriétés sur les valeurs et la nature des solutions, mais aussi de concevoir des algorithmes de résolutions puissants tel que l’algorithme de coupes et branchement.

L’équipe AOC s’intéresse énormément à ces approches polyédrales dans le but de concevoir des outils de résolutions dédiés à des problèmes donnés. Dans [↑RI-3], nous proposons une étude du polyèdre du problème de décision d’allumage ou d’extinction des centrales de production électrique, appelé Unit Commitment Problem (UCP). Cette étude a permis de mieux cerner la difficulté de ce problème central pour la production énergétique et dont nous avions étudié la complexité dans [↑RI-3]. Elle a également permis de proposer un algorithme de résolution efficace pour l’UPC basé sur des inégalités valides généralisant les inégalités du polyèdre du sac-à-dos en tenant compte de la dynamique temporelle de la prévision de consommation électrique. Une approche similaire a permis dans [RI-33] d’étudier et résoudre une relaxation du célèbre problème de coloration de graphe. Dans [↑RI-7], nous proposons une étude d’un polyèdre moins classique basée sur des variables continues indiquant les dates de début de tâches d’un problème d’ordonnancement juste à temps.

Un autre aspect plus fondamental des études polyèdrales est de donner la description de toutes les inégalités formant un polyèdre combinatoire, on appelle ceci donner la caractérisation d’un polyèdre. Si cette caractérisation est de taille polynomiale, ou si elle peut être associée à un algorithme de séparation qui soit polynomial, une telle approche permet de montrer la polynomialité du problème. Ce type de résultats permet d’établir la polynomialité de cas particuliers de problèmes sur une classe d’instances donnée, mais aussi de mettre en lumière des inégalités centrales dans la structure du problème. Dans [RI-19], nous donnons ainsi la caractérisation du polytope lexicographique, défini comme l’ensemble des points entiers qui sont lexicographiment entre deux points entiers donnés. Dans [RI-11], une caractérisation est donné pour le problème des st-bond (une st-coupe minimal au sens de l’inclusion) dans le cas des graphes série-parallèles.

Une approche originale a été débutée dans l’équipe afin d’utiliser les approches polyédrales pour réduire l’espace des solutions à explorer lors de la résolution des problèmes. Une première idée est de limiter l’exploration à un unique représentant par classe de symétrie dans l’espace des solutions. Cette technique consiste à intersecter le polyèdre combinatoire avec un polyèdre lexicographique sur les permutations d’indice

dans les vecteurs solutions : cette idée a été mise en œuvre algorithmiquement dans [↑RI-6] et dans une méthode de branch-and-cut dans [↑RI-8] pour le problème de l'UCP. Une deuxième idée est de couper une solution dominée au sens de l'existence d'une opération de voisinage construisant une meilleure solution à partir d'elle. Des inégalités de dominance ont ainsi été étudiées et utilisées dans [↑RI-10] pour un problème d'ordonnancement juste à temps.

En lien avec la thématique de décomposition, les approches polyèdrales sont également possibles pour les problèmes maîtres de la décomposition, en permettant de concevoir des algorithmes de branch-and-price comme dans [↑RI-2] pour un problème de planification.

1.2.2 Axe 2 : programmation mathématique

Les principaux résultats obtenus portent sur : l'application des approches par reformulation et décomposition, l'étude et la résolution de modèles non linéaires, les méthodes de *Machine learning* appliquées à *l'optimisation*.

- **Reformulation et decomposition**

La grande majorité des méthodes de résolution d'un modèle de programmation mathématique reposent sur l'idée de résoudre de manière répétée un problème "plus facile" à résoudre. Ce principe se retrouve, par exemple, dans l'algorithme "branch-and-bound", où un problème à variables discrètes est résolu en résolvant plusieurs fois un problème plus simple (souvent la relaxation continue du problème d'origine). Le but des techniques de reformulation et de décomposition est précisément de réécrire un modèle mathématique sous une forme plus facilement résoluble au moyen d'un algorithme approprié. Malheureusement leur résolution directe n'est pas envisageable à cause du nombre considérable de variables engendrées par la reformulation où la décomposition. Une approche classique dans le cadre linéaire (et plus récement en non-linéaire [**RI-58**]) consiste à résoudre une formulation restreinte à un sous-ensemble de variables, puis d'ajouter à cette formulation restreinte des variables permettant d'améliorer la solution courante, tant que cela est possible : la formulation finalement obtenue donne la solution optimale du problème. Cette approche peut ensuite être insérée dans des schémas de séparation et d'évaluation progressive donnant lieu à des méthodes exactes de type *b*ranch-and-price et *b*ranch-and-cut-and-price. Nous avons proposé différentes reformulations linéaires et non-linéaires de problèmes combinatoires afin de construire des schémas de résolution complémentaires. Ceux-ci s'appuient notamment sur des relaxations et décompositions des modélisations, qui peuvent ensuite être insérées dans des schémas de séparation et d'évaluation progressive. Voir notamment [RI-25, **RI-41**, RI-43, RI-4, **RI-48**, RI-16, CO-28, ↑RI-2].

Le grand nombre et la variété des articles témoignent de l'expertise de l'équipe dans les méthodes de décomposition. Au vu de la spécificité de chaque document, nous n'entrerons pas dans les détails de chacun. Nous nous limiterons à mettre en évidence les trois qui, selon nous, ont des spécificités plus complexes et intéressantes que les autres. En [RI-20], nous étendons et généralisons une méthode de résolution du maître restreint sous-jacent à une approche "branch-and-price". Nous démontrons mathématiquement et expérimentalement la stabilité et la rapidité de l'approche par rapport à un solveur linéaire. In [**RI-37**] nous proposons un schéma de décomposition emboîtée pour les programmes linéaires stochastiques à **horizon infini**. Notre approche peut être considérée comme une extension convergente de la programmation dynamique duale stochastique au cadre de l'horizon infini : nous explorons une séquence de problèmes d'horizon fini de longueur croissante jusqu'à ce que nous puissions garantir la convergence avec un niveau de confiance donné. La méthodologie alterne entre un passage vers l'avant pour explorer des exemples de chemins et déterminer des solutions d'essai, et un passage vers l'arrière pour générer une approximation polyédrique de la

fonction de valeur optimale en calculant des sous-gradients à partir des sous-problèmes de scénario. Enfin dans [**RI-27**] (cf. Sec10) nous proposons un modèle original pour le problème de diffusion dans un réseau : (virus, information, etc.). Nous proposons également un algorithme de résolution basé sur une succession de transformations et de décompositions qui passent par une formulation étendue et une formulation non-linéare, pour citer que les principales.

- **Optimisation non linéaire et quadratique**

L'optimisation non linéaire concerne l'optimisation où la fonction objectif et/ou les contraintes sont exprimées sous la forme de fonctions non linéaires. L'optimisation quadratique concerne quant à elle l'optimisation de problèmes avec des fonctions polynomiales de degré deux. L'optimisation non linéaire, et notamment l'optimisation quadratique, fait le lien entre l'optimisation combinatoire et l'optimisation continue. La possibilité d'exprimer l'objectif et/ou les contraintes sous des formes non linéaires permet d'augmenter l'expressivité des modèles mathématiques utilisés pour formuler les problèmes d'optimisation et ainsi de coller davantage à la réalité. Cette flexibilité permet d'augmenter le spectre des problèmes qui peuvent être modélisés.

Informations générales sur l'équipe AOC

Les liens entre optimisation non-linéaire et combinatoire sont de plus en plus explorés et exploités avec succès avec l'avènement et le développement d'outils numériques performants. Le cheminement est comparable à celui qui a été emprunté (et qui est toujours vivace) par l'optimisation linéaire pour l'étude et la résolution efficace de problèmes difficiles de l'optimisation combinatoire [RI-23]. Cette extension naturelle au contexte non-linéaire, presque mécanique, des approches connues dans le cadre linéaire (modélisations, relaxations, dualité, décomposition et génération de colonnes, hybridation d'approches numériques et combinatoires), est enrichi par le contexte non-linéaire et permet également l'élaboration de méthodes spécifiques. Outre la plus grande diversité qu'offrent les approches non-linéaires pour la modélisation, certains problèmes sont réputés pour être résistants aux techniques fondées sur la linéarisation (typiquement un saut d'intégralité important entre les modèles et leurs relaxations continues) et sont des candidats de choix pour ces nouvelles approches. Une contribution essentielle de l'équipe AOC est d'ordre méthodologique : elle concerne la proposition d'approches génériques pour la résolution de classes entières de problèmes combinatoires [RI-61] [**RI-17**] [RI-42] ou de problèmes fondamentaux de l'optimisation combinatoire liés à de nombreuses applications [RI-15]. Ce volet est prolongé par la constitution de bases de tests d'instances [**RI-22**] et la production et la diffusion de logiciels [RI-1] dans la communauté académique et auprès de partenaires industriels. Un second axe de recherche est l'exploitation de propriétés (implicites ou explicites) pour des classes d'instances de problèmes combinatoires. Ainsi, on peut choisir de réduire volontairement l'ensemble des instances considérées en effectuant des hypothèses structurelles supplémentaires. L'objectif est alors de rendre plus efficaces des méthodes dont la complexité dépend fortement de quelques paramètres ou d'une étape particulière [RI-34]. Une autre option est de considérer la spécificité (souvent implicite) d'instances de problèmes industriels qui présentent rarement un caractère uniforme [**RI-35**] [**RI-17**] [RI-38]. Formaliser ou rendre explicite cette hétérogénéité n'est pas nécessairement aisé et ses effets peuvent se révéler dans un premier temps via une étude expérimentale. Idéalement, dans un second temps, son analyse et sa compréhension peuvent permettre l'amélioration des méthodes développées. Parmi les problèmes ciblés par l'équipe AOC on peut citer ceux qui appartiennent à la programmation quadratique (voire polynomiale) en variables booléennes, entières ou mixtes, la programmation fractionnaire, et des problèmes appliqués fortement non-linéaires issus des domaines de l'énergie et du transport. Les contributions de l'équipe sont à la fois théoriques et expérimentales et presque toujours validées par une mise en œuvre informatique. Plusieurs membres de l'équipe ont également proposé des méthodes de décomposition pour l'optimisation non linéaire, notamment [RI-34], [CO-27], [**RI-58**] et [RI-61]. En ce qui concerne les résultats ob-

tenus par l'équipe [RI-61] pour la génération de matrices pour résoudre des problèmes binaires avec objectif et contraintes quadratiques via une nouvelle méthode de décomposition, nous souhaitons mettre en evidence la décomposition BQP (Boolean Quadratic Polytope). Ces résultats prometteurs permettent d'envisager la résolution exacte de ces problèmes via des approches de type branch-and-price où cette décomposition serait résolue en chaque nœud de l'arbre ; de résoudre les sous-problèmes issus de cette décomposition par un algorithme quantique, permettant ainsi d'accélérer par un facteur 1000 la résolution de ces sous problèmes ; et enfin de généraliser cette approche pour d'autres problèmes non linéaires.

- **Machine learning et optimisation**

L'apprentissage d'un réseau de neurones peut être considéré comme une simple interpolation d'une fonction très complexe. Dans de nombreuses applications, il est impossible de décrire la fonction objectif ou les contraintes par une formulation mathématique. Dans ces situations, il est nécessaire de recourir à des outils de boîte noire ou de méta-modélisation. La méta-modélisation est un sous-domaine de l'apprentissage automatique qui consiste à représenter un modèle complexe f par un modèle de substitution plus simple $\hat{f}$ au prix d'une certaine approximation. L'un des avantages de l'utilisation du modèle de substitution est de permettre un calcul plus rapide des nouvelles valeurs de $\hat{f}$, car les simulations sont généralement coûteuses à calculer dans le cadre initial.

Dans [**CI-47**], nous proposons d'utiliser une nouvelle fonction objectif pour un problème de transport à la demande qui tient compte de l'historique des demandes et de la politique utilisée par l'algorithme de résolution en ligne. L'apprentissage de la fonction objectif n'est qu'une partie du travail. L'étape suivante consiste à essayer de l'optimiser. Dans [**CI-47**], nous utilisons un algorithme heuristique de recherche locale développé par l'entreprise pour optimiser le méta-modèle appris. Dans le prolongement du travail effectué dans [**CI-47**], nous avons récemment proposé à la conférence TRISTAN 2022 [**CI-51**] une généralisation du Dial-a-Ride Problem (DaRP) [1] pour résoudre exactement la planification hors ligne. Le modèle proposé peut être amélioré en prenant mieux en compte l'incertitude des données. Dans [7] nous proposons une méthodologie ad hoc pour mieux estimer la demande, à utiliser dans le modèle proposé dans [**CI-51**].

Dans [CI-56], nous proposons d'apprendre la fonction qui renvoie son temps de résolution attendu pour une décomposition en cliques donnée. La décomposition en cliques est une technique bien établie pour résoudre les problèmes semi-définis peu denses [5]. Les résultats obtenus sont prometteurs. Ils montrent qu'il est possible d'apprendre une telle fonction, et qu'elle peut être utilisée pour choisir parmi différentes décompositions celle qui est la plus prometteuse.

1.2.3 Axe 3 : algorithmes, logiciel et architectures distribués

L'axe 3 étudie les domaines du calcul à hautes performances (HPC), l'environnement d'exécution distribué, l'informatique nuagique (cloud computing) et ses variantes (Edge et Fog Computing), avec des visions de convergence, sur certains aspects, des systèmes. De plus, cet axe étudie la tolérance aux pannées des machines à grand échelle.

- **Cloud**

Un aspect du travail traite des stratégies et des techniques permettant d'ouvrir la voie à l'exécution d'applications HPC traditionnelles dans le cloud[CI-45]. Nous suivons une direction de recherche qui peut sembler inhabituelle à première vue. En effet, nous nous intéressons à la problématique ou l'effort de cloudification est porté sur l'ordonnanceur HPC. En clair, les applications HPC ne sont pas réécrites, mais l'ordonnanceur HPC a été cloudifié. Avec nos travaux, l'ordonnanceur HPC est désormais disponible comme n'importe quel autre service Cloud à la demande. Ensuite, pour compléter ce

saut conceptuel, nous nous sommes intéressés à la problématique de la définition d'un contrôleur pour l'orchestrateur de cloud qui permette l'autoscaling des clusters HPC conteneurisés dans le Cloud[RI-36, **CI-54**]. La solution proposée vise à déclencher la création ou la suppression de nœuds de calcul HPC conteneurisés en fonction des métriques collectées au niveau de la file d'attente du planificateur HPC conteneurisé. En résumé, nous travaillons à l'unification du monde du HPC et du monde du Cloud, ce qui constitue le cœur du travail[CI-12].

Un deuxième aspect porte sur la définition d'algorithmes d'ordonnancement multi-critères pour placer des conteneurs dans les plateformes de Cloud [CI-29, RI-28]. Nous travaillons ici au niveau de la file d'attente des travaux soumis à l'orchestrateur. Une originalité du travail est de ne pas toucher aux composants de l'orchestrateur, mais de compter sur eux, pour définir de nouvelles politiques. Il ne s'agit pas de vouloir faire aussi bien que Kubernetes de Google, car Kubernetes est un projet très vaste. Par contre nous pouvons viser l'ajout de nouvelles fonctionalités d'orchestration, facilement intégrables dans les variantes de Kubernetes (K8s, K3s, micro-K8s).

Informations générales sur l'équipe AOC

- **Edge et Fog Computing**

Pour l'Edge Computing, l'idée ici est d'effectuer une partie des calculs en bordure de réseau (Edge)[CI-4] et d'envoyer un résumé au système Cloud, plutot que de systématiquement envoyer toutes les données au Cloud, afin d'apprendre sur les données. Le contexte applicatif est le bâtiment intelligent (Smart Building)[CI-21] pour lequel nous avons mis en place des collaborations avec le NII à Tokyo et le LIG (Grenoble)[CI-19]. Un des objectifs a été de mettre en place une méthodologie pour la phase d'exploration de jeux de données issus de batiments ainsi que de minimiser le nombre de capteurs à déployer au sein d'un bâtiment afin de garantir un niveau acceptable d'apprentissage.

L'activité en Fog Computing, se déroule depuis un an dans le cadre du projet FogSLA (AMI cloud) avec SMILE, CGI (ex Umanis), Gandi et Nokia pour les industriels. Pour relever le défi de l'orchestration des services de bout en bout sur la base de la qualité de service demandée et de l'accord de niveau de service (SLA), la plate-forme FogSLA s'appuiera sur un "accord de niveau de service dynamique" pour introduire et orchestrer les services de bout en bout.

L'idée est d'introduire et d'orchestrer des configurations flexibles basées sur le contrôle et l'orchestration du niveau fonctionnel. Ainsi, selon ces cas d'utilisation, la plateforme garantira les niveaux d'exécution dans des environnements hétérogènes et renforcera dynamiquement les liens réseau entre les composants en fonction de leurs besoins. Notre partie de travail se concentre sur deux aspects importants : l'optimisation de l'ordonnancement des conteneurs et les aspects liés au routage et au réseau, en collaboration avec le L2TI de l'Université Sorbonne Paris Nord.

- **Fault tolerance**

Les machines à très grande échelle sont sujettes aux défaillances, car elles sont composées d'un très grand nombre de composants. Ces pannes peuvent être matérielles (par exemple, un nœud de calcul qui surchauffe et s'éteint) ou logicielles (par exemple, un processus accidentellement déordonnancé suite à un bug de l'ordonnanceur de travaux). Il existe deux approches pour tolérer les défaillances et ainsi permettre aux applications parallèles distribuées de continuer leur exécution sans avoir à repartir du début. L'approche au niveau système est transparente pour l'application : il n'est pas nécessaire de la modifier, c'est l'intergiciel qui s'occupe de sauvegarder périodiquement l'état de l'application en enregistrant des points de reprise et en maintenant la cohérence de l'application distribuée et, en cas de défaillance, de repartir sur un état précédent. L'approche au niveau application est plus spécifique : l'application elle-même prend en compte la tolérance aux défaillances en définissant le comportement à suivre en cas de défaillance. Cette approche est souvent plus efficace, mais elle nécessite de prendre en compte la tolérance aux défaillances dans la conception de l'application.

Nous avons travaillé sur ces deux approches. L'approche au niveau système repose sur deux familles principales d'algorithmes : coordonnés (tous les processus se synchronisent et prennent un point de reprise au même moment) et non-coordonné (les processus prennent leurs points de reprise sans se coordonner, mais on enregistre les messages échangés pour pouvoir les rejouer en cas de retour en arrière). Les algorithmes impliqués reposent généralement sur l'hypothèse de canaux de communications ayant la propriété FIFO : par exemple, quand on fait circuler un marqueur, il ne double pas de message envoyé précédemment, et inversement. Cette hypothèse ne fonctionne pas dans un modèle de communications unilatérales comme celui suivi par OpenSHMEM. Nous avons montré cela avec un contre-exemple, et proposé des algorithmes de prise de point de reprise coordonné fonctionnant dans ce modèle, et implémentables à divers niveau de la pile de communications [CI-111, **CI-97**].

- **Environnement d'exécution distribué**

L'exécution d'une application parallèle sur des ressources distribuées est supportée par un intergiciel distribué qui orchestre l'utilisation de ces ressources : il lance la commande à exécuter sur les ressources, transfère les signaux et les entrée-sorties, monitore l'état de l'application (pour la détection de défaillances et l'analyse de performances), et permet les communications entre les processus. Il forme, avec une bibliothèque de communications à hautes performances, le support d'exécution. Ces deux composants du support d'exécution sont critiques pour les performances des applications parallèles distribuées.

Nous avons travaillé sur l'interface de programmation parallèle OpenSHMEM, [LO-4] qui suit un modèle de communications unilatérales. Le but d'OpenSHMEM est de fournir des communications qui s'implémentent en peu d'instructions machines, pour réduire le plus possible le chemin entre l'API et le matériel réseau. Nous avons implémenté la norme OpenSHMEM[CO-19, CO-12] dans un but de minimisation de ce chemin, en nombre d'instructions et d'indirections. Cette implémentation donne des communications avec une latence très basse. À cette époque, la communauté essayait diverses options d'implémentation parmi lesquelles figuraient également l'utilisation des communications unilatérales de MPI. Nous avons comparé ces implémentations, et notre approche s'est avérée la plus efficace. Aujourd'hui, l'approche que nous avons fournit a été adoptée par la communauté, avec l'initiative OpenUCX pour fournir une pile logicielle très fine tout en étant portable.

Nous avons également travaillé sur les outils d'analyse de performances, comme TAU, en implémentant des interfaces de programmation et un plugin pour les compilateurs de LLVM pour l'implémentation sélective par TAU. Ces travaux font l'objet d'une collaboration à long terme avec l'équipe de développement de TAU, qui se poursuit à travers des contributions à d'autres logiciels de l'écosystème de mesure TAU, ainsi que des concepteurs d'intergiciels.

Références complémentaires

[1] J.-F. Cordeau. A branch-and-cut algorithm for the dial-a-ride problem. *Operations Research*, 54(3) :573–586, 2006.

[2] G. Ding, L. Tan, and W. Zang. When is the matching polytope box-totally dual integral ? *Mathematics of Operations Research*, 43(1) :64–99, 2018.

[3] G. Ding, W. Zang, and Q. Zhao. On box-perfect graphs. *Journal of Combinatorial Theory, Series B*, 128 :17–46, 2018.

[4] L. R. Ford and D. R. Fulkerson. Maximal flow through a network. *Canadian J. Math.*, 8 :399–404, 1956.

[5] M. Fukuda, M. Kojima, K. Murota, and K. Nakata. Exploiting sparsity in semidefinite programming via matrix completion i : General framework. *SIAM Journal on optimization*, 11(3) :647–674, 2001.

[6] A. Schrijver. *Combinatorial optimization : polyhedra and efficiency*. Algorithms and combinatorics. Springer-Verlag, Berlin, Heidelberg, New York, N.Y., et al., 2003.

[7] L. Zigrand, P. Alizadeh, E. Traversi, and R. Wolfler Calvo. Optimization-driven Demand Prediction Framework for Suburban Dynamic Demand-Responsive Transport Systems. In *IJCAI 2023, the 32nd International Joint Conference on Artificial Intelligence*, 2023.

1.3 Profil d'activités liées à la recherche

Activités (répartir 100 points sur ces 7 items)	
Administration et animation de la recherche : pilotage de la recherche (VP, direction d'institut, DAS, par exemple), participation à des instances d'évaluation (CNU, CoNRS, CSS, Hcéres, par exemple), responsabilité de dispositifs Idex ou Isite, direction de projets (ANR, Horizon Europe, ERC, CPER, PIA, France 2030, par exemple), responsabilités éditoriales dans des revues ou collections nationales et internationales.	33
Aide aux politiques publiques et expertise technique : pouvoirs publics aux niveaux européen, national et régional, entreprises, instances internationales comme FAO, OMS, etc.	0
Contribution à l'adossement d'enseignements innovants à la recherche : EUR, SFRI, etc.	0
Dissémination de la recherche : partage de connaissances avec le grand public, médiation scientifique, interface sciences et société	1
Recherche et encadrement de la recherche.	33
Valorisation, transfert, innovation.	33
Autres activités. (à préciser en une ligne maximum).	

Informations générales sur l'équipe AOC

1.4 Prise en compte des recommandations du précédent rapport

Recommandation La répartition des productions scientifiques est très déséquilibrée entre les membres de l'équipe.

Réponse La production scientifique de l'équipe est un peu déséquilibrée entre les membres. Nous avons mis en place plusieurs mécanismes pour favoriser le travail conjoint et mettre les membres plus fragiles dans les meilleures conditions pour augmenter leur productivité. Parmi les mécanismes, nous pouvons lister : le co-encadrement de doctorant (tous les membres de l'équipe ont co-encadré au moins une thèse), les matinées scientifiques : partage de sujets sur lesquels démarrer éventuellement des collaborations et répondre aux appels externes (ex. le challenge ROADEF). En plus de ces mécanismes, il faut reconnaître que les personnes moins actives en recherche sont des personnes qui sont très investies dans l'enseignement et les responsabilités administratives.

Nous avons plusieurs membres très impliqués dans les responsabilités administratives de l'université, notamment :

- Frédéric Roupin a été directeur de l'Institut Galilée (plus de 3000 étudiants en licence, master et ingénieurs, 7 laboratoires de recherche) de 2017 à juillet 2022.
- Roberto Wolfler Calvo est responsable du pôle Math-STIC (qui regroupe les laboratoires de recherche du LIPN, LAGA et L2TI) depuis 2020.
- Sophie Toulouse a été directrice des études de la spécialité "Informatique" de l'école d'Ingénieurs Sup Galilée de 2009 à 2012 et membre du Conseil d'Administration

— Lucas Létocart a été directeur adjoint de l'Institut Galilée en charge des Licences de 2015 à janvier 2020. Lucas Létocart est actuellement directeur adjoint de l'Institut Galilée en charge de l'école d'ingénieur SupGalilée depuis juillet 2022.
— Sylvie Borne est directrice des études de la spécialité informatique à l'école d'ingénieurs Sup Galilée depuis septembre 2009.
— Jean-Christophe Dubacq a été chef du département Informatique de l'IUT de Villetaneuse de 2014 à août 2020 et est actuellement co-directeur des études des deux premières années du BUT informatique en apprentissage à l'IUT de Villetaneuse. Il est aussi président de la commission des locaux de l'IUT depuis 2020.
— Trois membres de l'équipe AOC siègent dans différents conseils centraux de l'Université Sorbonne Paris Nord (un au Conseil d'Administration, un à la Commission Recherche et un à la Commission de la Formation et de la Vie Universitaire).

Informations générales sur l'équipe AOC

Recommandation Si l'équipe compte quelques membres très reconnus à l'international, sa visibilité d'ensemble gagnerait à être améliorée

Réponse Nous avons augmenté la visibilité au niveau de l'équipe. Plusieurs indicateurs peuvent en témoigner comme, par exemple, le très bon niveau des candidats aux concours de professeur, de Maître de conférence et de CR CNRS de cette période. Le deuxième indicateur est la capacité à engager des contrats au niveau de l'équipe avec des entreprises et pas seulement au niveau individuel, l'exemple le plus flagrant étant le contrat avec Huawei.

Recommandation La création d'une identité d'équipe ne semble pas totalement aboutie à ce stade ; l'échange, puis la collaboration entre les trois axes est à mettre en place et à renforcer.

Réponse Les interactions entre les axes 1 et 2 se sont beaucoup renforcées grâce à l'arrivée de Pierre Fouilhoux car ses thématiques de recherche sont à cheval sur les deux. L'axe 3 interagit avec les autres par l'intermédiaire de projets et de contrats (un BQR, une PEPS et le FUI Wolphin).

Recommandation Les collaborations industrielles existantes devraient aller au-delà de thèses CIFRE, post-docs ou stages.

Réponse Pendant la période 2012-2017 qui vient de s'écouler, nous avons progressé dans le transfert technologique. Nous souhaitons faire remarquer que, pendant la période précédente, les thèses CIFRE avaient été indépendantes et ne concernaient que de contacts épisodiques avec des PME. Pendant la période qui vient de s'écouler (2017-2022), grâce aussi à l'arrivée du nouveau professeur, nous avons pu créer des liens forts et pérennes avec de grandes entreprises françaises et étrangères : EDF, Orange et Huawei.

Recommandation D'autre part, afin de gagner en visibilité à l'échelle mondiale le développement des partenariats internationaux devrait être renforcé

Réponse Nous avons des interactions suivies avec plusieurs pays dont le Mexique (avec en particulier le Tecnológico de Monterrey), l'Italie (relations avec plusieurs universités italiennes), les Emirats Arabes Units (avec en particulier l'institut de recherche TII) et la Chine. Nos relations avec ces quatre pays/continents ont vocation à perdurer.

Recommandation L'absence d'adossement à un master constitue une difficulté pour le recrutement de doctorants.

Réponse Une des faiblesses de l'équipe est l'absence de Master adossé à ses thématiques de recherche. Cela peut pénaliser le recrutement de bons doctorants, car il faut souvent les chercher à l'étranger, mais, jusqu'à présent, nous avons toujours eu un bon flux d'étudiants de très bon niveau. Nous devrions tirer avantage de la récente mise en place par USPN d'un Master EUR *Mathematic and Computer Science* (M&CS).

Recommandation L'encadrement d'une bonne partie des thèses repose sur un seul permanent et il serait opportun de rééquilibrer cette situation

Réponse Au niveau encadrement, nous pouvons nous enorgueillir de 2 HDR soutenues depuis 2017 et de l'arrivée d'un professeur. L'un des HDR a déjà un doctorant a son actif (Mathieu Vallée), de même que le nouveau professeur. Nous avons donc augmenté la capacité d'encadrement car nous sommes passés de trois professeurs et deux HDR à cinq professeurs et trois HDR.

Recommandation Activités de recherche à renforcer publis/permanent

Réponse Comme nous l'avons fait remarquer auparavant, le nombre des publications a augmenté, car nous sommes passé de 46 à 64 articles de revue et les articles à conférence sont passé de 53 à 55. Il est plus important de noter que les indicateurs sont maintenant à 1,68 et 1,44, articles par personne par an, sans compter les nombre de cosignataires par articles.

Chapitre 2
Introduction du portfolio

Nous présentons dans le portfolio de l'équipe "Algorithmes et Optimisation Combinatoire" (AOC) les 5 articles suivants.

- **Box-total dual integrality, box-integrality, and equimodular matrices**. Article publié dans le journal Mathematical Programming.
- **An exact algorithm for robust influence maximization**. Article publié dans le journal Mathematical Programming.
- **Méthodes de Décomposition Appliquées à des problèmes Non Linéaire**. Article publié dans le journal Mathematical Programming.
- **The Anchor-Robust Project Scheduling Problem**. Article publié dans le journal Operations Research.
- **A Methodology to Scale Containerized HPC Infrastructures in the Cloud**. Article présenté à la conférence Euro-Par.

L'importance de ces documents varie. Le premier est un papier fondamental car il a permis de mettre un point ferme sur certaines questions ouvertes sur les systèmes et les polyèdres Box-TDI et en même temps a ouvert de nouvelles voies de recherche. Le deuxième présente des aspects originaux dans la modélisation des processus de diffusion et présente une approche de résolution articulée et originale car basée sur une chaîne de reformulations. La troisième a permis de mettre un terme à plusieurs questions sur l'interaction et l'implication des approches de décomposition dans le traitement de problèmes quadratiques. Le quatrième est un article fondateur car il propose

l’ancrage comme une approche innovante pour résoudre des problèmes d’optimisation robustes. Le cinquième a reçu le prix Euro-Par car il a montré comment faire du calcul haute performance sur le cloud en utilisant des techniques transparentes pour l’utilisateur et indépendantes du système d’exploitation utilisé.

Chapitre 3
Auto-évaluation de l'équipe AOC

Sommaire

3.1 Domaine 2. Attractivité 140

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche. 140

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels. 142

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs. 144

3.1.4 Référence 4. L'équipe est attractive par la qualité de ses équipements et de ses compétences techniques. 145

3.2 Domaine 3. Production scientifique 146

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité. 146

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels. 146

3.2.3 Publications 147

3.3 Domaine 4. Inscription des activités de recherche dans la société . 160

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique. 160
3.3.2 Référence 3. L'équipe partage ses connaissances avec le grand public et intervient dans des débats de société. 160

3.1 Domaine 2. Attractivité

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche.

Auto-évaluation de l'équipe AOC

Le rayonnement de l'équipe AOC se traduit par plusieurs activités qui montrent l'insertion de l'équipe aussi bien au niveau national qu'international. Les membres de l'équipe ont été invités à parler dans plusieurs contextes différents et à différents niveaux : conférences internationales, écoles doctorales internationales, workshops internationaux et colloques nationaux ou locaux.

Plénière internationale
2019 52^o *Congreso Nacional de la Sociedad Matemática Mexicana* à Monterrey ;
2017 Journée *Franco-Japanese Days on Combinatorics and Optimization*.

Invitations à parler dans une école internationale pour doctorants
2022 orateur invité à *ELAVIO (Escuela Latinoamericana de Verano en Investigación Operativa) School for PhD* à Monterrey ;
2018 2 orateurs invités à *Column Generation School*.

Invitations à parler dans un Workshop international
2023 *Column Generation*, 2 membres ;
2018 *Column Generation*, 2 membres ;
2018 Zoli and his work "Zoli is 50 !", Grenoble.

Invitations à parler dans une conference internationale
2023 Semi-Plénière à la ROADEF 2023 ;
2021 European Conference on Operational Research ;
2021 INFORMS Conference ;
2021 Séminaire Polyèdres et Optimisation Combinatoire ;
2018 Cologne Twente Workshop on Graphs and Combinatorial Optimization ;
2018 International Symposium in Mathematical Programming ;
2018 EURO/ALIO Conference ;
2017 Cologne Twente Workshop on Graphs and Combinatorial Optimization ;

Colloques locaux
2019 Plénière du JFRO, Paris ;
2022 Journée du groupe de travail ATOM du GDR RO du CNRS, Paris ;
2019 Journée du groupe de travail POC du GDR RO du CNRS, Paris.

Invitations internationales
2019 À l'université d'*Edimbourg* ; Ecosse ;
2019 À l'université de *Vienne* ; Autriche ;
2019 À l'université de *Waseda* ; Japon ;
2019 Au *National SuperComputing Center* de *Jinan* ; China ;
2019 À l'université de *Wuhan* ; China ;
2019 À l'université de *Poznan* ; Poland ;
2018 À l'université de *Mascara* ; Algeria ;

Table 3.1 Organisation de manifestations scientifiques

Manifestation	Années	Lieu	Rôle
The 19th Annual IFIP International Conference on Network and Parallel Computing (NPC 2022)	2022	Jinan Chine	General co-chair
The International Conference on Future Computational Paradigms (FCP 2022)	2022	Iles Fidji	General co-chair
The 18th Annual IFIP International Conference on Network and Parallel Computing (NPC 2022)	2021	Paris	General co-chair
Triennial Symposium on Transportation Analysis	2022	Mauritius	Membre du comité
Odysseus	2020	Tangier	Membre du comité
Odysseus	2018	Cagliari	co-Organisateur
Column Generation School	2018	Paris	Organisateur
Journées Polyèdres et Optimisation Combinatoire	2021	*distanciel*	Organisateur
Cologne Twente Workshop on Graphs and Combinatorial Optimization	2018	Paris	Membre du comité
International Symposium on Combinatorial Optimization	2022	*distanciel*	Organisateur
Seminaire Polyèdres et Optimisation Combinatoire	2020	distanciel	co-Organisateur
Seminaire Polyèdres et Optimisation Combinatoire	2021	Paris	co-Organisateur
Seminaire Polyèdres et Optimisation Combinatoire	2022	Paris	co-Organisateur
Seminaire Polyèdres et Optimisation Combinatoire	2022	Paris	co-Organisateur
ROADEF	2017-2022		Membre du comité
Journées Polyèdres et Optimisation Combinatoire	2017	Paris	Membre du comité

2018 À l'université de *Tunis* ; Tunsia ;
2018 À l'université de *Ben Guerir* ; Marocco.
2018 À l'université Ludwig Maximilian de Munich, Allemagne
2018, 2019, 2021 Université de *Cagliari* ; Italie
2018, 2019, 2021 Centre de recherche TII ; Emirates Arabes

Membre d'un jury ou expertise

2023 Expert scientifique pour ANR
2023 Expert pour Hamad Bin Khalifa University
2020 Jury pour les prix des thèses de la chancellerie de Paris en sciences ;
2020 Expert scientifique pour le CNRS
2020 Expert scientifique projets "ESR 2020" région Nouvelle-Aquitaine
2019 Jury pour le prix de thèse national PGMO 2019
2019 Expert scientifique pour le programme STIC-AmSud
2016 – Expert pour le HCERES

Les membres de l'équipe ont également participé à plusieurs comités d'organisation et de programme tant au niveau national (ROADEF, JPOC) qu'international (Odysseus, TRISTAN, ISCO, CTW, Network and Parallel Computing). Toutes les manifestations auxquelles l'équipe a participé sont reportées dans la Table 3.1.

Le tableau 3.2 montre que les membres de l'équipe font aussi partie de comités éditoriaux de manière permanente et/ou ont été invités ponctuellement à gérer des numéros spéciaux pour des revues de premier plan.

Le Tableau 3.3 montre que l'équipe est largement investie dans les instances de pilotage de la recherche. Nous sommes très impliqués au niveau national (par exemple

Table 3.2 Comités éditoriaux

Revue	Années	Rôle
IEEE Transactions on Computers (IEEE TC)		Editorial Board Member
Springer Journal of Cloud Computing		Associate Editor
IEEE Cloud Computing		Member of the advisory committee
Discrete Applied Mathematics	2018	Guest editor
Annals of Operations Research	2018	Guest editor
Networks	2021	Guest editor
Transportations Science	2019	Guest editor
Networks	2018	Guest editor

Table 3.3 Participation aux instances de pilotage de la recherche

Instance de pilotage	Rôle	Periode
Groupe de Recherche Recherche Opérationnel (GDR RO) du CNRS	directeur-adjoint	depuis 2020
Data Intelligence Institute of Paris (DiiP)	executive committee member	depuis 2019
Technical Committee of iPOP-UP (integrative Platform for Omics Projects at Université de Paris)	member	depuis 2019
Executive Committee (Conseil de gestion) for the Sorbonne Paris Nord foundation	member	depuis 2019
Board for HPC computing with MAGI cluster at University Sorbonne Paris Nord	chair	depuis 2011
Programme STIC-AmSud 2019	expert scientifique	une fois
Conseil scientifique du GDR-RO, axe Programmation Mathématique Non-Linéaire	membre	depuis 2022
Conseil scientifique PGMO (Programme Gaspard Monge pour l'Optimisation), fondation mathématique Hadamard	membre	depuis 2018
Fédération de recherche MathStic (FR3734) of USPN	directeur	depuis 2020
Fédération de recherche MathStic (FR3734) of USPN	co-responsable Axe 1	2013-2020
Conseil scientifique de l'Université de Technologie de Troyes	membre	depuis 2019

dans le GDR-RO) et local (par exemple avec la fédération MathStic et le cluster de calcul MAGI).

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels.

L'équipe recrute les doctorants sur deux sources de financement : ressources propres grâce aux contrats CIFRE et bourses ministérielles. Malgré l'absence d'un Master, l'équipe recrute des doctorants d'excellent niveau et elle a beneficié d'une bourse ministérielle presque tous les ans pendant les six ans qui se sont écoulés.

Les candidats sont en majorité étrangers et ils sont très bien accompagnés pendant tout le processus de recrutement (relecture de dossier, préparation d'auditions, etc.).

Table 3.4 Prix pour l'activité en recherche

Prix	Date
Best paper award de WCO	2017
Award Recipient of IEEE Technical Committee on Cloud Computing Research Innovation Award	2020
Prix de thèse PGMO	2022
Best Artefact Award	2022

L'équipe a de plus une politique de co-encadrement pour garantir un excellent niveau de recherche et une richesse de points de vue aux doctorants.

L'équipe est particulièrement proactive vis-à-vis de la recherche de candidats. Pour les postes CR CNRS, nous contactons presque chaque année les candidat(e)s pressenti(e)s et les incitons à déposer un dossier en mentionnant le LIPN comme unité d'accueil. Nous proposons ensuite une relecture des dossiers. Pour les candidat(e)s auditionné(e)s, nous proposons une simulation d'audition. Le nombre de candidat(e)s est très variable d'une année sur l'autre, mais le mécanisme donne de bons résultats. Par exemple, l'équipe a accompagné dernièrement avec succès la candidature CNRS DR de Claudia D'Ambrosio (promotion) et très récemment la candidature de Sophie Huibert comme CR (classée première au dernier concours). La collaboration avec Claudia a demarré avant le concours CNRS car elle a decidé de se faire accompagner par l'équipe pendant son HDR. Les deux candidatures sont des témoignages évidents que l'actractivité et la visibilité de l'équipe sont remarquables.

L'équipe a aussi une activité soutenue d'invitation de chercheurs étrangers pour consolider des activités déjà entamées ou pour démarrer de nouvelles collaborations. Evidemment, l'activité s'est ralentie pendant le COVID (années 2019-2020), mais nous avons rapidement compensé par des rencontres (séminaires inclus) hebdomadaires en ligne.

- **Post-doctorants**

  — Garraffa Michele, 2016–2017, financement : MathSTIC, encadrement : Létocart & Achir (L2TI)
  — Rigonat Désirée, 2016–2017, financement : DecisionBrain , encadrement : Traversi & Wolfler
  — Menoueur Tarek, 2017–2018, financement : UMANIS, encadrement :Cérin
  — Delle Donne Diego, 2018–2019, financement : MathStic, encadrement :Wolfler & Valencia Pabon (CALIN)
  — Bettiol Enrico, 2022–2023, financement : Huawei, encadrement :Wolfler & Lacroix

- **Doctorants**

Dans les listes suivantes, par souci de concision, seuls les encadrants membres de l'université sont précisés.

Thèses soutenues :

— Pass-Lanneau Adèle, dir. Fouilhoux & Chretienne (LIP6), soutenue en mars 2021, durée : 36 mois, 2 revues et 8 confs, Ingénieure de Recherche à la Direction Générale des Armées (DGA).
— Sliwak Julie, dir. Létocart & Traversi, soutenue en juin 2021, durée : 39 mois, 1 revue et 1 conf, Ingénieure chercheuse GRTgaz.
— Torres Gonzales Daniel, dir. Petrucci & Coti, soutenue en déc. 2021, durée : 39 mois, 4 confs, Research Engineer Inria Paris.
— Chahbar Mohammed, dir. Cerin & Diaz (L2TI), soutenue en nov. 2021, durée : 38 mois, 1 revue et 3 confs, .
— Pan Stefania, dir. Wolfler & Létocart, soutenue en déc. 2018, durée : 36 mois, 2 revues et 2 confs, Consultante Artelys – Italie.
— Nguyen Thi Thanh Huyen, dir. Petrucci & Coti, soutenue en nov. 2018, durée : 44 mois (maternité dans la période), 1 conf, maitre de conférences et chef du département génie logiciel à l'université Hanoi.
— Bettiol Enrico, dir. Létocart & Traversi, soutenue en nov. 2019, durée : 38 mois, 2 revues et 2 confs, Post-doctorant LIPN.
— Lancini Emiliano, dir. Wolfler & Grappe, soutenue en déc. 2019, durée : 39 mois, 3 revues et 1 conf, MCF université Paris Dauphine.

— Torres Figueroa Juan José, dir. Wolfler & Traversi, soutenue en juillet 2022, durée : 46 mois, 1 acept. + 1 soum., Saint Gobain.
— Leclercq Etienne, dir. Roupin & Rouveirol (A3), soutenue en oct. 2022, durée : 60 mois, 5 confs., Ingénieur.
— Kissami Imad, dir. Cérin & Benkhaldoun (LAGA), soutenue en fév. 2017, durée : 41 mois, 2 confs, freelance.
— De Melo Da Silva Marcos, dir. Wolfler & Toulouse, soutenue en sep. 2017, durée : 48 mois, 2 revues, Decision Brain.

Thèses en cours :

— Dupont-Bouillard Alexandre, dir. Fouilhoux & Lacroix, depuis 2020.
— Grenèche Nicolas [1], dir. Cérin, depuis 2020.
— Pisanu Francesco, dir. Wolfler & Grappe, depuis 2020.
— Zigrand Louis, dir. Wolfler & Traversi, depuis 2020.
— Demelas Francesco, dir. Wolfler & Lacroix, depuis 2021.
— Vallée Mathieu, dir. Grappe, depuis 2022.
— Schneider Alexis, dir. Fouilhoux & Létocart, depuis 2022.
— Schulz Alexandre, dir. Wolfler & Le Roux (RCLN), depuis 2022.
— Zhang Yue, dir. Létocart & Fouilhoux, depuis 2022.
— Lahmer Seyyidahmed, dir. Cérin & Diaz (L2TI) & Boussetta (L2TI), depuis 2023.

Auto-évaluation de l'équipe AOC

Au total, 12 thèses ont été soutenues pendant la période d'évaluation. On peut constater que la durée moyenne des thèses dans l'équipe est d'environ 41 mois, ce qui est légèrement supérieur à la durée souhaitable. La politique de l'équipe est de viser 39 mois, correspondant à un début en septembre de l'année N et une soutenance en décembre de l'année N + 3. La pandémie de 2020–2022 a eu une influence négative sur la durée des thèses (par exemple, l'École Doctorale Galilée a introduit un dispositif semi-automatique de prolongement de 3 mois des allocations, applicable sur simple demande d'un(e) doctorant(e) en thèse pendant la pandémie). Une seule thèse dépasse largement la durée souhaitée (impactée par la pandémie), deux de très peu (44 mois) et deux un peu plus (48 et 52 mois, cette dernière egalement impactée par la pandémie). La majorité des thèses (les sept restantes) sont de la durée souhaitée.

- **Chercheurs et enseignants-chercheurs invités**

— Frangioni Antonio, 2017, Prof invité USPN
— Nannicini, Giacomo, 2017, Prof invité USPN
— Naoum-Sawaya, Joe, 2018, Prof invité MathStic
— Rinaldi, Francesco, 2018, Prof invité USPN
— Baldacci, Roberto, 2018, Prof invité USPN
— Jean-Luc, Gaudiot, 2019, Prof invité USPN
— Cüneyt, Güzelis, 2020, Prof invité USPN
— Ríos-Solís, Yasmín, 2019, 2 semaines
— Bomze, Immanuel, 2019, Prof invité USPN
— Marco, Caserta, 2020, Prof invité USPN
— Ríos-Solís, Yasmín, 2021, Prof invité USPN
— Baldacci, Roberto, 2021, Prof invité USPN
— Baldacci, Roberto, 2022, Prof invité USPN
— Maya Duque, Pablo Andres, 2023, 2 semaines

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs.

Le dynamisme et la notoriété de l'équipe ont permis d'obtenir de nombreux projets et contrats de différentes envergures, prouvant ainsi une visibilité nationale et interna-

1. décharge de temps IR par USPN

Table 3.5 Projets de l'équipe

Nom du projet	Années	Source	Montant	Rôle
Exploring the Virtual Network Embedding for 5G network slice design and placement	2022-2025	Orange	130	Porteur
FogSLA	2022-2025	BPI	220	Porteur
Smart Solver	2022-2024	Huawei	200	Porteur
Wolphin	2017-2019	FUI	150	Porteur
Wendelina	2014-2017	FUI	193	Porteur
Résolution des relaxations coniques pour l'optimisation appliquée aux réseaux électriques de grande taille	2018-2021	RTE	130	Participants
Data Mining for improving Dial-a-Ride	2019-2023	PADAM	130	Porteur
Solving workforce scheduling	2016-2019	Horizontal Software	130	Porteur
Optimizing urban transportation planning and management for developing countries	2021-2023	ECONORD	50	Porteur
Parallel Optimization in a Parameter Space (POPS)	2021-2022	PEPS (CNRS)	8	Porteur
Parallel Parametric Linear Programming (PPLip)	2018-2019	BQR	5	Porteur
Graph Learning for Optimization	2022-2023	PGMO	13	Porteur
Graph Learning for Optimization	2021-2022	PGMO	13	Porteur
Optimisation combinatoire pour la simulation numérique d'interfaces	2020-2021	PGMO	6	Porteur
Partitionnement connexe équicoloré des Smartgrids	2016-2019	PGMO	6	Porteur
DISTANCIA	2017-2022	ANR		Membre

tionale dans tous les axes de recherche de l'équipe.

Un des points forts de l'équipe est que la recherche de ses financements est très orientée vers le transfert de technologie et donc de recherche de contrats avec des enterprises. La majorité des contrats ont donné lieu à une thèse CIFRE : 5 sur la période.

Les jeunes chercheurs de l'équipe répondent de manière presque systématique aux appels à projets soit internes aux tutelles (comme BQR USPN et PEPS CNRS) soit de différentes institutions (comme la fondation mathématique Jaques Hadamard), avec un excellent taux de réussite, car l'équipe a déposé avec succès plusieurs projets, à savoir : 1 BQR, 4 PGMO, 1 PEPS.

Les membres plus expérimentés et avec des propositions suffisamment solides répondent aux appels multipartenaires de l'ANR ou à des projets FUI. Malheureusement, aucun financement de projets n'a été obtenu de l'ANR sur la période malgré le fait que presque toutes les demandes aient passé la première phase avec succès. Par contre, nous avons obtenu deux FUI sur la période. Les membres de l'équipe participent aussi aux projets ANR multipartenaires en tant que responsables de site (1 proposition financée sur la période).

3.1.4 Référence 4. L'équipe est attractive par la qualité de ses équipements et de ses compétences techniques.

L'équipe, comme plusieurs autres équipes de recherche en informatique, n'a besoin ni d'acheter ni de produire d'équipement lourd. Côté logiciel, notons la production de logiciels *Open Source* pour la communauté. En particulier, nous avons produit un des meilleurs solver pour la programmation semi-définie positive (logiciel BigCrunch, avec un accès libre en ligne) ainsi que la première librairie internationale d'instances pour la programmation quadratique (QPLib).

3.2 Domaine 3. Production scientifique

	2017	2018	2019	2020	2021	2022	2023	*Total*
Livres	0	0	0	0	0	0	0	*0*
Chapitres de livre	0	0	1	0	0	0	0	*1*
Revues internationales	5	14	7	10	14	13	2	*65*
Autres revues	0	0	0	0	1	1	0	*2*
Conférences internationales	16	17	7	6	4	6	0	*56*
Autres communications	8	2	6	2	2	8	0	*28*
Total	*29*	*33*	*21*	*18*	*21*	*28*	*2*	*152*

Auto-évaluation de l'équipe AOC

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité.

La production scientifique de l'équipe AOC a très bien progressé dans la période 2017-2022. Nous constatons tout d'abord une progression quantitative par rapport à la période 2012-2017 (nous sommes passés de 42 à 64 revues et de 52 à 55 conférences) à effectif constant. Le nouveau professeur est arrivé en 2020, mais en même temps il y a eu un départ à la retraite et deux détachements. Cette augmentation quantitative est naturellement liée à la progression de carrière des Maîtres de Conférence mise en évidence par les deux HDR et un promotion locale professeur.

Plus intéressante et remarquable est l'analyse de la qualité des publications. Plus de 20% des publications sont dans des revues A* (axes 1 et 2) et des conférences A* (axe 3).

Par exemple, huit publications se partagent les revues : *Mathematical Programming*, *Mathematical Programming Computation* et *Operations Research*. Un article a obtenu le très prestigieux prix Euro-Par.

L'avancée scientifique, comme expliqué dans la Section 1.2, nous a permis d'augmenter en visibilité internationale dans plusieurs domaines importants.

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels.

Le nombre total d'ETPT de l'équipe sur la période, limité au membres permanents et tenant compte de la présence partielle de certains membres, est 37.95. Cela donne, en ne considérant que les publications signées par des permanents :

— 64/37.95 = 1.68 articles de revue internationale par personne par an ;
— 55/37.95 = 1.44 articles de conférence internationale par personne par an.

L'analyse des données nous permet de faire remarquer trois choses importantes et significatives de l'équipe AOC :

— La production scientifique de l'équipe repose aussi sur la contribution de ses doctorant(e)s. Sur 12 doctorant(e)s, toutes catégories confondu, ayant soutenu leur thèse pendant la période d'évaluation, tous ont participé à au moins une publication dans une revue ou dans une conférence de premier plan, et le plus souvent deux ou plus.
— Une grande majorité des publications sont co-signées par au moins deux membres de l'équipe. Donc il est utile de souligner que, si chaque article était compté un nombre de fois égal au nombre de membres de l'équipe l'ayant co-signé, les chiffres (1.68 et 1.44) seraient beaucoup plus importants.

— Un certain nombre de membres de l'équipe (entre 20 et 25% selon le moment auquel on compte les effectifs) n'ont pas participé à la production scientifique, en dépit d'un rattachement à temps complet au LIPN tout le long de la période d'évaluation. Les membres non publiants ou peu publiants sont les membres plus impliqués dans des responsabilités collectives et administratives lourdes. Chaque cas est différent, mais les membres qui terminent leur mandat dans une activité administrative lourde sont rapidement integrées dans des projets d'équipe et/ou dans le co-encadrement de thèse, sauf si c'est une volonté personelle de réorienter sa carrière vers des directions alternatives à la recherche. Il s'agit en tout cas de choix personnels où l'équipe ne joue pas de rôle.

3.2.3 Publications

Les références :
- comportant « * » sont acceptées et à paraître,
- indiquées sur fond gris sont les productions phare de chaque catégorie,
- en gras constituent les 20% les meilleures de la production scientifique de chaque catégorie.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2023 [RI-65*] P. Bendotti, P. Chrétienne, P. Fouilhoux et A. Pass-Lanneau. The Anchor-Robust Project Scheduling Problem. *Operations Research*, XX :29 pages, 2023.

[RI-64] M. Lacroix, M. Barbato, R. Grappe et E. Lancini. Box-total dual integrality and edge-connectivity. *Mathematical Programming*, 197 :307–336, 2023.

2022 [RI-63] R. Baldacci, M. Caserta, E. Traversi et R. Wolfler. Robustness of solutions to the capacitated facility location problem with uncertain demand. *Optimization Letters*, 16(9) :2711–2727, 2022.

[RI-62] M. Barbato, R. Grappe, M. Lacroix, E. Lancini et R. Wolfler. The Schrijver System of the Flow Cone in Series-Parallel Graphs. *Discrete Applied Mathematics*, 308 :162–167, 2022.

[RI-61] E. Bettiol, I. Bomze, L. Létocart, F. Rinaldi et E. Traversi. Mining for diamonds - Matrix generation algorithms for binary quadratically constrained quadratic problems. *Computers & Operations Research*, 142 :105735 (12 pages), 2022.

[RI-60] C. Buchheim et E. Traversi. Convex optimization under combinatorial sparsity constraints. *Operations Research Letters*, 50(6) :632–638, 2022.

[**RI-59**] C. Cérin, K. Kimura et M. Sow. Data stream clustering for low-cost machines. *Journal of Parallel and Distributed Computing*, 166 :57–70, 2022.

[RI-58] A. Ceselli, L. Létocart et E. Traversi. Dantzig-Wolfe reformulations for binary quadratic problems. *Mathematical Programming Computation*, 14(1) :85–120, 2022.

[RI-57] S. Choi et M. Vallée. An algorithmic strategy for finding characteristic maps over wedged simplicial complexes. *Pacific Journal of Mathematics*, 320(1) :13–43, 2022.

[RI-56] S. Choi et M. Vallée. Cohomological rigidity of the connected sum of three real projective spaces. *Proceedings of the Steklov Institute of Mathematics*, 317(1) :178–188, 2022.

[RI-55] A. Greco et F. Pisanu. Improvements on overdetermined problems associated to the p-Laplacian. *Mathematics in Engineering*, 4(3) :1–14, 2022.

[RI-54] C. Jiang, Y. Qiu, W. Shi, Z. Ge, J. Wang, S. Chen, C. Cérin, Z. Ren, G. Xu et J. Lin. Characterizing Co-Located Workloads in Alibaba Cloud Datacenters. *IEEE Transactions on Cloud Computing*, 10(4) :2381–2397, 2022.

[RI-53] A. Khedimi, T. Menouer, C. Cérin et M. Boudhar. A cloud weather forecasting service and its relationship with anomaly detection. *Service Oriented Computing and Applications*, 16(3) :191–208, 2022.

[RI-52] Y. Li, D. Ou, X. Zhou, C. Jiang et C. Cérin. Scalability and performance analysis of BDPS in clouds. *Computing*, 104(6) :1425–1460, 2022.

[RI-51] D. Mishra, A. Trotta, E. Traversi, M. Di Felice et E. Natalizio. Cooperative Cellular UAV-to-Everything (C-U2X) communication based on 5G sidelink for UAV swarms. *Computer Communications*, 192(1) :173–184, 2022.

2021

[RI-50] L. Alfandari et S. Toulouse. Approximation of the Double Traveling Salesman Problem with Multiple Stacks. *Theoretical Computer Science*, 877 :74–89, 2021.

[RI-49] P. Alizadeh, E. Traversi et A. Osmani. Deterministic policies based on maximum regrets in MDPs with imprecise rewards. *AI Communications*, 34(4) :229–244, 2021.

[**RI-48**] M. Casazza, A. Ceselli et R. Wolfler. A route decomposition approach for the single commodity Split Pickup and Split Delivery Vehicle Routing Problem. *European Journal of Operational Research*, 289(3) :897–911, 2021.

[RI-47] M. Chahbar, G. Diaz, A. Dandoush, C. Cérin et K. Ghoumid. A Comprehensive Survey on the E2E 5G Network Slicing Model. *IEEE Transactions on Network and Service Management*, 18(1) :49–62, 2021.

[**RI-46**] P. Chervet, R. Grappe et L.-H. Robert. Box-total dual integrality, box-integrality, and equimodular matrices. *Mathematical Programming*, 188(1) :319–349, 2021.

[RI-45] C. Coti et A. D. Malony. DiPOSH : A portable OpenSHMEM implementation for short API-to-network path. *Concurrency and Computation : Practice and Experience*, 33(11) :e6179 (20 pages), 2021.

[**RI-44**] C. Coti, L. Petrucci, C. Rodriguez et M. Sousa. Quasi-Optimal Partial Order Reduction. *Formal Methods in System Design*, 57 :3–33, 2021.

[RI-43] D. Delle Donne, F. Furini, E. Malaguti et R. Wolfler. A branch-and-price algorithm for the Minimum Sum Coloring Problem. *Discrete Applied Mathematics*, 303 :39–56, 2021.

[RI-42] A. Frangioni, S. Pan, E. Traversi et R. Wolfler. A constraints-aware Reweighted Feasibility Pump approach. *Operations Research Letters*, 49(5) :671–675, 2021.

[**RI-41**] T. Gschwind, S. Irnich, F. Furini et R. Wolfler. A Branch-and-Price Framework for Decomposing Graphs into Relaxed Cliques. *INFORMS Journal on Computing*, 33(3) :1070–1090, 2021.

[RI-40] C.-H. Hsu, K. Amilineni, H. Wu, C. Cérin et Y.-C. Chung. A novel scalable network architecture for the evolution of robot swarm networks with dual connectivity application in control-data plane. *Swarm and Evolutionary Computation*, 64 :100889 (11 pages), 2021.

[RI-39] M. Lacroix, Y. A. Ríos-Solís et R. Wolfler Calvo. Efficient formulations for the traveling car renter problem and its quota variant. *0ptimization Letters*, 15(6) :1905–1930, 2021.

[RI-38] L. Létocart, J. Sliwak, E. Andersen, M. Anjos et E. Traversi. A Clique Merging Algorithm to Solve Semidefinite Relaxations of Optimal Power Flow Problems. *IEEE Transactions on Power Systems*, 36(2) :1641–1644, 2021.

[**RI-37**] G. Nannicini, E. Traversi et R. Wolfler. A Benders squared (B^2) framework for infinite-horizon stochastic linear programs. *Mathematical Programming Computation*, 13(4) :645–681, 2021.

2020

[RI-36] H. Abbes, T. Louati et C. Cérin. Dynamic replication factor model for Linux containers-based cloud systems. *The Journal of Supercomputing*, 76(9) :7219–7241, 2020.

[**RI-35**] R. Baldacci, A. Lim, E. Traversi et R. Wolfler. Optimal Solution of Vehicle Routing Problems with Fractional Objective Function. *Transport Science*, 54(2) :434–452, 2020.

[RI-34] E. Bettiol, L. Létocart, F. Rinaldi et E. Traversi. A conjugate direction based simplicial decomposition framework for solving a specific class of dense convex quadratic programs. *Computational Optimization and Applications volume*, 75(2) :321–360, 2020.

[RI-33] M. Braga, D. Delle Donne, M. S. Escalante, J. L. Marenco, M. E. Ugarte et M. C. Varaldo. The minimum chromatic violation problem : A polyhedral approach. *Discrete Applied Mathematics*, 281 :69–80, 2020.

[RI-32] C. Coti, D. Monniaux et H. Yu. A task-based approach to parallel parametric linear programming solving, and application to polyhedral computations. *Concurrency and Computation : Practice and Experience*, 33(6) :e6050 (23 pages), 2020.

[RI-31] F. Furini, I. Ljubic et E. Traversi. Preface : decomposition methods for hard optimization problems. *Annal of Operational Research*, 284(2) :483–485, 2020.

[RI-30] C. Jiang, Y. Qiu, W. Shi, Z. Ge, J. Wang, S. Chen, C. Cérin, Z. Ren, G. Xu et J. Lin. Characterizing Co-located Workloads in Alibaba Cloud Datacenters. *IEEE Transactions on Cloud Computing*, 10(4) :2381–2397,, 2020.

[RI-29] C. Jiang, T. Fan, H. Gao, W. Shi, L. Liu, C. Cérin et J. Wan. Energy aware edge computing : A survey. *Computer Communications*, 151 :556 – 580, 2020.

[RI-28] T. Menouer, C. Cérin et C.-H. Hsu. Opportunistic scheduling and resources consolidation system based on a new economic model. *The Journal of Supercomputing*, 76 :9942–9975, 2020.

[**RI-27**] G. Nannicini, G. Sartor, E. Traversi et R. Wolfler. An exact algorithm for robust influence maximization. *Mathematical Programming*, 183(1) :419–453, 2020.

2019 [RI-26] G. Beck, T. Duong, M. Lebbah, H. Azzag et C. Cérin. A distributed approximate nearest neighbors algorithm for efficient large scale mean shift clustering. *Journal of Parallel and Distributed Computing*, 134 :128–139, 2019.

[RI-25] D. Cornaz, F. Furini, M. Lacroix, E. Malaguti, R. Mahjoub et S. Martin. The vertex k-cut problem. *Discrete Optimization*, 31(31) :8–28, 2019.

[RI-24] D. Cornaz, R. Grappe et M. Lacroix. Trader multiflow and box-TDI systems in series-parallel graphs. *Discrete Optimization*, 31(31) :103–114, 2019.

[RI-23] F. Furini et E. Traversi. Theoretical and computational study of several linearisation techniques for binary quadratic problems. *Annals of Operational Research*, 279(1-2) :387–411, 2019.

[**RI-22**] F. Furini, E. Traversi, P. Belotti, A. Frangioni, A. Gleixner, N. Gould, L. Liberti, A. Lodi, R. Misener, H. Mittelmann, N. V. Sahinidis, S. Vigerske et A. Wiegele. QPLIB : a library of quadratic programming instances. *Mathematical Programming Computation*, 11(2) :237–265, 2019.

[RI-21] Y. Ngoko, C. Cérin et D. Trystram. Solving Sat in a Distributed Cloud : A Portfolio Approach. *International Journal of Applied Mathematics and Computer Science*, 29(2) :261–274, 2019.

[RI-20] S. Pan, R. Wolfler Calvo, M. Akplogan, L. Létocart et N. Touati. A dual ascent heuristic for obtaining a lower bound of the generalized set partitioning problem with convexity constraints. *Discrete Optimization*, 33 :146–168, 2019.

2018 [RI-19] M. Barbato, R. Grappe, M. Lacroix et C. Pira. Lexicographical polytopes. *Discrete Applied Mathematics*, 240 :3–7, 2018. 6 pages.

[RI-18] S. Borne, E. Gourdin, O. Klopfenstein et R. Mahjoub. A Branch-and-Cut Algorithm for the Capacitated Multi-Failure Survivable Network Design Problem. *Computers & Industrial Engineering*, 124 :582 – 603, 2018.

[**RI-17**] C. Buchheim et E. Traversi. Quadratic Combinatorial Optimization Using Separable Underestimators. *INFORMS Journal on Computing*, 30(3) :424–437, 2018.

[RI-16] M. Casazza, A. Ceselli, D. Chemla, F. Meunier et R. Wolfler. The multiple vehicle balancing problem. *Networks*, 72(3) :337–357, 2018.

[RI-15] C. D'Ambrosio, F. Furini, M. Monaci et E. Traversi. On the Product Knapsack Problem. *Optimization Letters*, 12(4) :691–712, 2018.

[**RI-14**] M. de Melo da Silva. The Block Retrieval Problem. *European Journal of Operational Research*, 265(3) :931–950, 2018.

[**RI-13**] M. de Melo da Silva, S. Toulouse et R. Wolfler. A new effective unified model for solving the Pre-marshalling and Block Relocation Problems. *European Journal of Operational Research*, 271(1) :40–56, 2018.

[RI-12] F. Furini, M. Monaci et E. Traversi. Exact approaches for the knapsack problem with setups. *Computers & Operations Research*, 90 :208–220, 2018.

[RI-11] R. Grappe et M. Lacroix. The st-bond polytope on series-parallel graphs. *RAIRO - Operational Research*, 52(3) :923–934, 2018.

[RI-10] G. Laporte, F. Meunier et R. Wolfler. Shared mobility systems : an updated survey. *Annals of Operational Research*, 271(1) :105–126, 2018.

[RI-9] L. Létocart, R. Wolfler Calvo et R. Baldacci. Preface : Emerging challenges in transportation planning. *Networks*, 72(3) :309–310, 2018.

[RI-8] T. Louati, H. Abbes et C. Cérin. LXCloudFT : Towards high availability, fault tolerant Cloud system based Linux Containers. *Journal of Parallel Distributed Computing*, 122 :51–69, 2018.

[RI-7] T. Louati, H. Abbes, C. Cérin et M. Jemni. LXCloud-CR : Towards LinuX Containers Distributed Hash Table based Checkpoint-Restart. *Journal of Parallel and Distributed Computing*, 111 :187–205, 2018.

[RI-6] W. van Ackooij, J. De Boeck, B. Detienne, S. Pan et M. Poss. Optimizing power generation in the presence of micro-grids. *Eur. J. Oper. Res.*, 271(2) :450–461, 2018.

2017 [RI-5] N. Avcu, N. Pekergin, F. Pekergin et C. Güzelis. Aggregation for Computing Multi-Modal Stationary Distributions in 1-D Gene Regulatory Networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(3) :813–827, 2017.

[RI-4] R. Baldacci, S. Ulrich et R. Wolfler. The Vehicle Routing Problem with Transhipment Facilities. *Transport Science*, 51(2) :592–606, 2017.

[RI-3] A. Bernáth, R. Grappe et Z. Szigeti. Partition constrained covering of a symmetric crossing supermodular function by a graph. *SIAM Journal on Discrete Mathematics*, 31(1) :335–382, 2017.

[**RI-2**] V. Denoyel, L. Alfandari et A. Thiele. Optimizing Healthcare Network Design under Reference Pricing and Parameter Uncertainty. *European Journal of Operational Research*, 263(3) :996–1006, 2017.

[RI-1] N. Krislock, J. Malick et F. Roupin. Biqcrunch : a semidefinite branch-and-bound method for solving binary quadratic problems. *ACM Transactions on Mathematical Software*, 43(4) :32 :1–23, 2017.

Articles dans des revues nationales ou internationales

2022 [RE-2] G. L. Israel, A. De Luca, M. Bachetti, A. Belfiore, P. Esposito, M. Imbrogno, A. Miraval Zanon, F. Pisanu, G. Rodriguez Castillo, R. Salvaterra, A. Tiengo et A. Wolter. The Extragalactic Population of Accreting Neutron Stars and the Ultraluminous X-ray Sources Paradigm : from ULTraS to PULSULTraS. *Journal of the Italian Astronomical Society*, 93(2) :34–43, 2022.

2021 [RE-1] J. Yin, C. Jiang, H. Mino et C. Cérin. Popularity-Aware In-Network Caching for Edge Named Data Network. *Wireless Communication and Mobile Computing*, 2021 :379185 (13 pages), 2021.

Chapitres de livre

2019 [CL-1] S. Benbernou, M. Bentounsi, C. Cérin, M. Ouziri, P. Garteiser, L. Abidi, M. Lebbah, H. Azzag et M. Smadja. *A Big Data Platform for Enhancing Life Imaging Activities*, J. Darmont et S. Loudcher, éditeurs, *Utilizing Big Data Paradigms for Business Intelligence*, pages 39–71. IGI Global, 2019.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2022 [CI-56] C. Alizadeh, P. Alizadeh, M. Anjos, L. Létocart et E. Traversi. How to Learn the Optimal Clique Decompositions in Solving Semidefinite Relaxations for OPF. *The 2022 International Joint Conference on Neural Networks (IJCNN 2022), Padua, Italy, July 18-23, 2022*, pages 1–8, 2022.

[CI-55] C. Coti, L. Petrucci et D. A. Torres Gonzalez. A Formal Model for Fault Tolerant Parallel Matrix Factorization. *The 26th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'22), Hiroshima City, Japan, March 26-30, 2022*, pages 62–70, 2022.

[**CI-54**] N. Grenèche, T. Menouer, C. Cérin et O. Richard. A Methodology to Scale Containerized HPC Infrastructures in the Cloud. J. Cano et P. Trinder, éditeurs, *The 28th International Conference on Parallel and Distributed Computing (Euro-Par 2022), Glasgow, UK, August 22-26, 2022.*, volume 13440 de *Lecture Notes in Computer Science*, pages 203–217, 2022.

[CI-53] E. Traversi, I. Zhilin, A. Almarzooqi, A. Alketbi, A. Al Mansoori et G. De Masi. Multi-Agent Routing Optimization for Underwater Monitoring. *The 20th International Conference on Advances in Practical Applications of Agents, Multi-Agent Systems, and Complex Systems Simulation (PAAMS 2022), L'Aquila, Italy, July 13-15, 2022*, volume 13616 de *Lecture Notes in Artifical Intelligence*, pages 397–409, 2022.

[CI-52] T. Zhang, D. Ou, Z. Ge, C. Jiang, C. Cérin et L. Yan. Interference-aware Workload Scheduling in Co-located Data Centers. *The 9th IFIP WG 10.3 International Conference on Network and Parallel Computing (NPC 2022), Jinan, China, September 24-25, 2022*, volume 13615 de *Lecture Notes in Computer Science*, pages 69–82, 2022.

[**CI-51**] L. Zigrand, P. Alizadeh, E. Traversi et R. Wolfler Calvo. A Scalable Logic-based Benders Decomposition to Optimize a Dynamic Demand-Responsive Transport System. *The XI TRISTAN Triennial Symposium on Transportation Analysis*, 2022.

2021 [**CI-50**] C. Coti, L. Petrucci et D. A. Torres Gonzalez. Fault-tolerant LU factorisation is low cost. *The 27th International European Conference on Parallel and Distributed Computing (Euro-Par'21), Lisbon, Portugal, September 1-3, 2021*, volume 12820 de *Lecture Notes in Computer Science*, pages 536–549, 2021.

[CI-49] T. Menouer, N. Grenèche, C. Cérin et P. Darmon. Towards an Optimized Containerization of HPC Job Schedulers Based on Namespaces. *The 18th IFIP WG 10.3 International Conference on Network and Parallel Computing (NPC 2021), Paris, France, November 3-5, 2021*, volume 13152 de *Lecture Notes in Computer Science*, pages 144–156, 2021.

[CI-48] T. Yuan, D. Ou, J. Wang, C. Jiang, C. Cérin et L. Yan. PPCTS : Performance Prediction-Based Co-located Task Scheduling in Clouds. *The 21st International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2021), Virtual Event, December 3-5, 2021*, volume 13157 de *Lecture Notes in Computer Science*, pages 245–257, 2021.

[**CI-47**] L. Zigrand, P. Alizadeh, E. Traversi et R. Wolfler Calvo. Machine Learning Guided Optimization for Demand Responsive Transport Systems. *The 2021 Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2021), Bilbao, Spain, September 13-17, 2021*, volume 12978 de *Lecture Notes in Computer Science*, pages 420–436, 2021.

2020 [CI-46] J. Ben Geloun, C. Coti et A. D. Malony. On-the-fly Optimization of Parallel Computation of Symbolic Symplectic Invariants. *The 19th International Symposium on Parallel and Distributed Computing (ISPDC 2020), Warsaw, Poland, July 5-8, 2020*, pages 102–109, 2020.

[CI-45] C. Cérin, N. Grenèche et T. Menouer. Towards Pervasive Containerization of HPC Job Schedulers. *The 32nd IEEE International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD 2020), Porto, Portugal, September 9-11, 2020*, pages 281–288, 2020.

[CI-44] M. Lacroix, M. Barbato, R. Grappe et E. Lancini. On k-edge-connected Polyhedra : Box-TDIness in Series-Parallel Graphs. *The 6th International Symposium on Combinatorial Optimization (ISCO 2020), Montréal, QC, Canada, May 4-6, 2020*, volume 12176 de *Lecture Notes in Computer Science*, pages 27–41. Springer, 2020.

[CI-43] T. Menouer, A. Khedimi et C. Cérin. Smart Network Slices Scheduling in Cloud. *The 5th IEEE International Conference on Smart Cloud (SmartCloud 2020), Washington, DC, USA, November 6-8, 2020*, pages 49–54, 2020.

[CI-42] T. Menouer, A. Khedimi, C. Cérin et M. Chahbar. Scheduling Service Function Chains with Dependencies in the Cloud. *The 9th IEEE International Conference on Cloud Networking (Cloudnet 2020), Piscataway, NJ, USA, 9-11 November 2020*, pages 1–3, 2020.

[CI-41] T. Menouer, A. Khedimi, C. Cérin et C. Jiang. Cloud Allocation and Consolidation Based on a Scalability Metric. *The 20th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2020), New York City, NY, USA, October 2-4, 2020,*, volume 12454 de *Lecture Notes in Computer Science*, pages 381–395, 2020.

2019

[CI-40] M. Chahbar, G. Diaz, A. Dandoush, C. Cérin et K. Ghoumid. NESSMA : Network Slice Subnet Management Framework. *The 10th International Conference on Networks of the Future (NoF 2019), Rome, Italy, October 1-3, 2019*, pages 54–57, 2019.

[**CI-39**] C. Coti, D. Monniaux et H. Yu. Parallel parametric linear programming solving, and application to polyhedral computations. *The 19th International Conference on Computational Science (ICCS 2019), Faro, Portugal, June 12-14, 2019*, volume 11540 de *Lecture Notes in Computer Science*, pages 566–572, 2019.

[**CI-38**] J. Le Roux, A. Rozenknop et M. Lacroix. Representation Learning and Dynamic Programming for Arc-Hybrid Parsing. *The 23rd Conference on Computational Natural Language Learning (CoNLL 2019), Hong Kong, China, November 3-4, 2019*, pages 238–248, 2019.

[CI-37] T. Menouer, C. Cérin et P. Darmon. Accelerated Promethee Algorithm Based on Dimensionality Reduction. *The 6th International Conference on Internet of Vehicles (IOV 2019), Kaohsiung, Taiwan, November 18-21, 2019*, volume 11894 de *Lecture Notes in Computer Science*, pages 190–203, 2019.

[CI-36] T. Menouer, C. Cérin, C. Jiang, J. Rivalan et T. Menouer. SAFC : Scheduling and Allocation Framework for Containers in a Cloud Environment. *The 17th International Conference on High Performance Computing & Simulation (HPCS 2019), Dublin, Ireland, July 15-19, 2019*, pages 825–832, 2019.

[**CI-35**] G. Nannicini, G. Sartor, E. Traversi et R. Wolfler. An Exact Algorithm for Robust Influence Maximization. *The 20th International Conference on Integer Programming and Combinatorial Optimization (IPCO 2019), Ann Arbor, MI, USA, May 22-24, 2019*, volume 11480 de *Lecture Notes in Computer Science*, pages 313–326, 2019.

[CI-34] J. Sliwak, M. Ruiz, M. Anjos, L. Létocart et E. Traversi. A Julia module for polynomial optimization with complex variables applied to optimal power flow. *The 2019 IEEE Milan PowerTech (PowerTech 2019), Milan, Italy, June 23-27, 2019*, pages 1–6, 2019.

2018

[CI-33] O. Ben Maaouia, H. Fkaier, C. Cérin et M. Jemni. Energy-Efficient Strategy for Placement of Online Services on Dynamic Availability Resources in Volunteer Cloud. *2018 IEEE International Conference on Parallel & Distributed Processing with Applications (ISPA 2018), Melbourne, VIC, Australia, December 11-13, 2018*, pages 629–636, 2018.

[CI-32] O. Ben Maaouia, H. Fkaier, C. Cérin, M. Jemni et Y. Ngoko. On Optimization of Energy Consumption in a Volunteer Cloud - Strategy of Placement and Migration of Dynamic Services. *The 18th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2018), Guangzhou, China, November 15-17, 2018*, volume 11335 de *Lecture Notes in Computer Science*, pages 388–398, 2018.

[CI-31] F. Butelle et C. Coti. Distributed snapshot for rollback-recovery with one-sided communications. *2018 International Conference on High Performance Computing & Simulation (HPCS), Orleans, France, July 16-20, 2018*, pages 614–620, 2018.

[CI-30] M. Casazza, A. Ceselli et R. Wolfler. A branch and price approach for the Split Pickup and Split Delivery VRP. *Joint EURO/ALIO International Conference 2018 on Applied Combinatorial Optimization (EURO/ALIO 2018), Bologna, Italy, June 25-27, 2018*, volume 69 de *Electronic Notes in Discrete Mathematics*, pages 189–196, 2018.

[CI-29] C. Cérin, T. Menouer et M. Lebbah. Accelerating the Computation of Multi-Objectives Scheduling Solutions for Cloud Computing. *The 8th IEEE International Symposium on Cloud and Service Computing (SC2 2018), Paris, France, November 18-21, 2018*, pages 49–56, 2018.

[**CI-28**] C. Coti, S. Evangelista et L. Petrucci. One-Sided Communications for more Efficient Parallel State Space Exploration over RDMA Clusters. *The 24th International European Conference on Parallel and Distributed Computing (Euro-Par'18), Torino, Italy, August 27-31, 2018*, volume 11014 de *Lecture Notes in Computer Science*, pages 432–446, 2018.

[CI-27] C. Coti, S. Evangelista et L. Petrucci. State Compression Based on One-Sided Communications for Distributed Model Checking. *The 23rd IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'18), Melbourne, Australia, December 12-14, 2018*, pages 41–50, 2018.

[CI-26] J.-F. Culus et S. Toulouse. 2 CSPs All Are Approximable Within a Constant Differential Factor. *The 5th International Symposium on Combinatorial Optimization (ISCO 2018),*

Marrakesh, Morocco, April 11-13, 2018, volume 10856 de *Lecture Notes in Computer Science*, pages 389–401, 2018.

[CI-25] J.-F. Culus et S. Toulouse. How Far From a Worst Solution a Random Solution of a k CSP Instance Can Be ? *The 29th International Workshop on Combinatorial Algorithms (IWOCA 2018), Singapore, July 16-19, 2018,*, volume 10979 de *Lecture Notes in Computer Science*, pages 374–386, Singapore, Singapore, 2018.

[CI-24] J. David, R. Grappe, M. Lacroix et E. Traversi. Self-sufficient sets in smartgrids. *The Joint EURO/ALIO International Conference 2018 on Applied Combinatorial Optimization (EURO/ALIO 2018), Viña del Mar, Chile, April 11–13, 2022*, volume 69 de *Electronic Notes in Discrete Mathematics*, pages 301–308, 2018.

[CI-23] C. Jiang, Y. Wang, D. Ou, Y. Qiu, Y. Li, J. Wan, B. Luo, W. Shi et C. Cérin. EASE : Energy Efficiency and Proportionality Aware Virtual Machine Scheduling. *The 30th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD 2018), Lyon, France, September 24-27, 2018*, pages 65–68, 2018.

[CI-22] T. Menouer, C. Cérin et E. Leclercq. New Multi-objectives Scheduling Strategies in Docker SwarmKit. *The 18th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2018), Guangzhou, China, November 15-17, 2018*, volume 11336 de *Lecture Notes in Computer Science*, pages 103–117, 2018.

[CI-21] T. Menouer, C. Cérin et Y. Ngoko. Efficient scheduling in a smart building. *The 10th International Conference on Management of Digital EcoSystems (MEDES 2018), Tokyo, Japan, September 25-28, 2018*, pages 80–86, 2018.

[**CI-20**] T. Menouer, C. Cérin, W. Saad et X. Shi. A Resource Allocation Framework with Qualitative and Quantitative SLA Classes. *The 24th International European Conference on Parallel and Distributed Computing (Euro-Par 2018), Turin, Italy, August 27-31, 2018*, volume 11339 de *Lecture Notes in Computer Science*, pages 69–81, 2018.

[CI-19] Y. Ngoko, N. Saintherant, C. Cérin et D. Trystram. Invited Paper : How Future Buildings Could Redefine Distributed Computing. *The 2018 IEEE International Parallel and Distributed Processing Symposium Workshops, (PDPSW 2018), Vancouver, BC, Canada, May 21-25, 2018*, pages 1232–1240, 2018.

[**CI-18**] T. T. H. Nguyen, C. Rodriguez, M. Sousa, C. Coti et L. Petrucci. Quasi-Optimal Partial Order Reduction. *The 30th International Conference on Computer Aided Verification (CAV 2018), Oxford, UK, July 14-17, 2018*, volume 10982 de *Lecture Notes in Computer Science*, pages 354–371, 2018.

[CI-17] S. Pan, M. Akplogan, N. Touati, L. Létocart, R. Wolfler Calvo et L.-M. Rousseau. A hybrid heuristic for the multi-activity tour scheduling problem. *Joint EURO/ALIO International Conference 2018 on Applied Combinatorial Optimization (EURO/ALIO 2018), Bologna, Italy, June 25-27, 2018*, volume 69 de *Electronic Notes in Discrete Mathematics*, pages 333–340. Elsevier, 2018.

2017 [CI-16] L. Abidi, C. Cérin et W. Saad. A Deployment System for highly Heterogeneous and Dynamic Environments. *The 2017 IEEE International Conference on High Performance Computing & Simulation (HPCS 2017), Genova, Italy, July 17-21, 2017*, pages 627–634, 2017.

[CI-15] M. Bekhti, M. Garraffa, N. Achir, K. Boussetta et L. Létocart. Assessment of Multi-UAVs Tracking for Data Gathering. *The 13th International Wireless Communications and Mobile Computing Conference (IWCMC 2017), Valencia, Espagne, June 26-30, 2017*, pages 1004–1009, 2017.

[CI-14] O. Ben Maaouia, C. Cérin, H. Fkaier et M. Jemni. A Novel Optimization Technique for Mastering Energy Consumption in Cloud Data Center. *The 2017 IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA 2017), December 12-15, 2017*, pages 475–480, 2017.

[CI-13] O. Ben Maaouia, M. Jemni, H. Fkaier et C. Cérin. Towards optimizing energy consumption in Cloud. *The 2017 International Conference on Engineering & MIS (ICEMIS 2017), Monastir, Tunisia, May 8-10, 2017*, pages 1–7, 2017.

[CI-12] F. Benkhaldoun, C. Cérin, I. Kissami et W. Saad. Challenges of Translating HPC codes to Workflows for Heterogeneous and Dynamic Environments. *The 15th International Conference on High Performance Computing and Simulation (HPCS 2017), Genoa, Italy July 17–21, 2017*, pages 858–863, 2017.

[CI-11] C. Cérin, J.-L. Gaudiot, M. Lebbah et F. Yuehgoh. Return of experience on the mean-shift clustering for heterogeneous architecture use case. *The 2017 IEEE International Conference on Big Data (BigData 2017), Boston, MA, USA, December 11-14, 2017*, pages 3499–3507, 2017.

[CI-10] C. Cérin, T. Menouer, W. Saad et W. Ben Abdallah. A New Docker Swarm Scheduling Strategy. *The 7th IEEE International Symposium on Cloud and Service Computing, (SC2 2017), Kanazawa, Japan, November 22-25, 2017*, pages 112–117, 2017.

[**CI-9**] C. Corro, J. Le Roux et M. Lacroix. Efficient Discontinuous Phrase-Structure Parsing via the Generalized Maximum Spanning Arborescence. *The 2017 Conference on Empirical Methods in Natural Language Processing (EMLNP 2017), Copenhagen, Denmark, September 7–11, 2017*, pages 1644 – 1654, 2017.

[CI-8] C. Coti, E. Leclercq, F. Roupin et F. Butelle. Solving 0-1 Quadratic Problems with Two-Level Parallelization of the BiqCrunch Solver. *The 2017 Federated Conference on Computer Science and Information Systems (FedCSIS 2017), Prague, Czech Republic, September 3-6, 2017*, volume 11 de *Annals of Computer Science and Information Systems*, pages 445–452, 2017. (Best Paper Award).

[CI-7] C. Coti, J.-V. Loddo et E. Viennet. An Overview of the Options Available for Practical Activities in Distributed Computing. *The 16th International Conference on Information Technology Based Higher Education and Training (ITHET 17), Ohrid, Macedonia, Jumy 10-12, 2017*, pages 1–2, 2017.

[CI-6] T. Louati, H. Abbes, C. Cérin et M. Jemni. GC-CR : A Decentralized Garbage Collector Component for Checkpointing in Clouds. *29th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD 2017), Campinas, Brazil, October 17-20, 2017*, pages 97–104, 2017.

[CI-5] T. Menouer et C. Cérin. Scheduling and Resource Management Allocation System Combined with an Economic Model. *The 2017 IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA 2017), Guangzhou, China, December 12-15, 2017*, pages 807–813, 2017.

[CI-4] Y. Ngoko et C. Cérin. An edge computing platform for the detection of acoustic events. *The 2017 IEEE International Conference on Edge Computing (EDGE 2017), Honolulu, HI, USA, June 25-30, 2017*, pages 240–243, 2017.

[CI-3] Y. Ngoko et C. Cérin. Automated Planner for Energy-Efficient Buildings. *The 10th IEEE Conference on Service-Oriented Computing and Applications (SOCA 2017), Kanazawa, Japan, November 22-25, 2017*, pages 147–154, 2017.

[CI-2] Y. Ngoko et C. Cérin. Reducing the number of comatose servers : automatic tuning as an opportunistic cloud-service. *The 2017 IEEE International Conference on Services Computing (SCC 2017), Honolulu, HI, USA, June 25-30, 2017*, pages 487–490, 2017.

[CI-1] Y. Ngoko, D. Trystram et C. Cérin. A Distributed Cloud Service for the Resolution of SAT. *The 7th IEEE International Symposium on Cloud and Service Computing (SC2 2017), Kanazawa, Japan, November 22-25, 2017*, pages 1–8, 2017.

Autres articles dans des conférences ou workshops à comité de lecture

2022 [CO-28] P. Bendotti, P. Fouilhoux, C. Rottner et T. Vignon. Overlapping decomposition in column generation. *PGMODAYS 2022, Palaiseau, France, November 29-30, 2022*, 2022.

[CO-27] E. Bettiol, L. Létocart, F. Rinaldi et E. Traversi. A simplicial decomposition - Branch and Price for convex quadratic mixed binary problems. *The 18th Cologne-Twente Workshop on Graphs and Combinatorial Optimization (CTW 2020), Virtual Event, September 14-16, 2020*, pages 1–4, 2022.

[CO-26] L. Closson, C. Cérin, D. Donsez et D. Trystram. Towards a Methodology for the Characterization of IoT Data Sets of the Smart Building Sector. *The 8th IEEE International Smart Cities Conference (ISC2 2022), Pafos, Cyprus, September 26-29, 2022*, pages 1–7, 2022.

[CO-25] A. Dupont-Bouillard, P. Fouilhoux, R. Grappe et M. Lacroix. Cut and Price pour le problème de coloration. *23ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF 2022), Lyon, France, February 23-25, 2022*, pages 1–2, 2022.

[CO-24] P. Fouilhoux, L. Létocart et Y. Zhang. Adding Valid Inequalities within Bi-objective Binary Branch & Bound. *The 2022 Workshop Recent Advances in Multi-Objective Optimization (RAMOO 2022), Coimbra, Portugal, September 15, 2022*, 2022.

[CO-23] N. Grenèche et C. Cérin. Autoscaling of containerized HPC Clusters in the Cloud. *The 2022 IEEE/ACM International Workshop on Interoperability of Supercomputing and Cloud Technologies (SuperCompCloud 2022), Dallas, TX, USA, November 13-18, 2022*, pages 1–7, 2022.

[CO-22] E. Leclercq, J. Rivalan, F. Roupin et C. Rouveirol. Allocation de ressources par une méthode hybride machine learning - optimisation en contexte de conteneurs. *Extraction et Gestion des Connaissances (EGC 2022), Blois, France, 24-28 janvier 2022*, volume E-38 de *Revue des Nouvelles Technologies de l'Information*, pages 499–500, 2022.

[CO-21] A. Pass-Lanneau, P. Bendotti, P. Chrétienne et P. Fouilhoux. Ancrage et robustesse pour le RCPSP : outils exacts et heuristiques. *23ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF 2022), Lyon, February 23-25, 2022*, pages 1–6, 2022.

2021

[CO-20] C. Cérin, F. Andrès et D. Geldwerth-Feniger. Towards an emulation tool based on ontologies and data life cycles for studying smart buildings. *The 2021 International Workshop on Big Data in Emergent Distributed Environments in conjunction with the 2021 ACM SIGMOD/PODS Conference (BiDEDE@SIGMOD 2021), Virtual Event, China, 20 June, 2021*, pages 8 :1–8 :15, 2021.

[CO-19] C. Coti et A. D. Malony. SKaMPI-OpenSHMEM : Measuring OpenSHMEM Communication Routines. *The 8th Workshop on OpenSHMEM and Related Technologies (OpenSHMEM 2021), Virtual Events, Spetmber 14-16, 2021*, volume 13159 de *Lecture Notes in Computer Science*, 2021.

2020

[CO-18] C. Coti, J. E. Denny, K. Huck, S. Lee, A. D. Malony, S. Shende et J. S. Vetter. OpenACC Profiling Support for Clang and LLVM using Clacc and TAU. *The 2020 IEEE/ACM International Workshop on HPC User Support Tools and Workshop on Programming and Performance Visualization Tools (HUST/ProTools@SC 2020), Atlanta, GA, USA, November 18, 2020*, pages 38–48, 2020. Workshop de supercomputing.

[CO-17] J. Sliwak, M. F. Anjos, L. Létocart, J. Maeght et E. Traversi. Improving Clique Decompositions of Semidefinite Relaxations for Optimal Power Flow Problems. *21ème congrès annuel de la société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF 2020), Montpellier, France, 19-21 Février 2020*, pages 1–8, 2020.

2019

[CO-16] P. Alizadeh, A. Osmani et E. Traversi. Calcul d'une politique déterministe dans un MDP avec récompenses imprécises. *Extraction et Gestion des connaissances (EGC 2019), Metz, France, January 21-25, 2019*, volume E-35 de *Revue des Nouvelles technologies*, pages 45–56, 2019.

[CO-15] C. Cérin, W. Saad, C. Jiang et E. Mekni. Where are the Optimization Potential of Machine Learning Kernels ? *The 5th International Conference on Big Data Intelligence and Computing (DATACOM 2019), Kaohsiung, Taiwan, November 18-21, 2019*, pages 130–136, 2019.

[CO-14] C. Coti, L. Petrucci et D. A. Torres Gonzalez. Fault-tolerant matrix factorisation : a formal model and proof. *The 6th International Workshop on Synthesis of Complex Parameters (SynCoP'19), Prague, Czech Republic, April 6-7, 2019*, pages 1–4, 2019.

[CO-13] C. Coti, L. Petrucci et D. A. Torres Gonzalez. Process scheduling on volatile nodes for fault-tolerant linear algebra. *The 3rd Workshop on Data Locality (COLOC'19), Göttingen, Germany, August 26, 2019*, 2019. (Poster).

[CO-12] C. Coti et A. D. Malony. On the road to DiPOSH : Adventures in high-performance OpenSHMEM. *The 13th International Conference on Parallel Processing and Applied Mathematics (PPAM 2019), Bialystok, Poland, September 8-11, 2019, Revised Selected Papers*, volume 12043 de *Lecture Notes in Computer Science*, pages 250–260, 2019.

[CO-11] T. Menouer, O. Manad, C. Cérin et P. Darmon. Power Efficiency Containers Scheduling Approach Based on Machine Learning Technique for Cloud Computing Environment. *The 16th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN 2019), Naples, Italy, September 16-20, 2019*, pages 193–206, 2019.

2018 [CO-10] S. Bejaoui, C. Cérin, T. Menouer et W. Saad. Towards a Two Layers Based Scheduling Schema for Data Aware Strategies. *The 4th IEEE International Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress (CyberSciTech 2018), Athens, Greece, August 12-15, 2018*, pages 1060–1067, 2018.

[CO-9] L. Létocart, M. Garraffa, M. Bekhti, N. Achir et K. Boussetta. Drones path planning for WSN data gathering : A column generation heuristic approach. *The 2018 IEEE Wireless Communications and Networking Conference (WCNC 2018), Barcelona, France, April 15-18, 2018*, pages 1–6, 2018.

2017 [CO-8] F. Benkhaldoun, C. Cérin, I. Kissami et W. Saad. HPC as a service for computational fluid dynamics problems. *The 11th International Workshop on Parallel Numerics (PARNUM 2017), Waischenfeld, Germany, April 19–21, 2017*, pages 8–9, 2017.

[CO-7] E. Bettiol, L. Létocart, F. Rinaldi et E. Traversi. Simplicial Decomposition for Large Scale Quadratic Convex Programming. *18ème congrès de la Société Françcaise de Recherche Opérationnelle et d'Aide à la Décision (ROADEF'2017), Metz, France, 22-24 Février 2017*, pages 1–4, 2017.

[CO-6] F. Butelle, C. Coti, E. Leclercq et F. Roupin. Une version Multithread du solveur Biq-Crunch. *18ème congrès de la Société Françcaise de Recherche Opérationnelle et d'Aide à la Décision (ROADEF'2017), Metz, France, 22-24 Février 2017*, pages 1–2, 2017.

[CO-5] M. Casazza, A. Ceselli et R. Wolfler Calvo. Inventory rebalancing in bike-sharing systems. *The 15th Cologne Twente Workshop on Graphs and Combinatorial Optimization (CTW 2017), COlogne, Germany, June 6 - 8, 2017*, pages 35–38, Köln, Allemagne, 2017.

[CO-4] A. Durand, Y. Ngoko et C. Cérin. Distributed and in-Situ Machine Learning for Smart-Homes and Buildings : Application to Alarm Sounds Detection. *ParLearning Workshop in conjunction with 2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPS Workshops 2017), Orlando / Buena Vista, USA, May 29, 2017*, pages 429–432, 2017.

[CO-3] F. Furini, T. Gschwind, S. Irnich et R. Wolfler Calvo. Social Network Analysis and Community Detection by Decomposing a Graph into Relaxed Cliques. *The 21st Combinatorial Optimization Workshop, Aussois, France, January 9-13, 2017*, 2017.

[CO-2] M. Garraffa, J. Malick et F. Roupin. On solving max-mean dispersion to optimality. *18ème congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF'2017), Metz, France, 22-24 Février, 2017*, pages 1–2, 2017.

[CO-1] S. Pan, R. Wolfler Calvo, L. Létocart, N. Touati et M. Akplogan. A column generation approach for a multi-activity tour scheduling problem. *18ème congrès de la Société Françcaise de Recherche Opérationnelle et d'Aide à la Décision (ROADEF'2017), Metz, France, 22-24 Février 2017*, pages 1–2, 2017.

Édition d'ouvrages collectifs

2022 [ED-6] C. Cérin, D. Qian, J.-L. Gaudiot, G. Tan et S. Zuckerman. *Proceedings of the 18th IFIP WG 10.3 International Conference on Network and Parallel Computing (NPC 2021), Paris, France, November 3-5, 2021*, volume 13152 de *Lecture Notes in Computer Science*. Springer, 2022.

2021 [ED-5] M. Di Francesco, E. Gorgone, R. Wolfler et P. Zuddas. *Special Issue on Freight Transportation and Logistics*, volume 78 de *Networks*. Wiley, 2021.

[ED-4] F. Furini, A. Lambert, L. Létocart, L. Liberti et E. Traversi. *16th Cologne-Twente Workshop on Graphs and Combinatorial Optimization (CTW 2018), Paris, France, June 18-20, 2018*, volume 296 de *Discrete Applied Mathematics*. Elsevier, 2021.

2020 [ED-3] J. André, E. Bourreau et R. Wolfler. *Special Section : ROADEF/EURO Challenge 2016 - Inventory Routing Problem*, volume 54 de *Transport Science*. INFORMS, 2020.

[ED-2] N. Brauner, C. Artigues, F. Clautiaux et P. Fouilhoux. *La Recherche Opérationnelle*, volume Hors-Série 75 de *Tangente*. 2020.

2018 [ED-1] A. M. J. Skulimowski, Z. Sheng, S. Khemiri-Kallel, C. Cérin et C.-H. Hsu. *The 5th International Conference on Internet of Vehicles. Technologies and Services Towards Smart City (IOV 2018), Paris, France, November 20-22, 2018*, volume 11253 de *Lecture Notes in Computer Science*. Springer, 2018.

Logiciels et Ressources (LO)

2021 [LO-4] C. Coti. SKaMPI 6 : an Extension of a Benchmark for MPI Libraries to support OpenSHMEM Libraries. https://github.com/coti/skampi, 2021.

2018 [LO-3] F. Furini, E. Traversi, P. Belotti, A. Frangioni, A. Gleixner, N. Gould, L. Liberti, A. Lodi, R. Misener, H. Mittelmann, N. Sahinidis, S. Vigerske et A. Wiegele. QPLIB : A Library of Quadratic Programming Instances. https://qplib.zib.de/, 2018.

[LO-2] J. Sliwak, M. Anjos, L. Létocart, E. Traversi et M. Ruiz. MathProgComplex : A Tool for for polynomial optimization problems with complex variables. https://github.com/JulieSliwak/mathprogcomplex.jl, 2018. MIT license.

2017 [LO-1] T. T. H. Nguyen, C. Rodríguez et C. Coti. Tool DPU. https://github.com/cesaro/dpu, 2017.

Thèses et Habilitations (TH)

Thèses

2022 [TH-8] E. Leclercq. Allocation de ressources par une approche hybride apprentissage automatique et optimisation pour l'hébergement et l'infogérance de services web. Thèse, Université Sorbonne Paris Nord, 2022.

[TH-7] J. J. Torres Figueroa. Algorithms for solving derivative-free optimization over mixed-integer domains. Thèse, Université Sorbonne Paris Nord, 2022.

2021 [TH-6] J. Sliwak. Résolution des SDP creux de grande taille issus de relaxations des ACOPF. Thèse, Université Sorbonne Paris Nord, 2021.

2019 [TH-5] E. Bettiol. Column generation methods for quadratic mixed binary programming. Thèse, Université Sorbonne Paris Nord, 2019.

[TH-4] E. Lancini. TDIness and Multicuts. Thèse, Université Sorbonne Paris Nord, 2019.

2018 [TH-3] S. Pan. Column generation methods for personnel scheduling problems. Thèse, Université Sorbonne Paris Cité, 2018.

2017 [TH-2] M. de Melo da Silva. Exact and Heuristic Algorithms for Container Rearrangement and Retrieval Problems at Container Terminals. Thèse, Université Sorbonne Paris Cité, 2017.

[TH-1] I. Kissami. High Performance Computational Fluid Dynamics on Clusters and Clouds : the ADAPT Experience. Thèse, Université Sorbonne Paris Cité, 2017.

Habilitations

2023 [HR-4] E. Traversi. Decompositions and Reformulations for Quadratic Problems. HdR, Université Sorbonne Paris Nord, 2023.

2021 [HR-3] R. Grappe. Polyhedra : Some Matricial Perspectives. HdR, Université Sorbonne Paris Nord, 2021.

2019 [HR-2] H. Abbes. Contributions à la Gestion Décentralisée des Ressources dans les Environnements de Grilles et Cloud. HdR, Université Sorbonne Paris Nord, 2019.

2018 [HR-1] C. D'Ambrosio. Solving well-structured MINLP problems. HdR, Université Sorbonne Paris Nord, 2018.

Rapports de contrats

2021 [RC-2] W. Saad et C. Cérin. Présentation de la solution Bridge-ITOPS, 2021.

2020 [RC-1] A. Amamou, M. Camey, C. Cérin, J. Rivalan et J. Sopena. Resources management for controlling dynamic loads in clouds environments. The Wolphin project experience. Rapport Technique hal-02481264, Université Sorbonne Paris Nord ; Sorbonne Université, 2020.

- **Publications antérieures à l'arrivée au LIPN**

Les références comportent « ↑ » et correspondent à des publications antérieures à l'arrivée au LIPN, non référencées par ailleurs.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2022 **[↑RI-10]** A.-E. Falq, P. Fouilhoux et S. Kedad-Sidhoum. Dominance inequalities for scheduling around an unrestrictive common due date. *European Journal of Operational Research*, 296(2) :453–464, 2022.

2021 **[↑RI-9]** P. Bendotti, P. Chrétienne, P. Fouilhoux et A. Pass-Lanneau. Dominance-based linear formulation for the Anchor-Robust Project Scheduling Problem. *European Journal of Operational Research*, 295(1) :22–33, 2021.

[↑RI-8] P. Bendotti, P. Fouilhoux et C. Rottner. Orbitopal fixing for the full (sub)-orbitope and application to the Unit Commitment Problem. *Mathematical Programming*, 186(1-2) :337–372, 2021.

[↑RI-7] A.-E. Falq, P. Fouilhoux et S. Kedad-Sidhoum. Mixed integer formulations using natural variables for single machine scheduling around a common due date. *Discrete Applied Mathematics*, 290 :36–59, 2021.

2020 **[↑RI-6]** P. Bendotti, P. Fouilhoux et C. Rottner. Symmetry-breaking inequalities for ILP with structured sub-symmetry. *Mathematical Programming*, 183(1-2) :61–103, 2020.

2019 **[↑RI-5]** P. Bendotti, P. Fouilhoux et C. Rottner. On the complexity of the Unit Commitment Problem. *Annals of Operations Research*, 274(1-2) :119–130, 2019.

2018 **[↑RI-4]** P. Bendotti, P. Fouilhoux et S. Kedad-Sidhoum. The Unit-capacity Constrained Permutation Problem. *European Journal of Operational Research*, 268(2) :463–472, 2018.

[↑RI-3] P. Bendotti, P. Fouilhoux et C. Rottner. The min-up/min-down unit commitment polytope. *Journal of Combinatorial Optimization*, 36(3) :1024–1058, 2018.

[↑RI-2] P. Fouilhoux, A. Mahjoub, A. Quilliot et H. Toussaint. Branch-and-Cut-and-Price algorithms for the preemptive RCPSP. *RAIRO - Operations Research*, 52(2) :513–528, 2018.

2017 **[↑RI-1]** P. Bendotti, P. Chrétienne, P. Fouilhoux et A. Quilliot. Anchored reactive and proactive solutions to the CPM-scheduling problem. *European Journal of Operational Research*, 261(1) :67–74, 2017.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2020 **[↑CI-2]** P. Bendotti, P. Chrétienne, P. Fouilhoux et A. Pass-Lanneau. Anchored Rescheduling Problems Under Generalized Precedence Constraints. *The 6th International Symposium on Combinatorial Optimization (ISCO 2020), Montreal, QC, Canada, May 4-6, 2020*, volume 12176 de *Lecture Notes in Computer Science*, pages 156–166, 2020.

2019 **[↑CI-1]** P. Bendotti, P. Fouilhoux et C. Rottner. Sub-Symmetry-Breaking Inequalities for ILP with Structured Symmetry. *20th Conference on Integer Programming and Combinatorial Optimization (IPCO 2019), Ann Arbor, MI, USA, May 22-24, 2019*, volume 11480 de *Lecture Notes in Computer Science*, pages 57–71, 2019.

Autres articles dans des conférences ou workshops à comité de lecture

2020 [↑CO-16] P. Bendotti, P. Chrétienne, P. Fouilhoux et A. Pass-Lanneau. Outils de résolution exacte pour l'ancrage de solutions en ordonnancement de projet. *21ème congrès annuel de la Société française de recherche opérationnelle et d'aide à la décision (ROADEF 2020), Montpellier, France, 19-21 Février 2020.*, Montpellier, France, 2020.

[↑CO-15] A.-E. Falq, P. Fouilhoux et S. Kedad-Sidhoum. Linear inequalities for neighborhood based dominance properties for the common due-date scheduling problem. *21ème congrès annuel de la Société française de recherche opérationnelle et d'aide à la décision (ROADEF 2020), Montpellier, France, 19-21 Février 2020*, pages 1–2, Montpellier, France, 2020.

2019 [↑CO-14] P. Bendotti, L. Brunod-Indrigo, P. Chrétienne, P. Fouilhoux et A. Pass-Lanneau. Anchor-Robust Solutions for the Resource-Constrained Project Scheduling Problem. *PGMO Days 2019, Palaiseau, France, 3-4 Décembre2019*, 2019.

[↑CO-13] P. Bendotti, P. Chrétienne, P. Fouilhoux et A. Pass-Lanneau. Stabilisation de solutions en optimisation combinatoire par des modèles d'ancrage ; application à l'arbre couvrant. *20ème congrès annuel de la Société française de recherche opérationnelle et d'aide à la décision (ROADEF 2019), Le Havre, France, 19-21 février 2019*, 2019.

[↑CO-12] P. Bendotti, P. Chrétienne, P. Fouilhoux et A. Pass-Lanneau. The Anchor-Robust Project Scheduling Problem. *14th Workshop on Models and Algorithms for Planning and Scheduling Problems (MAPSP), Renesse Netherlands, June 2-7, 2019*, 2019.

[↑CO-11] A.-E. Falq, P. Fouilhoux et S. Kedad-Sidhoum. Inégalités linéaires de dominance pour l'ordonnancement juste-à-temps avec date d'échéance commune non restrictive. *11ème Journées Polyèdres et Optimisation Combinatoire (JPOC11), Metz, France, 24-28 juin 2019*, Metz, France, 2019.

[↑CO-10] A.-E. Falq, P. Fouilhoux et S. Kedad-Sidhoum. MIP formulations for just-in-time scheduling around a common due-date. *14th Workshop on Models and Algorithms for Planning and Scheduling Problems (MAPSP 2019), Renesse, Netherlands, June 2-7, 2019*, 2019.

[↑CO-9] C. Rottner, P. Bendotti et P. Fouilhoux. Breaking structured symmetries and subsymmetries in Integer Linear Programming. *20ème congrès annuel de la société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF 2019), Le Havre, France, 19-21 Février 2019*, Le Havre, France, 2019.

2018 [↑CO-8] P. Bendotti, P. Chrétienne, P. Fouilhoux et A. Pass-Lanneau. Ancrage de solutions en ordonnancement PERT. *19ème congrès annuel de la Société française de recherche opérationnelle et d'aide à la décision (ROADEF 2018), Lorient, France, 21-23 février 2018*, 2018.

[↑CO-7] P. Bendotti, P. Chrétienne, P. Fouilhoux et A. Pass-Lanneau. Robust proactive maintenance planning with a deadline. *PGMO Days 2018, Palaiseau, France, 20-21 Novembre 2018*, 2018.

[↑CO-6] P. Bendotti, P. Fouilhoux et C. Rottner. Breaking full-orbitopal symmetries and subsymmetries. *The 23rd International Symposium on Mathematical Programming (ISMP 2018), Bordeaux, France, July 1-6, 2018*, 2018.

[↑CO-5] P. Bendotti, P. Fouilhoux et C. Rottner. Casser les symétries dans les PLNE à deux indices. *19ème congrès annuel de la Société française de recherche opérationnelle et d'aide à la décision (ROADEF 2019), Lorient, France, 21-23 février 2018*, 2018.

[↑CO-4] A.-E. Falq, P. Fouilhoux et S. Kedad-Sidhoum. Formulations PLNE et dominances pour l'ordonnancement juste-à-temps avec date d'échéance commune. *19ème congrès annuel de la Société française de recherche opérationnelle et d'aide à la décision (ROADEF 2018), Lorient, France 21-23 février 2018*, 2018.

2017 [↑CO-3] P. Bendotti, P. Fouilhoux et C. Rottner. Aspects polyédraux du Min-up/min-down Unit Commitment Problem. *10ème Journées Polyèdres et Optimisation Combinatoire (JPOC10), Villetaneuse, France, 14-16 Juin 2017*, Villetaneuse, France, 2017.

[↑CO-2] P. Bendotti, P. Fouilhoux et C. Rottner. Formulations PLNE et Branch & Cut pour le Min-up/Min-down Unit Commitment Problem. *18ème congrès annuel de la Société française de recherche opérationnelle et d'aide à la décision (ROADEF 2017), Metz, France, 22-24 février 2017*, Metz, France, 2017.

[↑CO-1] M. Casazza, P. Fouilhoux, M. Bouet et S. Secci. Securing Virtual Network Function Placement with High Availability Guarantees. *The 2017 IFIP Networking Conference (IFIP Networking) and Workshops, Stockholm, Sweden, June 12-16, 2017*, pages 1–9, Stockholm, Sweden, 2017.

3.3 Domaine 4. Inscription des activités de recherche dans la société

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique.

Comme mis en évidence à plusieurs reprises, l'équipe a la caractéristique très importante et très marquante d'avoir ses trois axes également développés, à savoir : une très forte composante théorique sur les fondamentaux, une composante algorithmique et un très bon niveau de transfert technologique. Sur ce dernier point, remarquons que les thèses CIFRE sont du même niveau que des thèses ayant reçu un financement ministériel et que le montant global des contrats est conséquent.

3.3.2 Référence 3. L'équipe partage ses connaissances avec le grand public et intervient dans des débats de société.

Nous n'avons pas de politique d'équipe sur ce point. Néanmoins, l'arrivée d'un nouveau professeur en 2020 nous permet d'afficher les deux activités listées dans la suite comme implication dans la diffusion de la science dans la société.

— deux articles de vulgarisation scientifique dans le numéro Bibliothèque Tangente consacré à la Recherche Opérationnelle, HS 75, octobre 2021
— interview dans le cadre du programme "Incroyables recherches" de la Cité des sciences et de l'industrie en décembre 2020 (vidéo en ligne sur Youtube, titre "Ma recherche : comment trouver la meilleure combinaison ?").

Chapitre 4
Trajectoire de l'équipe AOC

Sommaire

4.1 SWOT 162
4.2 Perspectives sur les prochaines années 163
4.2.1 Graphes box-parfaits 163
4.2.2 Algorithmes de décomposition et apprentissage 164
4.2.3 Optimiser dans l'incertain 164
4.2.4 Numérique responsable 165
4.2.5 Continuum Cloud Edge 166

Depuis sa création en 2010 et au cours des années qui se sont écoulées, l'équipe a recherché et maintenu, grâce aux trois axes de recherche et par leur interaction, un bon équilibre entre l'élaboration de fondements théoriques, leur implémentation algorithmique et leur mise en pratique pour la résolution de problèmes issus du monde industriel.

Au cours de la période précédente, l'équipe a constamment augmenté sa visibilité à tous les niveaux : local, national et international. Nous avons progressé et nous sommes internationalement reconnus dans des domaines variés à forte composante théorique, en particulier sur la caractérisation des systèmes Box-TDI et des polyèdres associés et sur les approches polyédrale.

L'équipe a gagné en visibilité dans le domaine du développement d'algorithmes de résolution de problèmes complexes d'optimisation combinatoire. Elle est plus particulierement renommée pour les approches basées sur la décomposition et la génération de colonnes, et sur les méthodes de résolution de problèmes non linéaires avec un focus sur problèmes quadratiques.

Enfin, l'équipe est reconnue pour ses compétences dans le développement sur des supports sophistiqués (Cloud, calcul haute performance).

Les résultats sont très visibles à travers l'augmentation du nombre de publications, leur positionnement, le nombre de thèses et de thèses CIFRE en particulier. Les objectifs fixés en 2017 ont été largement atteints.

4.1 SWOT

Trajectoire de l'équipe AOC

- **Forces et Atouts**
  - *richesse des thématiques* : ample spectre des thèmes de recherche et des compétences. Ces dernières permettent à l'équipe de valoriser ses résultats théoriques au travers de contrats, réalisations logicielles et collaborations aussi bien avec le milieu académique qu'avec des industriels ;
  - *rayonnement* : exceptionel niveau de publication, en hausse sur la période, dans des revues et des conférences internationales de premier plan ; responsabilités internationales, nationales, et locales ;
  - *collaborations* : collaborations au travers de structures (GdR RO, pôle MathStic), ciblées (CERMICS, LIX, INRIA, Bologne et Milan)
  - *attractivité* : candidatures internationales de niveau exceptionel sur les postes d'EC et de CR CNRS, séminaires de l'équipe et invitations ;
  - *insertion et contrats de recherche* : plusieurs financements, à la fois académiques (FUI, PGMO, BQR et PEPS) et industriels (PADAM, EDF, Orange, HUAWEI), opportunité pour nos docteurs de travailler dans les milieux académiques ou dans l'industrie ;
  - *valorisation de la recherche* : production de plusieurs logiciels *open source*, effort actuel de coordination et de standardisation des productions logicielles de l'équipe pour permettre une meilleure diffusion.

- **Faiblesses**
  - *formation à la recherche* : l'équipe n'est adossée à aucun master uniquement à une école d'ingénieurs ;
  - *projet européen* : manque de projets sur la période (mais dépôt prévu), support administratif insuffisant ;
  - *manque chronique de personnel permanent CNRS (CR-DR)* : amplifié par un manque de disponibilité pour la recherche (forte pression des responsabilités administratives et pédagogiques).

- **Opportunités**
  - *industrielles* : croissance des besoins des entreprises dans nos domaines d'expertises, et donc possibilités pérennes de trouver des financements ;
  - *nouveaux recrutements* : arrivée probable de deux nouveaux membres et développement de nouvelles activités de recherche associées à leur arrivée ;
  - *fédération de recherche MathSTIC* : elle contribue à initier et pérenniser des collaborations avec le L2TI et renforcer celles déjà existantes avec le LAGA et avec les autres équipes du LIPN (comme RCLN par exemple).

- **Menaces**
  - — *HDR* : les départs possibles de membres HDR pourraient déstabiliser l'équipe ;
  - — *recrutements de doctorants* : si le manque d'un Master adossé perdure, difficulté pour le recrutement des meilleurs doctorants ;

4.2 Perspectives sur les prochaines années

Tout d'abord nous souhaitons continuer à développer nos atouts et les domaines pour lesquels nous sommes visibles et reconnus comme par exemple les approches par décomposition appliquées à des problèmes linéaires et non linéaires en nombres entiers.

Nous souhaitons développer aussi de nouveaux thèmes de recherche comme décrit par la suite. Les lignes de recherche que nous souhaitons mettre en avant sont intéressantes parce qu'ancrées sur nos compétences, mais aussi relevant de nouveaux défis. Nous souhaitons par exemple aborder la relation entre Box-TDI et les graphes parfaits, améliorer les méthodes de décomposition avec l'introduction du machine learning, continuer à développer et étendre le concept de l'ancrage comme pont théorique et applicatif entre robustesse et réoptimisation, nous attaquer au Numérique responsable et enfin développer l'*Edge computing*.

Nous souhaitons de plus signaler que les projets et les perspectives pour les prochaines années sont fortement liés à l'évolution de l'équipe et aux arrivées et départs de ses membres. On peut constater que les départs de la période précédente ont affaibli l'axe 3 ; l'équipe accueillera probablement une chercheuse CNRS et une nouvelle MdC et perdra en revanche probablement deux HDRs.

Ces évolutions pourront nous amener à recentrer l'équipe et les prochains sujets de recherche sur les activités des Axes 1 et 2.

4.2.1 Graphes box-parfaits

Les graphes parfaits sont une classe de graphes où le comportement des stables (ensemble de sommets deux à deux non adjacents) est bien compris, et ont été extrêmement étudiés depuis les années 1950. En 1972, Lovász démontrait ce qui est maintenant connu sous le nom de théorème faible des graphes parfait : un graphe est parfait si et seulement si son complémentaire l'est aussi. Le point culminant sur ce sujet est le théorème fort des graphes parfaits, publié en 2006 et dû à Maria Chudnovsky, Neil Robertson, Paul Seymour, et Robin Thomas. Ce théorème fournit une caractérisation des graphes parfaits par sous-graphes interdits et résout une conjecture de 1961 due à Berge.

Les graphes box-parfaits sont une sous-classe des graphes parfaits où non seulement les stables se comportent bien, mais aussi les unions de stables. La question de leur caractérisation par sous-graphes interdits est un problème ouvert important du domaine, posé en 1982 par Cameron et Edmonds. Cameron a étudié les relations min-max dans les graphes box-parfaits en 1982. En 2018, Ding et al. ont caractérisé certaines sous-classes des graphes box-parfaits, mais la question de leur caractérisation en général reste largement ouverte.

Nous avons déjà mentionné que nos résultats nous ont permis de réfuter une conjecture de 2018 à leur propos. Il se trouve que, depuis, nos caractérisations des polyèdres box-TDI ont donné deux nouveaux résultats concernant les graphes box-parfaits. Le premier est une caractérisation polyédrale de ces graphes [1]. Le second commence par la remarque que le théorème faible des graphes parfaits ne s'étend pas aux graphes box-parfaits, et nous avons déterminé comment obtenir un théorème faible des graphes box-parfaits dans [2]. Ces résultats permettent deux nouvelles approches pour leur caractérisation : une approche polyédrale, et une approche purement théorie des graphes. La complémentarité de ces deux approches, et les compétences de l'équipe AOC sur

ces sujets, font de la caractérisation de graphes box-parfaits l'une des perspective importantes des années à venir pour l'axe 1. Mener à terme cette recherche résoudrait complètement un problème ouvert depuis plus de 40 ans.

Cet exemple est emblematique de nos souhaits sur ces thèmes : continuer l'étude des interations entre polyèdres et objets combinatoires de manière à développer de nouveaux théorèmes min-max forts.

4.2.2 Algorithmes de décomposition et apprentissage

Trajectoire de l'équipe AOC

L'équipe AOC est spécialisée depuis de nombreuses années dans le développement d'algorithmes de décomposition efficaces pour la résolution de problèmes linéaires en nombres entiers (génération de colonnes et/ou de contraintes, relaxation Lagrangienne, etc.) [RI-43, **RI-37**]. Ces algorithmes reviennent à résoudre de manière itérative une relaxation du problème original. Même si la convergence de tels algorithmes est assurée théoriquement, ces derniers peuvent en pratique mettre du temps à converger. Il faut alors développer différentes techniques d'accélération : méthodes de stabilisation, préprocessing, heuristiques de calcul de bornes, algorithmes de filtrage (des colonnes et/ou contraintes à ajouter), etc. Ces techniques d'accélération d'algorithmes exacts sont souvent heuristiques et il est difficile de définir des stratégies qui fonctionnent pour n'importe quelle instance. L'équipe AOC entend travailler sur l'utilisation de l'apprentissage automatique pour l'élaboration de ces techniques d'accélération.

Un des axes de recherche consiste à utiliser l'apprentissage automatique pour l'initialisation des algorithmes de décomposition. Les travaux actuels se focalisent sur l'initialisation du problème maître restreint dans la génération de colonnes (choix des colonnes initiales basé sur une prédiction de leur apparition dans la solution optimale), sur les valeurs initiales des multiplicateurs Lagrangiens dans un algorithme de faisceaux (variante de l'algorithme du sous-gradient), ou sur la décomposition des graphes en cliques pour l'optimisation de problèmes semi-définis creux (dans la continuité des travaux initiés dans [CI-56]).

Un deuxième axe de recherche consiste à utiliser l'apprentissage automatique pour guider la résolution des problèmes résolus à chaque itération. En effet, dans ces algorithmes itératifs, un même problème (tel que le problème de pricing en génération de colonnes par exemple) est résolu un grand nombre de fois. La variation du problème à chaque itération est faible et il est donc possible d'utiliser l'historique des précédentes résolutions pour accélérer la résolution courante. L'apprentissage automatique est une des pistes étudiées pour tirer partie de cet historique.

Ces travaux sont actuellement menés dans l'équipe AOC notamment à travers deux thèses. Un groupe de lecture bimensuel portant sur l'étude d'articles de recherche liant apprentissage automatique et programmation mathématique a par ailleurs été mis en place depuis fin 2022.

4.2.3 Optimiser dans l'incertain

Lors de la résolution d'un problème d'optimisation dans un contexte d'utilisation réelle, il est rare que les données de terrain restent identiques sur une longue période, ou alors qu'elles soient exactement les données fournies au logiciel de résolution. L'exemple le plus naturel est la planification de projet où la durée des tâches à positionner dans la durée du projet peuvent fluctuer au cours du temps ou ne pas être précisément connues à l'avance. Pour autant, il faut prévoir une solution initiale pour répondre à la demande du projet.

Une première question posée par cette résolution dans l'incertain est de pouvoir "ré-optimiser", c'est-à-dire fournir une solution réalisable ou optimale qui tient compte des variations sur les données. Dans [RI-63], l'incertitude sur les données du problème de localisation d'unité de production est abordée de manière classique pour montrer

le limite des approches classique. Dans [↑RI-1, ↑CI-2], il est montré que le problème de réoptimisation de la planification de projet est polynomiale, même en tenant compte des deadlines sur les dates des tâches à fixer.

Dans ce contexte d'incertitude, on emploie le mot *robustesse* pour désigner les approches permettant de fixer une solution initiale ayant la capacité de résister aux incertitudes sur les données. Pour cela, l'espace des incertitudes est supposé connu, par exemple avec un intervalle de variation sur chaque donnée du problème, ainsi qu'une mesure réaliste sur la variation possible de l'ensemble des données. Une approche robuste particulièrement intéressante est de considérer un budget maximal total des variations et de produire ainsi une solution capable de résister à toutes les variations dans l'espace de ce budget. Nous avons proposé le concept d'ancrage permettant de produire une solution initiale qui, au lieu d'être valide pour toute variation peut être corrigée en une solution optimale en opérant un nombre minimum de changements parmi les décisions initiales. Ce nouveau concept a été mise en œuvre dans [↑RI-9, **RI-65***] au travers d'une étude de complexité et de techniques algorithmiques basées sur la programmation linéaire qui se sont avérées très performantes par rapport à d'autres approches de robustesse.

Des premiers travaux sont en cours dans l'équipe AOC pour mettre en œuvre ces techniques pour d'autres problèmes de recherche opérationnelle. C'est par exemple particulièrement important dans le cas du le principe de slicing qui apparaît dans les protocoles 5G pour les réseaux Télécom. Le slicing consiste à découper et réserver les ressources d'un réseau d'un opérateur entre différents clients en assurant une qualité importante en termes de capacité et de sécurité. Il est important de prévoir à l'avance des répartitions robustes pour résister aux variations de demande tout en assurant une stabilité pour les clients prioritaires (comme par exemple une communication sécurisée entre deux salles d'opérations dans deux hôpitaux distants). Des travaux sont en cours en commençant par l'étude de l'ancrage des multiflots dans un réseau.

Une autre application où l'ancrage est également pertinente est de construire des grilles d'horaires robustes pour les transports publics, en particuliers les bus de villes qui sont soumis à différents aléas concernant le personnel ou la circulation. Maximiser l'ancrage de certains horaires à des stations peut améliorer considérablement la qualité du service public. Une première étude commence dans le cadre d'une collaboration.

Au delà du coté applicatif, l'ancrage s'avère être un concept assez naturel en optimisation combinatoire et les applications citées ci-dessus permettent d'aborder ce concept au travers de différents cadres exacts de l'optimisation multi-niveau et de la programmation mathématique.

4.2.4 Numérique responsable

En 15 ans, le trafic Internet été multiplié par environ 500, de 2002 à 2017[1]. Les émissions de CO_2 associées ont été évaluées à 762 millions de tonnes en 2018[2]. Pour éviter les conséquences dramatiques du changement climatique, tous les secteurs de l'économie mondiale, y compris les technologies de l'information et de la communication (TIC), doivent maintenir leurs émissions de gaz à effet de serre (GES) en conformité avec l'accord de Paris. La littérature contient des estimations des émissions de GES des TIC, qui situent la part des TIC dans les émissions mondiales de GES entre 1,8% et 2,8% en 2021, soit une augmentation de 2,5% par rapport à l'année précédente et 2,8% en 2022[3].

Le Cloud Computing, étudié dans l'équipe, qui est l'accès à des ressources distantes via Internet, a un impact important sur le réchauffement climatique. Les centres de données sont par exemple tout particulièrement pointés du doigt même si des efforts importants ont été consentis ces dernières années. Nous pouvons citer la mise en place

1. https://global-internet-map-2022.telegeography.com/
2. https://theshiftproject.org/wp-content/uploads/2019/07/2019-01.pdf
3. https://www.youtube.com/watch?v=UM3EW01_PUY&t=143s&ab_channel=Jean-MarcJancovici

de rack réfrigérés plutôt que de réfrigérer la salle machine, l'utilisation de processeurs de moins en moins énergivores et la consolidation des serveurs pour éviter de laisser allumé inutilement des équipements.

Nous allons poursuivre le travail entamé en 2022. Ce travail étudie l'impact environnemental des requêtes HTTP, et la partie expérimentale est disponible en ligne [4]. L'idée maîtresse est de plonger le problème dans une problématique de science des données, plutôt que de continuer avec des approches analytiques, comme pour l'EcoIndex [5]. En effet, la science des données se prête plus facilement, par exemple, au traitement d'un grand nombre d'attributs dans les modèles.

4.2.5 Continuum Cloud Edge

Trajectoire de l'équipe AOC

Les architectures Cloud et Edge sont les architectures dominantes de ces dernières années. Le Edge peut être vu comme un prolongement du Cloud où certains calculs sont effectués en bordure de réseau, par exemple directement ou au plus près des capteurs produisant des données. On parle alors de continuum et nous proposons de continuer à investir des sujets suivants.

Premièrement, pour le Cloud, nos résultats antérieurs permettent de déployer à la demande tout un écosystème de cluster de calcul, c'est à dire un ordonnanceur, des noeuds de calcul et des applications, au dessus d'un orchestrateur de Cloud, pour nous, Kubernetes. Nous allons pouvoir aider à travailler à la reproductibilité des grandes expériences scientifiques, celles qui tournent sur nos clusters nationaux au CINES par exemple. En effet, dans 15 ans, comment pourrions-nous rejouer une expérience réalisée en 2023 ? Précisément en pouvant redéployer le cluster du CINES disponible en 2023.

Deuxièmement, pour l'Edge, nous proposons deux directions. La première vise à continuer à travailler sur des approches frugales permettant de calculer et d'apprendre sur du matériel disponible en bordure de réseau. Nous avons par exemple mis en place un schéma de clustering en ligne adapté à une machine de type Raspberry. La seconde se placera dans le cadre du projet FogSLA (AMI Cloud Computing, BPI France) ou nous avons commencé à travailler sur de l'ordonnancement multi-critères pour les orchestrateurs de l'Edge (K3s et MicroK8s).

Par ailleurs, dans l'environnement de la plateforme FogSLA, plusieurs micro-services coexisteront dans des environnements hétérogènes. Par conséquent, une orchestration des services de bout en bout est nécessaire pour maintenir un contrôle continu des services déployés en fonction de la qualité de service demandée par les cas d'utilisation. Ce point nous amène à la question de savoir comment exprimer dynamiquement les demandes de qualité de service des cas d'utilisation dans un accord de niveau de service (SLA), ce qui pourrait alors être utilisé pour définir des objectifs d'orchestration de bout en bout (SLO). Pour les questions d'orchestration, lorsqu'un objectif est fixé, il existe plusieurs façons de l'atteindre en déployant plusieurs services et, en général, chaque étape introduit des propriétés qui peuvent être plus ou moins attrayantes pour l'objectif. L'une des difficultés réside dans le fait que ces propriétés sont découvertes à la volée et doivent être appliquées dynamiquement. L'examen des questions de SLA pour la 5G du point de vue des SLO et de la dynamicité de l'orchestration et du cycle de vie des services reste une question ouverte. Dans la solution FogSLA, le déploiement de micro-services sera basé sur des technologies de conteneurs au niveau du "fog".

Notons que cette dernière partie est conjointement étudiée avec le laboratoire L2TI de l'Université Sorbonne Paris Nord.

Références complémentaires

[1] P. Chervet and R. Grappe. The essence of a box-totally dual integral polytope. *Submitted.*

4. https://github.com/christophe-cerin/Ecoindex-Revisited
5. https://github.com/cnumr/GreenIT-Analysis

[2] P. Chervet and R. Grappe. A weak box-perfect graph theorem. *Submitted*.

Partie IV

CALIN

Combinatoire, Algorithmique et Interactions

Sommaire

Informations générales sur l'équipe CALIN **175**
1.1 Membres de l'équipe CALIN 176
1.2 Les thématiques scientifiques et leurs enjeux 178
1.3 Profil d'activités liées à la recherche 186
1.4 Prise en compte des recommandations du précédent rapport 186

Introduction du portfolio **189**

Auto-évaluation de l'équipe CALIN **193**
3.1 Domaine 2. Attractivité 194
3.2 Domaine 3. Production scientifique 199
3.3 Domaine 4. Inscription des activités de recherche dans la société 214

Trajectoire de l'équipe CALIN **217**
4.1 Atouts, faiblesses, opportunités, menaces 217
4.2 Positionnement par rapport au projet proposé en 2017 219
4.3 Projet 219

Overview

In 2010, the LIPN decided to create a combinatorics team in an effort to bring together people working on various aspects of combinatorics within different communities, including computer scientists, mathematicians, and physicists. This combination of different skills and backgrounds allows to tackle difficult problems arising from each of these communities using an extended set of techniques. This philosophy is reflected in the team's acronym, CALIN, that stands for *Combinatorics, ALgorithmics, and their INteractions*, and is French for a *big hug*.

The research is organized into two interacting sub-groups. One of them focuses on the *theory of algorithms*, and the rigorous *analysis of these algorithms*. This task is conducted by statistically understanding random structures and their asymptotics. This sub-group is primarily interested in the combinatorial structures arising in computer science: words, and their link to algebra; graphs and trees, appearing for instance as search trees, networks, or transition structures in automata; random walks and other dynamics on such graphs; partitions and permutations, to name only a few. Their methods are bijective, probabilistic, or arising from an analytic approach through generating functions and singularity analysis.

The other sub-group is interested in the *interface between combinatorics, geometry, and physics*. Many combinatorial problems arise from physics often via geometric models. Among such models, one finds for instance tilings, and decompositions of surfaces (or higher-dimensional manifolds) such as triangulations or cell complexes. The study of these and associated objects conducted by this sub-group from a combinatorial point of view, also has strong links with more classical computer science problems, as for instance the triangulations of topological bordered surfaces that, in their simplest manifestation, give rise to the associahedron, a polytope with physical roots (via string theory), but whose face complex can also be thought of in terms of binary search trees.

Our research is supported by several ANR projects, and is based on a large number of fruitful scientific collaborations, with both French and foreign research groups. The team is involved on a regular basis in the organisation of wide-ranging scientific events, including for instance the GT ALEA and its European counterpart, the *Alea in Europe* network, the conferences LAGOS and AofA, the *Séminaire Lotharingien de Combinatoire*, and two regional seminars that take place at the Institut Henri Poincaré: the *Philippe Flajolet combinatorics seminar* and the *séminaire francilien de géométrie algorithmique et combinatoire*.

Our scientific results are published in journals among the very best in computer science (ACM Transactions on Algorithms, Random Structures and Algorithms, Algorithmica, ...), in mathematics (Journal of the European Mathematical Society, Advances in Mathematics, Journal of Combinatorial Theory A, Annals of Probability, Israel Journal of Mathematics, Journal d'Analyse Mathématique, ...), and in physics (Journal of High Energy Physics, Physical Review D, European Physical Journal C, ...). They are also presented in conferences with the highest international standards as for instance FPSAC.

Besides scientific publications, our team is also involved in the Sage mathematics software thanks to T. Monteil.

Chapitre 1

Informations générales sur l'équipe CALIN

Sommaire

1.1 Membres de l'équipe CALIN 176
1.1.1 Membres permanents 176
1.1.2 Membres associés 177
1.1.3 Doctorants au 1er janvier 2023 177
1.1.4 Postdoctorants au 1er janvier 2023 177
1.1.5 Autres personnels temporaires du 01/01/2017 au 31/12/2022 178
1.1.6 Évolution de l'équipe (permanents) 178
1.2 Les thématiques scientifiques et leurs enjeux 178
1.2.1 Axe analyse d'algorithmes et de structures combinatoires 179
1.2.2 Axe interactions entre combinatoire, physique et géométrie 182
1.2.3 Collaborations avec d'autres équipes du LIPN 184
1.3 Profil d'activités liées à la recherche 186
1.4 Prise en compte des recommandations du précédent rapport . . . 186

Nom de l'équipe : Combinatoire, Algorithmique et Interactions (CALIN)

Responsable de l'équipe : Lionel Pournin

1.1 Membres de l'équipe CALIN

Effectifs au 1^{er} janvier de chaque année.

Corps	2017	2018	2019	2020	2021	2022
DR	0	0	0	0	0	0
CR	4	4	4	4	4	4
PU	3	3	3	3	3	4
MCF	6	6	5	6	6	6
Post-docs et ATER	3			5	3	
Doctorants	5	6	9	7	9	7
Total	*21*	*19*	*21*	*25*	*25*	*21*

1.1.1 Membres permanents

(membres ou, *en italique, anciens membres*, et, **en gras, responsable de l'équipe** au 31/12/2022)

Informations générales sur l'équipe CALIN

	Nom	*Prénom*	*Situation*	*Institution*
	Bacher	Axel	MCF	IUTV
	Banderier	Cyril	CR section 6	CNRS
	Bassino	Frédérique	PU	IUTV
	Ben Geloun	Joseph	MCF	IG
	Bodini	Olivier	PU	IG
	Bonzom	Valentin	MCF	IUTSD
	David	Julien	MCF	IG
	Duchamp	Gérard	PU émérite	IG
	Fernique	Thomas	CR section 6	CNRS
	Monteil	Thierry	CR section 41	CNRS
	Mustafa	Nabil	PU	IG
	Poinsot	*Laurent*	*MCF*	*IG*
	Pournin	**Lionel**	**PU**	**IG**
	Sportiello	Andrea	CR section 6	CNRS

	Tollu	Christophe	MCF	IG
	Valencia-Pabon	*Mario*	*MCF*	*IG*

1.1.2 Membres associés

	Nom	*Prénom*	*Situation*	*Institution*
	Akhavi	Ali	MCF	Université de Caen
	Hoang Ngoc Minh	Vincel	PU	Université de Lille

1.1.3 Doctorants au 1er janvier 2023

	Nom	*Prénom*	*Situation*	*Institution*
	Buffière	Théophile	CD	USPN MathSTIC
	Chinaud-Chaux	Clémence	CD	80Prime
	Louvet	Alexandre	CD	ANR ADDS
	Matrosova	Kristina	CD	ANR Records
	Nguyen Dinh	Vu	CD	USPN
	Pchelina	Daria	CDSN	Ens Paris 13
	Porrier	Carole	Allocataire	UQAM (Canada)

1.1.4 Postdoctorants au 1er janvier 2023

	Nom	*Prénom*	*Situation*
	Nurligareev	Khaydar	ATER USPN
	Pépin	Martin	Post-doctorant ANR LambdaComb
	Singh	Alexandros	ATER USPN

1.1.5 Autres personnels temporaires du 01/01/2017 au 31/12/2022

Nom	*Prénom*	*Situation*	*Début*	*Fin*	*Financement*
Bean	Christian	post-doctorant	01/12/2019	31/07/2020	FR MathSTIC
Courtiel	Julien	post-doctorant	01/09/2016	31/08/2017	ANR MetaConc
De Mourgues	Quentin	doctorant	01/10/2013	05/12/2017	ENS Paris
Dovgal	Serguey	doctorant	01/10/2016	30/09/2019	ANR MetaConc
		ATER	01/10/2019	31/08/2020	USPN
Dub	Nicolas	doctorant	01/12/2016	28/06/2021	USPN
Galanov	Ilya	doctorant	01/11/2016	17/12/2019	USPN
		ATER	18/12/2019	31/08/2021	USPN
Lutfalla	Victor	doctorant	01/10/2018	31/08/2021	ENS Lyon
Naima	Medhi	doctorant	01/10/2017	18/12/2020	USPN
		ATER	19/12/2020	31/08/2021	USPN
Nurligareev	Khaydar	doctorant	01/10/2018	20/10/2022	USPN
Pardal	Nina	doctorante	20/12/2018	30/03/2020	Univ. Buenos Aires
Rakotoarimalala	Tsinjo Tony	doctorant	01/10/2015	08/07/2019	USPN
		ATER	09/07/2019	31/08/2020	USPN
Rakotonarivo	Rado	doctorant	01/11/2016	18/09/2019	USPN
		ATER	19/109/2019	31/08/2021	USPN
Singh	Alexandros	doctorant	01/10/2019	15/11/2022	USPN
Ugolnikova	Alexandra	Post-doctorante	01/01/2017	31/03/2017	ANR QuasiCool
Wallner	Michaël	Post-doctorant	01/09/2017	31/12/2017	FR MathSTIC

1.1.6 Évolution de l'équipe (permanents)

Informations générales sur l'équipe CALIN

Créée en mai 2010, l'équipe CALIN (Combinatoire, ALgorithmique et INteractions), réunit les combinatoristes auparavant répartis entre deux équipes du LIPN. Il s'agit de la concrétisation d'un projet scientifique autour de la combinatoire sous toutes ses formes, de ses applications à l'analyse d'algorithmes et de structures de données et de ses interactions avec la physique.

Après avoir connu un fort renouvellement dans la période précédente, l'équipe s'est stabilisée. Depuis 2017, il y a eu le recrutement d'un PU et le départ en mutation d'un MCF. Le nouvelles mobilités prévues (promotion, mutation et disponibilité, possible recrutement) devraient faire sensiblement évoluer le profil de l'équipe.

Il est à noter que Laurent Poinsot est en détachement à l'école de l'Armée de l'Air depuis 2014 et Julien David est en échange de service avec Ali Akhavi (Université de Caen) depuis 2021.

Départs

- **2018** J. David, MCF IG (en disponibilité chez Qwant)
- **2022** M. Valencia Pabon, MCF IG (mutation à l'Université de Lorraine)

Arrivées

- **2019** J. David, MCF IG (retour de chez Qwant)
- **2021** N. Mustafa, PU IG (Professeur à l'ESIEE)

1.2 Les thématiques scientifiques et leurs enjeux

Le LIPN a créé en mai 2010 une équipe dédiée à la combinatoire, baptisée CALIN (Combinatoire, ALgorithmique et INteractions), réunissant les combinatoriciens auparavant répartis entre deux autres équipes. Cette équipe réunit des chercheurs ayant des compétences dans les divers domaines de la combinatoire (analytique, bijective, géométrique et algébrique) avec un intérêt pour la complexité des algorithmes dont ils cherchent à déterminer le comportement en moyenne ou au pire, ainsi que pour la physique dont ils utilisent des méthodes ou étudient certains problèmes.

Les recherches de l'équipe s'organisent autour de deux axes : l'un abordant l'analyse d'algorithmes et de structures combinatoires, l'autre tourné vers les interactions entre

combinatoire, géométrie et physique. Une journée hebdomadaire de séminaires largement suivie favorise les discussions scientifiques entre membres de léquipe. Nos activités s'inscrivent dans le cadre de multiples collaborations qui jouent un rôle essentiel au niveau local avec la fédération de rechercheMathSTIC (CNRS FR 3734), et l'organisation de deux séminaires régionaux (le séminaire de combinatoire Philippe Flajolet et le séminaire Francilien de Géométrie Algorithmique et Combinatoire), au sein des GT Aléa et CombAlg du GDR Informatique Mathématique (au niveau national), de la communauté du Séminaire Lotharingien de Combinatoire (au niveau européen), des communautés Analysis of Algorithms et FPSAC (au niveau international).

Nos résultats ont fait l'objet de nombreuses publications, dont une partie significative d'articles longs dépassant les 40 pages, dans les meilleures revues de *combinatoire* (Journal of Combinatorial Theory A, Combinatorial Theory, European Journal of Combinatorics, Combinatorics Probability and Computing), d'*algorithmique* (ACM Transactions on Algorithms, Random Structures and Algorithms, Algorithmica), de *physique* (Journal of High Energy Physics, Physical Review D, European Physical Journal C), de *mathématiques généralistes* (Journal of the European Mathematical Society, Advances in Mathematics, Israel Journal of Mathematics, Proceedings of the American Mathematical Society) ou *spécialisées* (Annals of Probability, Journal d'Analyse Mathématique, Discrete & Computational Geometry), ainsi que de communications dans la *meilleure conférence internationale* du domaine (FPSAC).

1.2.1 Axe analyse d'algorithmes et de structures combinatoires

Informations générales sur l'équipe CALIN

Nous cherchons à étudier de manière fine des structures combinatoires fondamentales de l'informatique (mots, automates, marches, permutations, arbres, graphes) ainsi que les algorithmes opérant sur ces structures. Une partie des méthodes que nous utilisons sont bijectives, probabilistes ou issues de la combinatoire analytique. Ce dernier domaine, dont le développement doit beaucoup à Philippe Flajolet [25], recourt à des techniques d'analyse complexe. Nous étudions aussi la complexité d'algorithmes géométriques (comme l'algorithme du simplexe) ou dont l'analyse s'appuie sur des techniques et des objets géométriques ou topologiques (rééquilibrage d'arbres par des opérations de rotation).

- **Génération aléatoire**

La génération aléatoire joue un rôle important dans les domaines manipulant d'énormes volumes de données. Elle permet de comparer des modèles, d'évaluer leur adéquation avec les données réelles, mais aussi de tester des programmes.

Les générateurs de Boltzmann [18] sont une méthode simple et automatisable d'engendrer des structures combinatoires, dès lors que l'on peut calculer la série génératrice de ces structures. En développer des versions multivariées fait partie des principaux défis. Ainsi nous avons conçu [RI-98] et implémenté [LO-1] une solution efficace au problème difficile du calibrage des générateurs de Boltzmann multivariés. Cela permet, par exemple, la génération aléatoire rapide de λ-termes de grande taille, ouvrant ainsi la porte à des études empiriques de paramètres non-triviaux.

Les générateurs de Boltzmann, s'ils sont extrêmement efficaces pour la génération aléatoire approchée, ont (dans la plupart des cas) une complexité d'ordre $n^{\frac{3}{2}}$ ou n^2, là où la borne inférieure de Shannon ne serait que linéaire. Dans ce contexte, des algorithmes efficaces voire optimaux en nombre de *bits aléatoires* ont été développés. Nous [**RI-22**] avons ainsi optimisé, en gagnant un facteur $\log n$, la génération aléatoire de permutations. Leur algorithme `MergeShuffle`[1] est actuellement le plus rapide pour engendrer uniformément des permutations aléatoires [CO-10]. En lien avec des applications en

1. https://github.com/axel-bacher/mergeshuffle

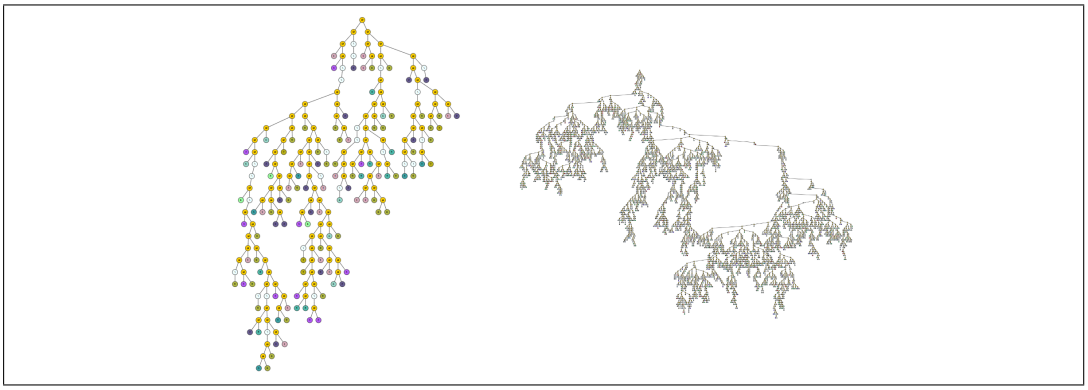


Figure 1.1: Des λ-termes aléatoires de grande taille ($n \approx 500$ à gauche et $n \approx 10000$ à droite)

théorie des systèmes concurrents, où les contraintes d'ordre du poset capturent naturellement les contraintes d'évolution temporelle, nous [CI-4] avons également étudié la génération aléatoire d'extensions linéaires de posets série-parallèle, et notamment leur coût en termes de bits aléatoires requis.

Nous [**RI-51**, CO-9, **CI-26**] avons aussi développé une nouvelle méthode de génération aléatoire, *la méthode des densités*, qui repose sur un plongement dans le continu de structures discrètes, aboutissant alors à des récurrences multivariées encodant des volumes de polytopes. Cette méthode permet d'engendrer aléatoirement des permutations contraintes et diverses familles de posets et d'obtenir des formules efficaces (de coût polynomial) pour calculer différentes structures qui n'avaient jusqu'ici que des énumérations de coût exponentiel. Ces formules font intervenir des récurrences d'un nouveau type : $p_{n+1}(x) = \int_{y \in P_n} Q(x,y) p_n(y) dy$, et des intégrales multiples sur des polytopes P_n. L'énumération est alors donnée par la suite $a_n = n! \int_0^1 p_n(x) dx$.

Informations générales sur l'équipe CALIN

3	5	8	9	11	13	14	15
2	10	4	7	1	16	6	12

Figure 1.2: Combien y a-t-il de telles structures de taille $2 \times n$ (les étiquettes sont croissantes par lignes et par colonnes, sauf éventuellement de part et d'autre d'un mur rouge). La méthode des densités permet la génération et le dénombrement de tels posets.

Notons pour finir que certaines de nos méthodes de générations aléatoires ont été utilisées pour engendrer des œuvres artistiques (voir la Section 3.3.1).

- **Analyse de structures combinatoires contraintes**

L'analyse de structures combinatoires simples vérifiant certaines contraintes, telles que vivre dans un réseau sous-jacent donné d'un certain réseau ou éviter certains motifs est une question naturelle en combinatoire. Ces structures contraintes apparaissent en effet dans de nombreuses applications, comme l'analyse d'algorithmes (pour modéliser les entrées ou les exécutions d'un programme), la bio-informatique, la physique théorique, etc.

Les λ-termes jouent un rôle fondamental en programmation fonctionnelle, leur étude statistique est destinée à la vérification de programmes et de compilateurs. La structure combinatoire particulière des λ-termes leur confère des propriétés jusqu'à présent peu rencontrées dans les structures plus classiques ; nous en avons fait une étude approfondie [RI-49]. De plus, des liens forts ont été trouvé entre cartes combinatoires et certains lambda-termes [CI-8] ouvrant la voix à une "géométrie" du λ-calcul. Plus récemment,

nous [4] avons réussi la première étude quantitative sur la dynamique du calcul sur les λ-termes, en bornant la profondeur moyenne de la β-réduction sur les λ-termes linéaires. Cela constitue une avancée majeure pour la compréhension de la dynamique du λ-calcul.

Les permutations *contraintes* à éviter des motifs modélisent des phénomènes biologiques ainsi que divers tris par piles et sont à l'origine de nombreuses études combinatoires et algorithmiques [6, 28, 35], ainsi que, plus récemment, probabilistes. Nous avons ainsi étudié la forme limite des permutations aléatoires uniformes dans des classes ayant une spécification combinatoire finie pour la décomposition par substitution, analogue de la décomposition modulaire des graphes [**RI-38**, **RI-65**, **RI-100**]. Nous nous sommes sont intéressés à leur comportement limite au sens des permutons, *i.e.* des mesures de probabilité sur le carré unitaire ayant des marginales uniformes, qui peuvent être vus comme des diagrammes de permutation normalisés. Le permuton limite est, avec une certaine universalité, le permuton séparable brownien introduit dans [**RI-38**] (voir portfolio 1).

- **Complexité des algorithmes de pivot**

En optimisation linéaire, une des méthodes les plus populaires est l'algorithme du simplexe qui consiste à explorer le graphe d'un polyèdre en utilisant une *règle de pivot*. Les sommets des polyèdres traités en pratique ont des coordonnées rationnelles et à homothétie près, un tel polyèdre est en nombres entiers (c'est-à-dire que les coordonnées de ses sommets sont des entiers). Afin de pouvoir étudier le neuvième problème de Smale (qui demande si l'optimisation linéaire peut être résolue en temps fortement polynomial), nous avons considéré le diamètre maximal $\delta(d,k)$ que peut avoir un polytope en nombre entiers contenu dans l'hypercube $[0,k]^d$ (ainsi le paramètre k quantifie la taille de l'instance traitée, dans l'esprit du problème de Smale). Nous avons obtenu des bornes sur $\delta(d,k)$ [RI-27] et plusieurs résultats sur des questions voisines provenant de la physique théorique et de l'optimisation [RI-58, RI-74, RI-86, RI-113*]. Nous avons aussi établi une borne inférieure asymptotique lorsque k tend vers l'infini et que d est fixé [**RI-57**]. Cette borne inférieure, obtenue en calculant le diamètre maximal $\delta_z(d,k)$ d'un *zonotope* en nombre entiers contenu dans l'hypercube $[0,k]^d$ est conjecturalement exacte [14]. à la suite de ce résultat, nous avons obtenu une formule explicite exacte pour $\delta_z(d,k)$ pour tout d et k, qui donne conjecturalement la valeur de $\delta(d,k)$ [RI-85] (voir portfolio 5). Si cette conjecture est vraie, cela fournit une valeur exacte de la complexité de l'algorithme du simplexe dans le cas le pire si la règle de pivot produit des chemins géodésiques.

- **Complexité et géométrie des graphes**

La complexité de nombreux algorithmes peut être étudiée dans le cadre de la géométrie des graphes (distances, convexité). Différents types de convexité peuvent être étudiées dans les graphes en remplaçant le caractère géodésique des chemins considérés dans la convexité usuelle par une autre propriété. Par exemple, si l'on considère, au lieu de chemins géodésiques, des chemins induits, on obtient la *convexité monophonique* [16] et si on considère uniquement des chemins de longueur 3 [13] on obtient la P_3*-convexité*. Des exemples d'applications des différents types de convexité dans les graphes se trouvent dans [17] et plus récemment dans [RI-24, 33]. Dans [RI-61, RI-71], nous avons obtenu des résultats prometteurs concernant la P_3-convexité des graphes de Kneser et des graphes de Hamming, respectivement. Nous avons aussi étudié les graphes de flips des surfaces topologiques de type fini (en lien avec la géométrie du *mapping class group* de ces surfaces [10]) et des arbres d'élimination du point de vue de la convexité et plus généralement de leurs propriétés métriques, obtenant des bornes sur les diamètre [**RI-2**, **RI-4**, RI-24, RI-25, RI-90, CI-25] et des résultats de complexité algorithmique pour le calcul de distances [12, 33]. Un de nos résultats particulièrement

marquants est la preuve que les triangulations de dimensions 3 introduites par Sleator, Tarjan et Thurston [24] pour modéliser les chemins géodésiques dans le graphe des flips d'un polygone sont *flag* [33]. Nous avons en outre obtenu des résultats sur certains graphes de flips qui modélisent la construction par opérations locales des polytopes en nombres entiers [**RI-59**].

1.2.2 Axe interactions entre combinatoire, physique et géométrie

Une thématique de l'axe interactions entre combinatoire, physique et géométrie de l'équipe CALIN est l'étude structurelle et énumérative de variétés discrétisées de dimensions 2, 3 et 4. En dimension 2, on parle de *cartes combinatoires*, un domaine classique de la combinatoire, initié par Tutte et fortement développé ensuite au sein de la communauté informaticienne (notamment par Cori via des décompositions à base de grammaires algébriques) et de la communauté probabiliste (carte brownienne, etc.). Nous étudions aussi les liens avec la géométrie convexe ou avec la combinatoire des polyèdres.

Informations générales sur l'équipe CALIN

- **Variétés discrétisées de dimension 2**

Un défi de la combinatoire des cartes est la mise en place de structures universelles reliant les propriétés de cartes de genres différents. S'inspirant des travaux en physique théoriques des années 80, Goulden et Jackson [14] ont prouvé que certaines familles de cartes satisfont de remarquables *formules de récurrence sur le nombre de cartes en fonction de la taille et du genre*. Bien que leur travaux aient largement essaimé, leur universalité n'est pas bien comprise. C'est dans cette optique que nous avons étendu leur travaux aux cartes non-orientées dans [**RI-93**, **RI-94**] (voir portfolio 3). Dans [9], nous avons prouvé une conjecture qui supposait qu'une autre structure universelle bien connue, la *récurrence topologique* [20], est satisfaite par la plus large classe de cartes, celles des constellations de Lando–Zvonkin [30]. Notre preprint est paru le même jour que celui d'une autre équipe internationale [11] !

- **Variétés discrétisées de dimensions supérieure**

En dimensions 3 et 4, la topologie et la combinatoire sont notoirement plus difficiles à contrôler [19]. Par exemple, on ne sait pas si le nombre de triangulations de la 3-sphère est exponentiellement borné par le nombre de tétraèdres. Nous avons beaucoup développé la technologie des graphes réguliers aux arêtes proprement coloriées comme une représentation graphique de (pseudo-)variétés triangulées [RI-11, RI-12, 8]. Cette approche a permis des avancées dans un domaine où elles sont rares. Nous avons prouvé dans [7] que les triangulations obtenues par collage de 3-boules qui ont la propriété de maximiser le nombre d'arêtes sont topologiquement des 3-sphères mais possèdent la structure combinatoire d'arbres ! Ce résultat a une portée universelle, puisqu'il tient pour tout choix de 3-boules discrétisées. Nous avons aussi développé une forme de récurrence topologique pour ces triangulations [RI-29, 10].

L'énumération de ces triangulations grâce à des méthodes de combinatoire algébrique [**RI-15**] a permis de mettre à jour une *construction et une interprétation combinatoire des coefficients de Kronecker* [**RI-118**], une question combinatoire de Murnaghan en théorie des représentations, ouverte depuis 1938 [32]. Nous avons même construit des opérateurs dont les noyaux ont pour dimension des valeurs des coefficients de Kronecker (voir portfolio 2).

En 2002, Bollobás et Riordan [5] avaient étendu le polynôme de Tutte, un polynôme défini sur les graphes et satisfaisant de remarquables identités sous les transformations de suppressions et contractions d'arêtes, à un polynôme défini sur des cartes et satisfaisant des identités similaires. Nous avons à notre tour étendu ce polynôme aux

(pseudo-)variétés de dimension 3 [RI-103], ce qui nécessite de prendre en compte plus de paramètres tout en maintenant les identités de suppressions et contractions.

L'équipe est aussi active dans l'expertise auprès de la physique théorique. Elle a fourni des classifications de "graphes de Feynman" indiquant aux physiciens comment ordonner les termes de séries génératrices pour certaines théories quantiques des champs associées aux trous noirs [RI-47, RI-48, RI-91, RI-115]. Associées aux triangulations de dimension 3 et 4, elle a aussi développé des théories quantiques des champs satisfaisant des propriétés profondes et rares comme la liberté asymptotique, suggérant ainsi des *transitions de phase vers de nouvelles géométries* [**RI-36**].

- **Pavages apériodiques et empilements de sphères**

Les empilements de sphères et les pavages apparaissent en physique, en mathématiques, mais aussi dans le cadre de la calculabilité en informatique comme généralisation des mots en dimension supérieure. Nous voulons comprendre comment des contraintes locales structurelles telles que l'imbrication de tuiles d'un pavages ou de sphères dans un empilement se propage par un phénomène de transfert d'information et peut engendrer des structures globales complexes. Une série de conditions pour qu'un pavage apériodique par coupe et projection puisse être engendré de cette façon ont été obtenus, en termes algébriques [RI-17, RI-64] ou de calculabilité [**RI-42**]. De premiers résultats sur la dynamique du transfert d'information (auto-assemblage du pavage) ont aussi été trouvés [CI-14, 22].

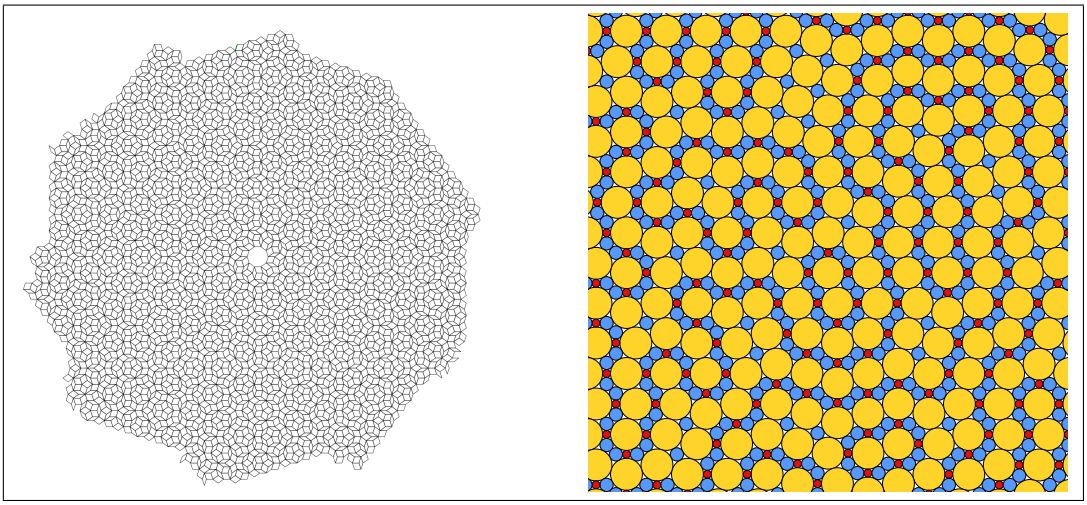


Figure 1.3: Auto-assemblage d'un pavage de Penrose à partir d'une graine (à gauche). Un des nombreux arrangements triangulés permis par trois tailles de disques bien spécifiques (à droite).

Par ailleurs, la recherche d'arrangements de disques apériodiques qui seraient caractérisés par le seul critère de maximisation de la densité, bien que pour l'instant infructueuse, a permis de donner des preuves de densité maximale pour certains arrangements à deux ou trois tailles de disques [RI-99, 23, 24] et de découvrir des candidats prometteurs [RI-72]. Les techniques utilisées, inspirées de celles utilisées par Hales pour démontrer la conjecture de Kepler [15], ont aussi conduit à de nouvelles bornes sur la densité maximale pour un continuum de cas [RI-112*] ou quand des conditions supplémentaires comme la stœchiométrie des différents disques entrent en jeu [**RI-111**].De premiers résultats sur les empilements de sphères bi-disperses et tri-disperses ont aussi été obtenus [RI-73, 21].

- **Combinatoire et géométrie convexe**

Nous étudions également des questions pertinentes à l'interface de la géométrie convexe et de la combinatoire. Vitali Milman a posé au début des années 2000 une question sur l'extrêmalité du volume de l'intersection d'un hypercube avec un hyperplan qui ne passe pas par le centre de l'hypercube [29], dans l'esprit du résultat de Keith Ball pour les hyperplans centraux [3]. Il a aussi conjecturé que ce volume n'est maximal ou minimal que lorsque l'hyperplan est orthogonal à une diagonale ou à une sous-diagonale de l'hypercube. Nous avons prouvé cette conjecture dans le cas où l'hyperplan est suffisamment éloigné du centre de l'hypercube [**RI-104***] et, localement, que cette conjecture revient à trouver le signe d'une certaine somme alternée et pondérée de coefficients binomiaux, faisant le lien entre ce problème de géométrie convexe et des questions classiques mais difficiles de combinatoire [**RI-105***]. Nous avons obtenu, comme conséquences une preuve locale de la conjecture de Milman pour des hyperplans relativement proches du centre de l'hypercube et, en utilisant l'ordinateur, pour des hyperplans arbitraires en dimension suffisamment faible (jusqu'en dimension 300).

- **Combinatoire algébrique**

Toute algèbre associative devient une algèbre de Lie pour le crochet de commutation. Cette relation fonctorielle admet un adjoint à gauche : l'algèbre enveloppante universelle. Une conséquence du théorème de Poincaré–Birkhoff–Witt établit qu'une algèbre de Lie, libre en tant que module sur son anneau de base, se plonge dans son algèbre enveloppante universelle. De façon similaire, une algèbre *différentielle* commutative (A, d) devient une algèbre de Lie différentielle (A, W_d, d) pour le wronskien $W_d(x, y) = xd(y) - d(x)y$. Cette construction est en fait un foncteur algébrique entre modèles de théories de Lawvere, et donc admet également un adjoint à gauche, l'*enveloppe wronskienne*. Dans [31] figure la conjecture de Lawvere selon laquelle toute algèbre de Lie se plonge dans son algèbre wronskienne. Si cette conjecture est vraie pour $\mathfrak{sl}_2(\mathbb{K})$, pour un corps $\mathbb{K}$ de caractéristique zéro, elle est fausse en générale : dans [RI-23] nous avons établi les conditions sous lesquelles une algèbre de Lie se plonge dans son enveloppe wronskienne.

Informations générales sur l'équipe CALIN

Les algèbres commutatives †-Frobenius sont les monoïdes commutatifs dans la catégorie des espaces hilbertiens pour lesquels l'adjoint de la multiplication satisfait la relation de compatibilité de Frobenius. Ces algèbres, nécessairement de dimension finie (sur le corps des complexes) ont été étudiées en mathématiques et en physique pour leur relation avec les théories quantiques des champs topologiques [2] bidimensionnelles et les invariants de surfaces. Pour disposer de telles algèbres en dimension infinie, il faut nécessairement considérer des algèbres non-unitaires. Dans [RI-83] nous avons caractérisé les conditions sous lesquelles de telles algèbres sont semi-simples (au sens des algèbres de Banach), étendant (et corrigeant) ainsi le résultat principal de [1].

1.2.3 Collaborations avec d'autres équipes du LIPN

Les chercheurs de l'équipe CALIN ont des liens avec ceux des autres équipes, en particulier avec les équipes AOC, LoVe et A^3.

La recherche sur les polytopes du point de vue de l'optimisation est un pont naturel avec l'équipe AOC ; L. Pournin (CALIN) a ainsi eté garant de l'HDR de R. Grappe (AOC). En outre, on compte deux co-publications avec des membres l'équipe AOC [CI-16, CI-6]. Sophie Huiberts, classée 1ère sur le concours CR CNRS de la section 6 du CoNRS de 2023, et dont les thématiques de recherches sont à cheval entre les deux équipes renforcerait si elle était affectée au LIPN.

Des chercheurs de l'équipe CALIN participent à l'axe transverse Complexités initié par des membres de l'équipe LoVe, et des chercheurs des deux équipes participent au projet l'ANR LambdaComb (2022-26).

La collaboration avec l'équipe A³ porte sur les tenseurs en dimension 3 : J. Ben Geloun (CALIN), M. Lebbah (A³) et leur doctorant ont conçu *une nouvelle méthode de clustering de données tridimensionelles*, le *Multi-Slice Clustering* [CI-27], qui, de manière remarquable, ne prend en hyperparamètres ni les tailles de clusters, ni leur nombre. Cet algorithme efficace garantit de plus la forte similarité des données dans les clusters et la forte séparation de ces clusters.

Références complémentaires

[1] S. Abramsky and C. Heunen. H^*-algebras and nonunital Frobenius algebras : first steps in infinite-dimensional categorical quantum mechanics. *Clifford Lectures*, 71 :1–24, 2012.

[2] M. F. Atiyah. Topological quantum field theory. *Publications Mathématiques de l'IHÉS*, 68 :175–186, 1988.

[3] K. Ball. Cube slicing in R^n. *Proceedings of the American Mathematical Society*, 97 :465–473, 1986.

[4] O. Bodini, A. Singh, and N. Zeilberger. Asymptotic distribution of parameters in trivalent maps and linear lambda terms. *arXiv :2106.08291*, 2022.

[5] B. Bollobás and O. Riordan. A polynomial of graphs on surfaces. *Mathematische Annalen*, 323 :81–96, 2002.

[6] M. Bóna. *Combinatorics of permutations*. Chapman-Hall and CRC Press, 2012.

[7] V. Bonzom. Maximizing the number of edges in three-dimensional colored triangulations whose building blocks are balls. *arXiv :1802.06419*, 2019.

[8] V. Bonzom. Tensor models with generalized melonic interactions. *arXiv :1905.01903*, 2019.

[9] V. Bonzom, G. Chapuy, S. Charbonnier, and E. Garcia-Failde. Topological recursion for Orlov–Scherbin tau functions, and constellations with internal faces. *arXiv :2206.14768*, 2022.

[10] V. Bonzom and N. Dub. Blobbed topological recursion for correlation functions in tensor models. *arXiv :2011.09399*, 2020.

[11] B. Bychkov, P. Dunin-Barkowski, M. Kazarian, and S. Shadrin. Symplectic duality for topological recursion. *arXiv :2206.14792*, 2022.

[12] J. Cardinal, L. Pournin, and M. Valencia-Pabon. The rotation distance of brooms. *arXiv :2211.07984*, 2022.

[13] M. Changata and J. Mathew. On triangle path convexity in graphs. *Discrete Mathematics*, 206 :91–95, 1999.

[14] A. Deza, G. Manoussakis, and S. Onn. Primitive zonotopes. *Discrete & Computational Geometry*, 60 :27–39, 2018.

[15] V. Disarlo and H. Parlier. The geometry of flip graphs and mapping class groups. *Transactions of the American Mathematical Society*, 372 :3809–3844, 2019.

[16] M. Dourado, F. Protti, and J. Szwarcfiter. Complexity results related to monophonique convexity. *Discrete Applied Mathematics*, 158 :1268–1274, 2010.

[17] P. Duchet and H. Meyniel. Ensembles convexes dans les graphes I : Théorèmes de Helly et de Radon pour graphes et surfaces. *European Journal of Combinatorics*, 4 :127–132, 1983.

[18] P. Duchon, P. Flajolet, G. Louchard, and G. Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. *Combinatorics, Probability and Computing*, 13(4-5) :577–625, 2004.

[19] N. M. Dunfield and W. P. Thurston. Finite covers of random 3-manifolds. *Inventiones Mathematicae*, 166 :457–521, 2006.

[20] B. Eynard. *Counting Surfaces*, volume 70 of *Progress in Mathematical Physics*. Birkhäuser/Springer, 2016.

[21] T. Fernique. Compact packings of space with three sizes of spheres. *arXiv :1912.02293*, 2019.

[22] T. Fernique and I. Galanov. Purely local growth of a quasicrystal. *arXiv :2201.00789*, 2022.

[23] T. Fernique and D. Pchelina. Compact packings are not always the densest. *arXiv :2104.12458*, 2021.

[24] T. Fernique and D. Pchelina. Density of triangulated ternary disc packings. *arXiv :2211.02905*, 2022.

[25] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[26] I. P. Goulden and D. M. Jackson. The KP hierarchy, branched covers, and triangulations. *Advances in Mathematics*, 219 :932–951, 2008.

[27] T. C. Hales. A proof of the Kepler conjecture. *Annals of Mathematics*, 162 :1065–1185, 2005.

[28] S. Kitaev. *Patterns in permutations and words*. Springer, 2011.

[29] H. König and A. Koldobsky. Volumes of low-dimensional slabs and sections in the cube. *Advances in Applied Mathematics*, 47 :894–907, 2011.

[30] S. K. Lando and A. K. Zvonkin. *Graphs on Surfaces and Their Applications*, volume 141 of *Encyclopaedia of Mathematical Sciences*. Springer, 2004.

Informations générales sur l'équipe CALIN

[31] W. Lawvere. Functorial semantics of algebraic theories, (summary of Ph.D. thesis, Columbia). *Proceedings of the National Academy of Sciences of the United States of America*, 50 :869–872, 1963.

[32] F. D. Murnaghan. The analysis of the kronecker product of irreducible representations of the symmetric group. *American Journal of Mathematics*, 60 :761–784, 1938.

[33] L. Pournin and Z. Wang. Strong convexity in flip-graphs. *arXiv :2106.08012*, 2021.

[34] D. Sleator, R. Tarjan, and W. Thurston. Rotation distance, triangulations, and hyperbolic geometry. *Journal of the American Mathematical Society*, 1 :647–681, 1988.

[35] V. Vatter. *Handbook of Enumerative Combinatorics*, chapter Permutation classes. Chapman-Hall and CRC Press, 2015.

1.3 Profil d'activités liées à la recherche

Activités (répartir 100 points sur ces 7 items)	
Administration et animation de la recherche : pilotage de la recherche (VP, direction d'institut, DAS, par exemple), participation à des instances d'évaluation (CNU, CoNRS, CSS, Hcéres, par exemple), responsabilité de dispositifs Idex ou Isite, direction de projets (ANR, Horizon Europe, ERC, CPER, PIA, France 2030, par exemple), responsabilités éditoriales dans des revues ou collections nationales et internationales.	25
Aide aux politiques publiques et expertise technique : pouvoirs publics aux niveaux européen, national et régional, entreprises, instances internationales comme FAO, OMS, etc.	0
Contribution à l'adossement d'enseignements innovants à la recherche : EUR, SFRI, etc.	10
Dissémination de la recherche : partage de connaissances avec le grand public, médiation scientifique, interface sciences et société	5
Recherche et encadrement de la recherche.	60
Valorisation, transfert, innovation.	0
Autres activités. (à préciser en une ligne maximum).	0

Informations générales sur l'équipe CALIN

1.4 Prise en compte des recommandations du précédent rapport

Recommandation : "Le comité recommande de continuer sur la lancée, avec un regard vers les grandes conférences d'informatique fondamentale qui peuvent apporter une visibilité plus rapide sur les résultats et à l'équipe."

Réponse : *Dans les thématiques de l'équipe, les journaux sont significativement mieux valorisés que les conférence et notre objectif est d'avoir une qualité de publication la meilleure possible.*

Recommandation : "Il faut veiller à ce que l'activité de développement d'un paquet de combinatoire dans le logiciel Sage repose sur un nombre adapté de contributeurs."

Réponse : *Au sein du LIPN, un seul membre est administrateur de Sage et fait un travail significatif de formation et de maintien autour du logiciel.*

L'équipe a par ailleurs accueilli 2 postdoctorantss qui ont également contribué à du code libre autour de l'énumération et la génération de structures combinatoires.

Recommandation : "Il est important de susciter d'excellentes candidatures sur les postes de chercheurs (au recrutement et à la mutation) et d'enseignants-chercheurs, en particulier au niveau professeur et directeur de recherche."

Réponse : *L'équipe a chaque année des candidats CNRS (3 au printemps 2023, dont un est premier sur liste complémentaire à l'issue de la phase d'admissibilité). C'est sur un poste ouvert à tout profil qu'une personne a été recrutée pour l'équip CALIN.*

Recommandation : "Tout en poursuivant l'implication dans la recherche de doctorants à l'international, l'équipe aura à explorer une solution locale aux problèmes de recrutement de doctorants via la construction d'un parcours de master adapté : la croissance actuelle des effectifs d'étudiants de premier cycle, tout particulièrement dans les filières scientifiques, va prochainement atteindre le master et il convient d'avoir anticipé ceci plutôt que d'y réagir tardivement."

Réponse : *L'EUR M&CS (École Universitaire de Recherche en math-info, voir section 3.1.2) a démarré en octobre 2020 avec de meilleurs étudiants. Ses effectifs montent en puissance (environ 50% de ses étudiants de M2 continuent en thèse en France, soit 1 doctorant en 2021, 2 en 2022, 3 en 2023) et ses doctorants se répartissent sur toutes les thématiques du LAGA et du LIPN. Pour l'instant, aucun doctorant issu de cette filière n'a encore été recruté dans l'équipe CALIN.*

Recommandation : "L'activité de séminaire doit être poursuivie de façon aussi dynamique que la disponibilité des membres de l'équipe le permet."

Réponse : *Les membres de l'équipe participent au séminaire de façon régulière.*

Recommandation : "Les riches pratiques d'échange entre les différentes thématiques représentées dans l'équipe ne doivent pas être perdues."

Réponse : *Entre 2018 et 2023, cette pratique ne s'est pas perdue (par exemple, plusieurs co-encardements de thèses ont eu lieu depuis entre des membres de l'équipe de culture scientifique relativement différente). Les échanges se sont mêmes amplifiés sur des sujets faisant des ponts entre les différents thèmes de l'équipe, notamment sur des sujets de combinatoire géométrique ou de la géométrie discrète et algorithmique.*

Recommandation : "Le projet actuel est dans une certaine forme de continuité : il faut réfléchir dès maintenant à l'avenir lointain et au renouvellement."

Réponse : *Le projet a évolué avec l'arrivée (2021) d'un PU. Les aspects géométriques sont devenus transverses aux activités de l'équipe. Un recrutement qui renforcerait cette dimension de l'activité de l'équipe est prioritaire.*

Chapitre 2
Introduction du portfolio

Nous présentons ici brièvement les cinq contributions que nous avons retenues pour le portfolio de l'équipe CALIN. Une présentation plus détaillée de chacune de ces contributions se trouve dans le portfolio lui-même, disponible en annexe. Nous avons sélectionné un échantillon de quatre articles et un livre, qui nous semblent mettre en évidence la qualité, la diversité et la cohérence des recherches menées dans l'équipe CALIN. La qualité de ces résultats se traduit notamment par des publications dans des journaux du meilleur niveau international.

1. **Universalité du permuton séparable brownien.** Nous avons étudié d'un point de vue probabiliste des objets bien connus en combinatoire, les classes de permutations. Ils ont exhibé un permuton limite, le *permuton séparable brownien*, qui est la limite des permutations séparables aléatoires uniformes [**RI-38**]. La déformation à un paramètre de ce permuton s'avère être la limite de permutations tirées aléatoirement uniformément dans de nombreuses classes [**RI-65**, **RI-100**].

   Cette ligne de recherche, qui a immédiatement reçu un certain écho dans la communauté scientifique [5], est une illustration des fructueuses recherches menées à l'interface entre combinatoire et probabilités dans l'équipe CALIN et en collaboration avec l'équipe Probabilités et Statistiques du LAGA au sein de la fédération MathSTIC.

2. **Complexité et combinatoire algébrique : une interprétation combinatoire des symboles de Kronecker.** Interpréter combinatoirement les symboles de Kronecker est un problème ouvert depuis 1937 et qui s'est révélé au cœur des recherches en théorie géométrique de la complexité. Cet article [**RI-118**], aboutissement de recherches au long cours [**RI-15** , 1], y parvient via une construction extrêmement habile. Ces recherches à l'interface de l'informatique, des mathématiques et de la physique sont une belle illustration de la richesse des approches interdisciplinaires développées dans l'équipe CALIN.
3. **Cartes et hiérarchie BKP.** Les cartes combinatoires sont des objets topologiques apparaissant en informatique, en physique et en mathématiques. Goulden et Jackson [14] ont montré que l'énumération de ces cartes est liée à un ensemble d'équations aux dérivées partielles, la *hiérarchie KP* qui a conduit à plusieurs résultats remarquables ces dernières années [2, 3, 7].

   Nous avons montré, dans le même esprit, qu'un autre ensemble d'équations aux dérivées partielles, la hiérarchie BKP est liée au cas des cartes *non-orientées* [**RI-94**]. La difficulté majeure de la preuve réside dans un délicat changement de base (irréalisable en toute généralité) dans l'espace des fonctions symétriques, (car connaître le développement des séries sur les polynômes de Schur est) nécessaire pour prouver l'intégrabilité. Ce résultat profond et unificateur de combinatoire algénbrique a de nombreux corollaires telles les preuves d'une conjecture de Féray [4], d'une autre due à Oliveira et Novaes [8] ainsi que des formules de récurrence pour calculer le nombre de cartes non-orientées.
4. **Échantillonage aléatoire : aspects probabilistes, combinatoires et géométriques.**

   La quantité de données traitées par les algorithmes a explosé ces dernières années et afin de préserver leur efficacité, une méthode très utilisée consiste à produire un sous-ensemble de données représentatif auquel appliquer ces algorithmes. Ces méthodes d'approximation sont à l'interface de l'informatique théorique, des mathématiques et des statistiques et se basent sur des outils combinatoires. Le livre "Sampling in combinatorial and geometric set systems" [**LI-2**] publié par l'American Mathematical Society en 2022 fournit un traitement rigoureux de ces méthodes, qui introduit et explique la recherche de la dernière décennie sur un sujet de l'informatique dont la visibilité est devenue exceptionnelle et dont l'importance est fondamentale. Ce livre est à destination des étudiants et des spécialistes en combinatoire, en géométrie algorithmique ou en apprentissage automatique.

   En bref, il s'agit d'une monographie sur un sujet brulant à destination d'un larger public.
5. **Zonotopes et complexité de l'optimisation linéaire.** L'algorithme du simplexe, dont l'objectif est de résoudre les problèmes d'optimisation linéaire, fait partie des algorithmes les plus importants de l'informatique. Dans l'esprit du problème de Steve Smale [9] sur la complexité de ce type d'algorithmes, Deza, Pournin et Sukegawa ont obtenu une expression exacte du nombre de pivots nécessaires pour résoudre un problème d'optimisation linéaire [**RI-57**, RI-85] dans le cas le pire, lorsque la taille en mémoire de l'exécution est bornée. Ce résultat de complexité obtenu par des outils de géométrie, de combinatoire et de théorie des nombres illustre l'apport de méthodes croisées pour attaquer ce grand problème d'informatique théorique.

Références complémentaires

[1] J. Ben Geloun and S. Ramgoolam. Counting tensor model observables and branched covers of the 2-sphere. *Annales de l'Institut Henri Poincaré D*, 1(1) :77–138, 2014.

[2] T. Budzinski and B. Louf. Local limits of uniform triangulations in high genus. *Inventiones mathematicae*, 223 :1–47, 2020.

[3] S. R. Carrell and G. Chapuy. Simple recurrence formulas to count maps on orientable surfaces. *Journal of Combinatorial Theory, Series A*, 133 :58–75, 2015.

[4] V. Féray. On complete functions in Jucys–Murphy elements. *Annals of Combinatorics*, 16 :677–707, 2012.

[5] E. Ghys. *A singular mathematical promenade*. CNRS Editions, 2017. https://perso.ens-lyon.fr/ghys/promenade/.

[6] I. P. Goulden and D. M. Jackson. The KP hierarchy, branched covers, and triangulations. *Advances in Mathematics*, 219 :932–951, 2008.

[7] S. Janson and B. Louf. Unicellular maps vs hyperbolic surfaces in large genus : simple closed curves. *arXiv :2111.11903*, 2021.

[8] L. H. Oliveira and M. Novaes. Immanants of blocks from random matrices in some unitary ensembles. *Journal of Physics A*, 54 :295205, 2021.

[9] S. Smale. Mathematical problems for the next century. *Mathematical Intelligencer*, 20 :7–15, 1998.

Chapitre 3

Auto-évaluation de l'équipe CALIN

Sommaire

3.1 Domaine 2. Attractivité 194

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche 194

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels 196

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs 198

3.2 Domaine 3. Production scientifique 199

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité. 199

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels. 201

3.2.3 Publications 202

3.3 Domaine 4. Inscription des activités de recherche dans la société . 214

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique. 214

3.3.2 Référence 3. L'équipe partage ses connaissances avec le grand public et intervient dans des débats de société. 215

3.1 Domaine 2. Attractivité

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche

- **Invitations dans des institutions et congrès internationaux**

Les membres de l'équipe CALIN sont fréquemment invités dans dans des institutions académiques ou des congrès internationaux. Parmi ces institutions académiques figurent bien-sûr des universités (en Afrique, en Amérique du Nord et du Sud, en Asie et en Europe), ainsi que des centres de recherche internationaux (Dagstuhl, Fields Institute, ICERM, Oberwolfach, etc).

Voici un tableau résumant ces invitations.

Institution	Pays	date
African Institute for Mathematical Sciences	Sénégal	février 2017, février 2018 février 2019, février 2020
Fields Institute	Canada	février 2017 novembre 2019
Oberwolfach	Allemagne	avril 2017
Queen Mary University London	Grande Bretagne	mai 2017 juillet 2017
Technische Universität Wien	Autriche	septembre 2017 septembre 2018
Banff International Research Station	Canada	septembre 2017 mars 2019
Academia Sinica	Taïwan	octobre 2017
National Chiao Tung University	Taïwan	octobre 2017
Université de Maribor	Slovénie	avril 2018
Perimeter Institute	Canada	août 2018 février 2019 octobre 2019
Université du Luxembourg	Luxembourg	novembre 2018
Université de Belo Horizonte	Brésil	juin 2019
Alpen–Adria–Universität Klagenfurt	Autriche	juillet–septembre 2019
Yukawa Institute Kyoto	Japon	mars 2019
Université Fédérale de Rio de Janeiro	Brésil	juillet 2019
Université d'Heidelberg	Allemagne	juillet 2019
IMPA	Brésil	juillet 2019
Université du Chili	Chili	novembre 2019
Freie Universität Berlin	Allemagne	janvier 2020
Leibniz-Zentrum für Informatik Dagstuhl	Allemagne	février 2022 février 2023 février 2024
ICERM, Brown University	États-Unis	mars–avril 2023

- **Participation à comités de programme de conférences**

Les membres de l'équipe participent aux comités de programme de conférences du domaine parmi lesquelles on trouve pare exemple les conférences récurrentes LATIN ou LAGOS qui rassemblent des centaines de participants. Voici un tableau récapitulatif des conférences pour lesquelles des membres de l'équipe CALIN ont participé au comité de programme.

Conférence	Lieu	Date
CSR 2017	Kazan, Russie	juin 2017
LAGOS 2017	CIRM, Marseille	septembre 2017
ANALCO 2018	New Orleans, États-Unis	janvier 2018
Applied Mathematics and Random Structures	Birzeït, Palestine	août 2018
LATIN 2020	São Paulo, Brésil	janvier 2021
LAGOS 2021	São Paulo, Brésil	mai 2021
Lattice Paths 2021	CIRM, Marseille	juin 2021
CPM'21	Wrocław, Pologne	juillet 2021
GASCom 2022	Varèse, Italie	juin 2022
CPM'22	Prague, Tchéquie	juin 2022
ICGT 2022	Montpellier, France	juillet 2022
LATIN 2022	Guanajuato, Mexique	octobre 2022
GASCom 2024	Bordeaux, France	juin 2024

- **Organisations de conférences et d'évènements**

L'équipe est également impliquée dans l'organisation de conférences internationales ou d'évènements. Le tableau suivant en donne la liste.

Conférence	Lieu	Date
Combinatoire, Arithmétique, Physique	IHÉS, Bures-sur-Yvette	annuel, 2017–2022
LAGOS 2017	CIRM, Marseille	septembre 2017
Groupe de travail Pytheas Fogg	Marseille	depuis 2018
MathExp 2018	Saint Flour	mai 2018
Applied Math. and Random Structures	Birzeït, Palestine	août 2018
International Science Fair	Addis-Abeba, Éthiopie	décembre 2018
Journée Cartes	LIPN/LAGA	avril 2019
AofA'19	CIRM, Marseille	juin 2019
Lost in Translation Surfaces	Univ. Luxembourg	janvier 2020
Journées Cartes	CIRM, Marseille	janvier 2021
Quantum Gravity in Bordeaux	Univ. Bordeaux	octobre 2019
Lattice Paths 2021	CIRM, Marseille	juin 2021
Structures on Surfaces	CIRM, Marseille	mai 2022
Group Theoretical Methods in Physics	Cotonou, Bénin	reporté en 2024

- **Activités éditoriales**

Le journal international Annales de l'Institut Henri Poincaré D Combinatorics, Physics and their Interactions, publié par l'EMS (European Mathematical Society) a été co-fondé en 2014 par des membres de l'équipe CALIN. Deux membres de l'équipe en sont éditeurs depuis 2014.

Un membre de l'équipe est co-éditrice en chef du journal international de combinatoire Annals of Combinatorics.

- **Participation à des instances de pilotage de la recherche**

L'équipe est investie au niveau du pilotage de la recherche, avec notamment

Administration de la recherche :

— Direction du LIPN depuis 2018.
— Direction (adjointe) de l'ED Galilée (depuis 2018) ;
— Responsabilité du GT ALEA du GDR IM de 2012 à 2018.
— Responsabilité du GDRI AleaNetwork de 2015 à 2018.

Comités d'expertise ou d'évaluation de la recherche : :

— Groupe d'Evaluation des Sciences Informatiques pour le programme de subventions à la découverte (NSERC) (l'équivalent Canadien de l'ANR) de 2019 à 2022.
— Jury de l'Institut Universitaire de France en 2017.
— Jury du CAPES NSI (informatique) depuis 2020.
— Comité HCERES (présidence, en 2019).
— Conseil scientifique du LabEx Milyon depuis 2019.

- **Prix, distinctions**

Les membres de l'équipe ont reçu le *Young Scientist Prize de l'International Union of Pure and Applied Physics 2015-2017*, le *Next Einstein Forum Prize 2015-2017*, le *SageMath Prize 2015-2021* et le *prix Artex des pépinières européennes de création en 2021*.

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels

Auto-évaluation de l'équipe CALIN

- **Formations de niveau master**

Pour attirer des étudiants (aussi bien en master qu'en doctorat) avec une formation théorique solide, la fédération MathSTIC a été à l'initiative d'un ambitieux programme d'École Universitaire de Recherche (EUR) en mathématiques et informatique. Elle a ainsi obtenu en 2020 un financement européen pour cette EUR Math & Computer Sciences Graduate school, ce qui lui permet d'offrir une bourse d'étude à des étudiants (de master 1 et master 2) internationaux sélectionnés au mérite. Ils suivent alors le master de leur choix parmi le panel des formations en maths et/ou en info de l'université. Certains cours et formations propres à l'EUR leurs sont également donnés (en anglais).

Après des difficultés de mise en place en période de pandémie, la formation est passée de 4 M1/M2+1 doctorant en 2020/2021, à 24 M1/M2+3 doctorants en 2022/2023, et devrait continuer à monter en effectif dans les années à venir jusqu'en 2028. Chaque année, la majorité de ces élèves de M2 continuent en thèse au LIPN et au LAGA.

Les étudiants de l'EUR constituent ainsi un vivier local particulièrement intéressant pour l'équipe CALIN toujours en quête de doctorants en informatique ayant un solide bagage en mathématiques (analyse complexe, probabilités, géométrie, topologie, algèbre...).

Dans la dynamique MathSTIC insufflée par les collaborations au sein de la fédération MathSTIC, des membres de l'équipe CALIN et de l'équipe Probabilités et Statistiques du LAGA interviennent depuis 2022 dans deux cours de cette EUR, d'une part via un cours mutualisé de *Combinatoire et Probabilités* (également suivi par des doctorants), et d'autres part via des séances d'initiation à la présentation de résultats scientifiques (latex, lecture d'article, petits séminaires). Lionel Pournin intervient également dans l'EUR via un cours de géométrie algorithmique en première année de master.

- **Formation doctorale**

Les membres de l'équipe CALIN participent de façon très active à la formation par la recherche. Ils interviennent comme enseignants dans le Master de Parisien de Recherche en Informatique, dans le Master 2 de la MISA à Madagascar et dans l'École Universitaire de Recherche (EUR) de l'université Sorbonne Paris Nord. Ils organisent des groupes de travail ciblés pour les doctorants comme le groupe de travail Combinatoire, Informatique et Physique (voir la page du séminaire d'équipe) ou le groupe tilearn. Le séminaire d'équipe, largement suivi est aussi un lieu privilégié d'échanges pour tous les membres de l'équipe et en particulier pour les doctorants et post-doctorants. Les membres de l'équipe CALIN donnent aussi des cours dans des écoles d'été, par exemple dans une école d'été consacrée aux questions combinatoires et algorithmiques en algèbre à Trissur dans le Kérala (Inde) : outre les 200 participants présents, l'ensemble des conférences était retransmis en direct dans tous les collèges du Kérala. Un membre de l'équipe est orateur invité à l'école d'été *Convex and Discrete Geometry summer school* qui aura lieu au centre Erdő de l'institut Alfréd Rényi à Budapest du 28 août au 1er septembre 2023.

Depuis 2017, l'équipe a accueilli, comme post-doctorants, Christian Bean, Julien Courtiel, Martin Pépin et Michael Wallner. L'équipe dispose d'un recrutement de doctorants de grande qualité au niveau international (Russie, Madagascar, Vietnam) ou national (Polytechnique, École Normale Supérieure). Le tableau suivant récapitule les doctorants de l'équipe CALIN depuis 2017.

Nom	Provenance	Date de soutenance
Quentin De Mourgues		décembre 2017
Tsinjo Rakotoarimalala	MISA (Madagascar)	juillet 2019
Sergey Dovgal	Russie	septembre 2019
Ilya Galanov	Russie	décembre 2019
Nina Pardal		mars 2020
Rado Rakotonarivo	MISA (Madagascar)	mars 2020
Matteo D'Achille	University of Milan	octobre 2020
Mehdi Naima	Master informatique	décembre 2020
Nicolas Dub		juin 2021
Victor Lutfalla	ÉNS Lyon	juillet 2021
Khaydar Nurligareev	Russie	octobre 2022
Alexandros Singh		novembre 2022
Clémence Chinaud-Chaix		janvier 2023
Asma Althagafi	King Saud University	en cours
Théophile Buffière	École Polytechnique	en cours
Théo Lenoir	ÉNS Ulm	en cours
Alexandre Louvet		en cours
Kristina Matrossova	Master informatique	en cours
Victor Nador		en cours
Vu Nguyen Dinh	Vietnam	en cours
Daria Pchelina	Russie	en cours
Carole Porrier		en cours

L'équipe de chercheurs et enseignants-chercheurs permanents a légèrement évolué depuis 2017 avec l'arrivée de Nabil Mustafa (en septembre 2021) et le départ de Mario Valencia (en septembre 2022). Valentin Bonzom, un membre de l'équipe très actif en recherche vient d'être promu professeur à l'université Gustave Eiffel à l'automne 2023. Les candidats à une intégration dans l'équipe sont de grande qualité : Nabil Mustafa a été recruté sur un poste ouvert à tous les profils thématiques du laboratoire. En outre, chaque année, des candidats sérieux au concours CR CNRS demandent leur affectation dans l'équipe. Ceci est illustré par la campagne de recrutement du printemps 2023, lors de laquelle 3 candidats CNRS ont demandé une affectation dans l'équipe CALIN, et l'un

d'eux (Florent Koechlin) est arrivé 1er sur liste complémentaire. Lors de cette même campagne, Sophie Huiberts qui a demandé une affectation dans l'équipe AOC mais dont les thèmes de recherche sont proches de ceux de l'équipe CALIN a été classée première.

Depuis 2017, l'équipe a accueilli de nombreux professeurs invités de renom travaillant à l'étranger. Une partie d'entre eux sont venus par le programme de professeurs invités de l'université Paris 13 et d'autres financés directement par le laboratoire ou par des projets ANR. Le tableau suivant récapitule les chercheurs invités dans l'équipe CALIN depuis 2017

Nom	Provenance	Date
Imre Bárány	Budapest, Hongrie	2021
Jérémie Bouttier	Lyon, France	2021
Brešar, Boštjan	Maribor, Slovénie	2017
Jean Cardinal	Bruxelles, Belgique	2020, 2021 et 2022
Michael Drmota	Wien, Autriche	2021
Samantha Fairchild	Leipzig, Allemagne	2021
Florian Girelli	Waterloo, Canada	2018
Amir Hashemi	Ispahan, Iran	2019
Wolfgang Mulzer	Berlin, Allemagne	2022
Gleb Koshevoy	Moscou, Russie	2018
Florian Luca	Johannesburg, Afrique du Sud	2018
Marton Naszodi	Budapest, Hongrie	2022
Alantha Newman	Grenoble, France	2022
Hugo Parlier	Esch-sur-Alzette, Luxembourg	2018 et 2022
Benedikt Stufler	Wien, Autriche	2021
Roland Van der Veen	Groningen, Pays Bas	2019

Auto-évaluation de l'équipe CALIN

Les moyens mis en œuvre par l'équipe et plus largement le laboratoire pour l'accueil de ses chercheurs permanents, de ses doctorants, de ses post-doctorants et de ses invités comprennent l'achat d'ordinateurs portables et de matériels informatique, la mise à disposition de moyens de calcul (cluster MAGI), l'incitation à participer aux écoles thématiques (comme par exemple les journées ALEA qui ont lieu au CIRM chaque année) et aux conférences du domaine.

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs

Les membres de l'équipe CALIN participent à de nombreux projets compétitifs. Le tableau suivant regroupe ces projets depuis 2017.

Type	Nom	Dates
ANR	QuasiCool (ANR-12-JS02-0011)	2013–2018
ANR franco–taïwanaise	MétAConC (ANR-15-CE40-0014)	2015–2020
ANR franco–luxembourgeois	SoS (ANR-17-CE40-0033)	2018–2023
ANR	ADDS (ANR-19-CE48-0005)	2019–2023
ANR	DIMERS (ANR-18-CE40-0033)	2019–2023
ANR	RECORDS (ANR-19-CE38-0013)	2019–2023
80PRIME	ASAP–Assemblage de Supercristaux par Approche Prédictive	2019–2023
PHC franco–belge	Tournesol (42703TD)	2020–2021
MathAmSud	STALGRAPH (MathAmSud 20MATH-09)	2020–2022
ERC	From Geometry to Combinatorics and Back: Escaping the Curse of Dimensionality	2020–2025
ANR	3DMaps (ANR-20-CE48-0018)	2021–2024
ANR	LambdaComb (ANR-21-CE48-0017)	2022–2025

3.2 Domaine 3. Production scientifique

Nous fournissons ci-dessous un tableau de synthèse des publications de l'équipe, organisées par année et par typologie :

	2017	2018	2019	2020	2021	2022	2023	*Total*
Livres	0	0	0	1	0	1	0	*2*
Chapitres de livre	0	0	0	0	0	0	0	*0*
Revues internationales	22	17	15	13	15	21	19	*122*
Autres revues	0	0	0	0	0	0	0	*0*
Conférences internationales	4	7	2	7	7	4	0	*31*
Autres communications	3	7	1	1	1	2	0	*15*
Total	*29*	*31*	*18*	*22*	*23*	*28*	*19*	*170*

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité.

La production scientifique de l'équipe repose sur des fondements théoriques et méthodologiques solides. En témoignent les journaux dans lesquels les résultats obtenus sont publiés, qui comptent parmi les meilleurs de trois champs différents, la combinatoire, les mathématiques et la physique. Certains de ces journaux ont une aura exceptionnelle et des critères de sélection extrêmement difficile comme par exemple *Journal of the European Mathematical Society (JEMS)* dans lequel nous avons publié deux articles depuis 2017 [**RI-65**, **RI-4**] ou *Advances in Mathematics* [**RI-100**]. Parmi les autre journaux prestigieux dans lesquels les membres de l'équipe ont publié on trouve des journaux de mathématiques généralistes comme *Israel Journal of Mathematics* [**RI-2**, **RI-17**, **RI-104***], *Proceedings of the American Mathematical Society* [**RI-57**], *International Mathematics Research Notices* [**RI-94**], *Mathematika* [RI-85] ou *Canadian Mathematical Bulletin* [RI-44], des journaux de mathématiques spécialisés comme *Annals of Probability* [**RI-38**, **RI-51**] ou *Journal d'Analyse Mathématique* [**RI-105***], des journaux de combinatoire comme *Journal of Combinatorial Theory, Series A* [**RI-59**, **RI-111**], *Combinatorial Theory* [**RI-102**], *Combinatorics, Probability and Computing* [RI-98, RI-103], *Journal of Graph Theory* [RI-40], *Random structures and Algorithms* [**RI-101**,

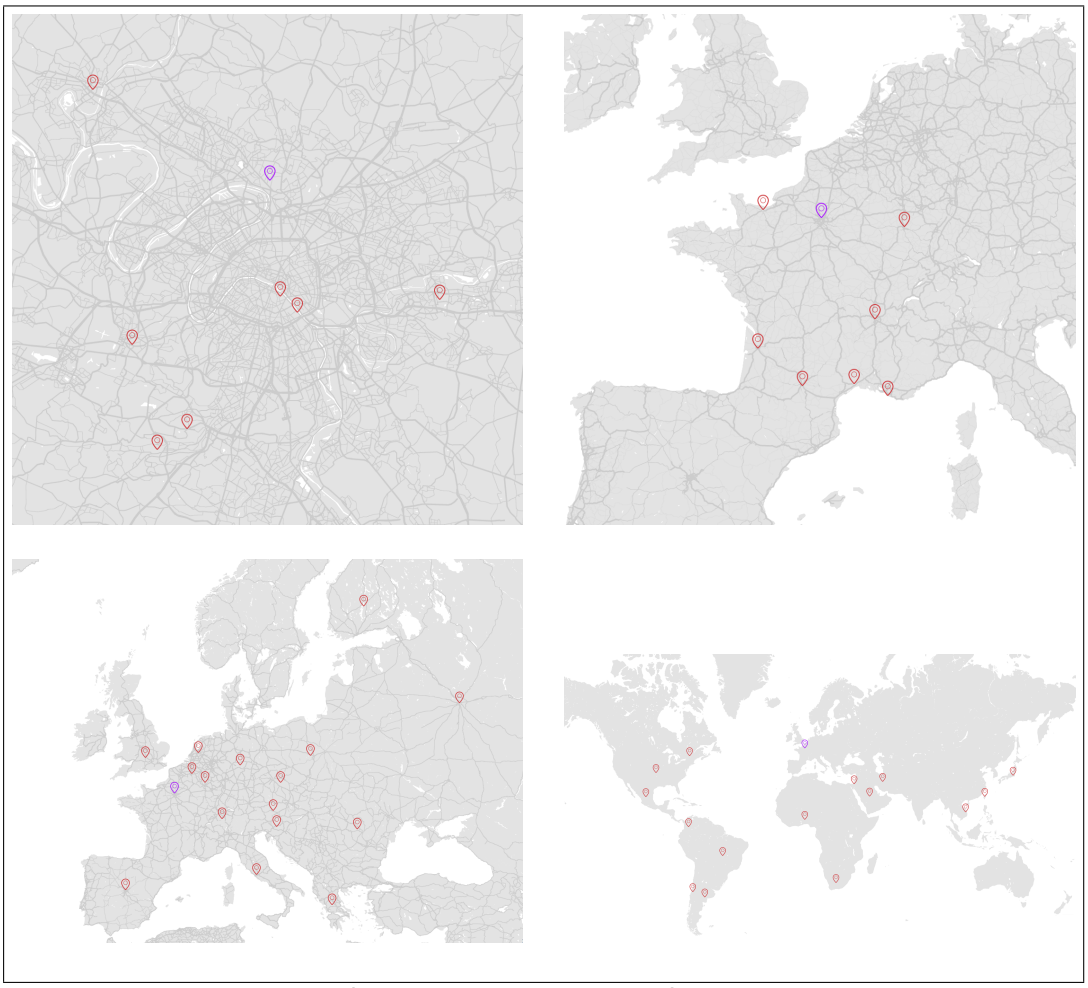


Figure 3.1: Implantation régionale, nationale, européenne et mondiale des co-auteurs de l'équipe CALIN pour les articles publiés entre 2017 et 2022.

RI-96] ou *European Journal of Combinatorics* [RI-1, RI-25] ainsi que des journaux de physique comme *Journal of High Energy Physics* [**RI-15**], *Physical Review D* [**RI-36**] ou *European Physical Journal C* [RI-37].

Ces articles sont co-écrits par les membres de l'équipe CALIN avec de nombreux chercheurs Français et étrangers. Nous collaborons avec des laboratoires d'informatique, de mathématiques et de physique du monde entier dont l'implantation est indiquée sur la figure 3.1.

Le tableau suivant récapitule le nombre de publications co-écrites depuis 2017 avec des laboratoires Français par discipline (les articles co-écrits avec plusieurs laboratoires de disciplines différentes sont comptés plusieurs fois).

Table 3.1 Nombre de co-publications depuis 2017 entre l'équipe CALIN et des laboratoires Français (hors LIPN), par discipline de ces laboratoires.

Pays du laboratoire	Nombre de publications co-écrites
Informatique	36
Mathématiques	13
Physique	4

Le tableau suivant regroupe le nombre d'articles co-écrits depuis 2017 avec des chercheurs travaillant dans des laboratoires de continents différents (les articles co-

écrits avec des chercheurs de continents différents sont comptés plusieurs fois).

Table 3.2 Nombre de co-publications depuis 2017 entre l'équipe CALIN et des laboratoires Européeans (hors France) ou sur d'autres continents.

Localisation du laboratoire	Nombre de publications co-écrites
Afrique	13
Amérique du Nord	17
Amérique du Sud	12
Asie	7
Europe (hors France)	55

L'équipe contribue aussi à l'administration, au maintien et au développement du logiciel de mathématiques *open source* SageMath. Elle contribue au code source de Sage, a développé et maintient la clef USB live officielle de Sage qui permet à toute personne disposant d'un ordinateur d'utiliser le logiciel en évitant les contraintes d'installation, administre le miroir pour la distribution de sage-on-gentoo (au LIPN), le serveur de sauvegarde (à Orsay), le serveur de patchbot (à Strasbourg) et les serveurs asksage et asksage-db (à l'Ohio State University), voir https://wiki.sagemath.org/Infrastructure. En plus de cela, l'équipe assure le support en ligne du logiciel, organise des formations à SageMath dans le cadre d'écoles et de conférences. On pourra se reporter à la page SageMath au LIPN.

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels.

Auto-évaluation de l'équipe CALIN

Avec *122* articles de journaux internationaux, *31* articles de conférences internationales et *deux* livres publiés depuis 2017 pour 15 membres permanents (dont un en détachement, un en échange de service et un en disponibilité) l'équipe montre un excellent dynamisme et sa production scientifique est proportionnée à son potentiel de recherche. Onze de ces articles ont été coécrits par au moins deux membres permanents de l'équipe, deux en collaborations avec des membres d'autres équipes du LIPN, trois avec des membres du laboratoire de mathématiques (LAGA) de l'Université Sorbonne Paris Nord. Tous les membres de permanents de l'équipe ont participé à cette production depuis 2017 (voir références). Le nombre d'articles par permanent et par an (équivalent temps plein, en comptabilisant les enseignants-chercheurs à un taux de 50% et en prenant en compte les dates de départs et arrivées) est de *1.97* pour les articles de journaux et de *0.61* pour les conférences.

Au sujet de l'intégration des non-permanents dans la production scientifique, mentionnons que 21 des articles publiés depuis 2017 ont été co-écrits avec un doctorant ou un post-doctorant dans l'équipe et que quatre de ces articles ont été publiés par un doctorant sans autre membre de l'équipe [RI-87, CI-28, RI-69, RI-84].

L'accompagnement des jeunes chercheurs est l'objet de différentes actions spécifiques de la part de l'équipe (hors accompagnement des doctorants par leurs directeurs de thèse). Le séminaire d'équipe hebdomadaire permet à tous et notamment à nos jeunes chercheurs d'étayer leur culture scientifique et de se mettre au courant des avancées dans nos domaines. Certaines séances de notre séminaire sont du type "groupe de travail".L'équipe met en place au besoin des groupes de travail ciblés thématiquement, nous avons notamment eu un groupe sur les pavages, le groupe tilearn, à destination des doctorants et chercheurs de l'équipe. Nous participons aussi au groupe de travail inter-équipes du LIPN axé sur la théorie de la complexité. Enfin des soutenances de mi-parcours ont été mises en place pour les doctorant afin de détecter les problèmes éventuels au niveau scientifique et au niveau de l'encadrement doctoral. L'activité de

l'équipe sur le plan de la formation à la recherche est décrite de façon plus détaillée dans la section 3.1.2.

Tous les membres de l'équipe sont actifs en recherche, il n'est par conséquent pas nécessaire de mettre en place des actions d'accompagnement des chercheurs qui reprennent une activité recherche. Cependant, si la situation venait à se produire, le séminaire d'équipe et les séminaires thématiques seraient une façon efficace de leur permettre de reprendre une activité de recherche. Nous veillons également à solliciter les différents membres de l'équipe sur les thématiques du moment pour les inciter à travailler dessus.

3.2.3 Publications

Les références
- comportant « * » sont acceptées en 2022 et à paraître.
- indiquées sur fond gris sont les productions phare de chaque catégorie.
- en gras constituent les 20% les meilleures de la production scientifique de chaque catégorie.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2023

[RI-122] R. C. Avohou, J. Ben Geloun et M. N. Hounkonnou. Universality for polynomial invariants for ribbon graphs with half-ribbons. *Discrete Applied Mathematics*, 327 :1–17, 2023.

[RI-121*] C. Banderier, M. Kuba et M. Wallner. Phase transitions of composition schemes : Mittag-Leffler and mixed Poisson distributions. *Annals of Applied Probability*, XX :57 pages, 2023.

[RI-120*] N. Behr, G. H. Duchamp et K. A. Penson. Explicit formulae for all higher order exponential lacunary generating functions of Hermite polynomials. *Séminaire Lotharingien de Combinatoire*, XX :29 pages, 2023.

[RI-119*] J. Ben Geloun et S. Ramgoolam. Kronecker coefficients from algebras of bi-partite ribbon graphs. *Noncommutativity and Physics*, ST :13 pages, 2023.

[**RI-118**] J. Ben Geloun et S. Ramgoolam. Quantum mechanics of bipartite ribbon graphs : Integrality, Lattices and Kronecker coefficients. *Algebraic Combinatorics*, 6(2) :547–594, 2023.

[**RI-117**] V. Bonzom, M. Dupuis, F. Girelli et Q. Pan. Local observables in $\mathrm{SU}_q(2)$ lattice gauge theory. *Physical Review D*, 107(2) :026014 (38 pages), 2023.

[**RI-116**] V. Bonzom, M. Dupuis et Q. Pan. Spinor representation of the Hamiltonian constraint in 3D loop quantum gravity with a nonzero cosmological constant. *Physical Review D*, 107(2) :026015 (27 pages), 2023.

[RI-115] V. Bonzom, V. Nador et A. Tanasa. Double scaling limit of multi-matrix models at large D. *Journal of Physics A : Mathematical and Theoretical*, 56(7) :075201 (41 pages), 2023.

[RI-114] V. C. Bui, G. Duchamp, Q. H. Ngo, V. Hoang Ngoc Minh et V. Nguyen Dinh. Towards a Theory of Domains for Harmonic Functions and its Symbolic Counterpart. *Mathematics in Computer Science*, 17 :Article 4 (40 pages), 2023.

[RI-113*] A. Deza et L. Pournin. Sizing the white whale. *Fields Institute Communications*, XX :29 pages, 2023.

[RI-112*] T. Fernique. Density of binary disc packings : lower and upper bounds. *Experimental Mathematics*, XX :12 pages, 2023.

[**RI-111**] T. Fernique. Density of binary disc packings : playing with stoichiometry. *Journal of Combinatorial Theory - Series A*, 194 :105690 (18 pages), 2023.

[RI-110*] J. Kari et V. Lutfalla. Planar Rosa : a family of quasiperiodic substitution discrete plane tilings with 2n-fold rotational symmetry. *Natural Computing*, XX :23 pages, 2023.

[RI-109] J. Kari et V. Lutfalla. Substitution Discrete Plane Tilings with 2n-Fold Rotational Symmetry for Odd n. *Discrete & Computational Geometry*, 69(2) :349–398, 2023.

[RI-108] K. Nurligareev et A. Povolotsky. Watermelons on the half-plane. *Journal of Statistical Mechanics : Theory and Experiment*, 2023(1) :013101 (30 pages), 2023.

[RI-107*] K. A. Penson, K. Gorska, A. Horzela et G. Duchamp. Hausdorff moment problem for combinatorial numbers of Brown and Tutte : exact solution. *Enumerative Combinatorics and Applications*, XX :8 pages, 2023.

[RI-106*] L. Poinsot et H.-E. Porst. The algebra of Feistel-Toffoli construction. *Journal of Algebra and its Applications*, XX :24 pages, 2023.

[**RI-105***] L. Pournin. Local extrema for hypercube sections. *Journal d'Analyse Mathématique*, XX :38 pages, 2023.

[**RI-104***] L. Pournin. Shallow Sections of the Hypercube. *Israel Journal of Mathematics*, XX :20 pages, 2023.

2022 [RI-103] R. C. Avohou, J. Ben Geloun et M. N. Hounkonnou. Extending the Tutte and Bollobás-Riordan polynomials to rank 3 weakly coloured stranded graphs. *Combinatorics, Probability and Computing*, 31(3) :507–549, 2022.

[**RI-102**] F. Bassino, M. Bouvel, M. Drmota, V. Feray, L. Gerin, M. Maazoun et A. Pierrot. Linear-sized independent sets in random cographs and increasing subsequences in separable permutations. *Combinatorial Theory*, 2(3) :15 (35 pages), 2022.

[**RI-101**] F. Bassino, M. Bouvel, V. Féray, L. Gerin, M. Maazoun et A. Pierrot. Random cographs : Brownian graphon limit and asymptotic degree distribution. *Random Structures and Algorithms*, 60(2) :166–200, 2022.

[**RI-100**] F. Bassino, M. Bouvel, V. Féray, L. Gerin, M. Maazoun et A. Pierrot. Scaling limits of permutation classes with a finite specification : a dichotomy. *Advances in Mathematics*, 405 :108513 (84 pages), 2022.

[RI-99] N. Bédaride et T. Fernique. Density of binary disc packings : the nine compact packings. *Discrete & Computational Geometry*, 67 :787–810, 2022.

[RI-98] M. Bendkowski, O. Bodini et S. Dovgal. Tuning as convex optimisation : a polynomial tuner for multi-parametric combinatorial samplers. *Combinatorics, Probability and Computing*, 31(5) :765–811, 2022.

[RI-97] J. Ben Geloun et S. Ramgoolam. All-orders asymptotics of tensor model observables from symmetries of restricted partitions. *Journal of Physics A : Mathematical and Theoretical*, 55(43) :435203 (37 pages), 2022.

[**RI-96**] O. Bodini, A. Genitrini, B. Gittenberger, I. Larcher et M. Naima. Compaction for two models of logarithmic-depth trees : Analysis and experiments. *Random Structures and Algorithms*, 61(1) :31–61, 2022.

[RI-95] O. Bodini, A. Genitrini, C. Mailler et M. Naima. Strict monotonic trees arising from evolutionary processes : Combinatorial and probabilistic study. *Advances in Applied Mathematics*, 133 :102284 (43 pages), 2022.

[**RI-94**] V. Bonzom, G. Chapuy et M. Dołęga. b-monotone Hurwitz numbers : Virasoro constraints, BKP hierarchy, and $O(N)$-BGW integral. *International Mathematical Research Notices*, rnac(177) :59 pages, 2022.

[**RI-93**] V. Bonzom, G. Chapuy et M. Dołęga. Enumeration of non-oriented maps via integrability. *Algebraic Combinatorics*, 5(6) :1363–1390, 2022.

[RI-92] V. Bonzom et E. R. Livine. Self-duality of the $6j$-symbol and Fisher zeros for the Tetrahedron. *Annales de l'Instut Henri Poincaré D : Combinatorics, Physics and their Interactions*, 9(1) :73–118, 2022.

[RI-91] V. Bonzom, V. Nador et A. Tanasa. Double scaling limit for the $O(N)^3$-invariant tensor model. *Journal of Physics A : Mathematical and Theoretical*, 55(13) :135201 (33 pages), 2022.

[RI-90] J. Cardinal, L. Pournin et M. Valencia Pabon. Diameter estimates for graph associahedra. *Annals of Combinatorics*, 26(4) :873–902, 2022.

[RI-89] M. Csikós et N. Mustafa. Optimal approximations made easy. *Information Processing Letters*, 176 :106250 (6 pages), 2022.

[RI-88] C. M. H. de Figueiredo, C. S. R. Patrão, D. Sasaki et M. Valencia Pabon. On total and edge coloring of some Kneser graphs. *Journal of Combinatorial Optimization*, 44 :119–135, 2022.

[RI-87] Q. De Mourgues. A combinatorial approach to Rauzy-type dynamics II : The labelling method and a second proof of the KZB classification theorem. *Discrete and Continuous Dynamical Systems (DCDS)*, 42(7) :3465–3538, 2022.

[RI-86] A. Deza et L. Pournin. A linear optimization oracle for zonotope computation. *Computational Geometry*, 100 :101809 (22 pages), 2022.

[RI-85] A. Deza et L. Pournin. Primitive point packing. *Mathematika*, 68(3) :979–1007, 2022.

[RI-84] K. Nurligareev et I. Reshetnikov. Decompositions of functions defined on finite sets in $\mathbb{R}^d$. *Journal of Knot Theory and its ramifications*, 31(2) :2250011 (16 pages), 2022.

[RI-83] L. Poinsot. Hilbertian Frobenius algebras. *Communications in Algebra*, 50(1) :334–363, 2022.

2021 [RI-82] A. Asinowski, C. Banderier et B. Hackl. Flip-sort and combinatorial aspects of pop-stack sorting. *Discrete Mathematics & Theoretical Computer Science*, 22(2) :Article 4 (39 pages), 2021.

[RI-81] C. Banderier, F. Luca, F. Pappalardi, E. Trevino et C. A. Gòmez Ruiz. On Egyptian fractions of length 3. *Revista de la Union Matematica Argentina*, 62(1) :257–274, 2021.

[RI-80] F. Bassino, C. Nicaud et P. Weil. Statistics of subgroups of the modular group. *International Journal of Algebra and Computation*, 31(8) :1691–1751, 2021.

[RI-79] C. Bean, E. Nadeau et H. Ulfarsson. Enumeration of Permutation Classes and Weighted Labelled Independent Sets. *Discrete Mathematics & Theoretical Computer Science*, 22(2) :2 (38 pages), 2021.

[RI-78] D. Benedetto, E. Caglioti, S. Caracciolo, M. D'Achille, G. Sicuro et A. Sportiello. Random Assignment Problems on 2d Manifolds. *Journal of Statistical Physics*, 183 :Article 34 (40 pages), 2021.

[RI-77] O. Bodini, M. Dien, A. Genitrini et F. Peschanski. Quantitative and Algorithmic aspects of Barrier Synchronization in Concurrency. *Discrete Mathematics & Theoretical Computer Science*, 22(3) :36 pages, 2021.

[RI-76] S. Caracciolo, V. Erba et A. Sportiello. The Number of Optimal Matchings for Euclidean Assignment on the Line. *Journal of Statistical Physics*, 183 :Article 3 (27 pages), 2021.

[RI-75] A. Deza, J.-B. Hiriart-Urruty et L. Pournin. Polytopal balls arising in optimization. *Contributions to Discrete Mathematics*, 16(3) :125–138, 2021.

[RI-74] A. Deza, L. Pournin et R. Rakotonarivo. The vertices of primitive zonotopes. *Contemporary Mathematics : Polytopes and Discrete Geometry*, 764 :71–81, 2021.

[RI-73] T. Fernique. Compact packings of space with two sizes of spheres. *Discrete & Computational Geometry*, 65(4) :1287–1295, 2021.

[RI-72] T. Fernique, A. Hashemi et O. Sizova. Compact packings of the plane with three sizes of discs. *Discrete & Computational Geometry*, 66 :613–635, 2021.

[RI-71] L. N. Grippo, A. Pastine, P. Torres, M. Valencia Pabon et J. C. Vera. On the P_3-hull number of Kneser graphs. *Electronic Journal of Combinatorics*, 28 :P3.32 (9 pages), 2021.

[RI-70] S. Johri, S. Debnath, A. Mocherla, A. Singh, A. Prakash, J. Kim et I. Kerenidis. Nearest centroid classification on a trapped ion quantum computer. *npj Quantum Information*, 7 :122 (11 pages), 2021.

[RI-69] A. Leivaditis, A. Singh, G. Stamoulis, D. M. Thilikos et K. Tsatsanis. Minor obstructions for apex-pseudoforests. *Discrete Mathematics*, 344(10) :112529 (45 pages), 2021.

[RI-68] L. Poinsot et H.-E. Porst. Internal coalgebras in cocomplete categories : Generalizing the Eilenberg-Watts Theorem. *Journal of Algebra and Its Applications*, 20(9) :2150165 (35 pages), 2021.

2020 [RI-67] A. Asinowski, A. Bacher, C. Banderier et B. Gittenberger. Analytic combinatorics of lattice paths with forbidden patterns, the vectorial kernel method, and generating functions for pushdown automata. *Algorithmica*, 82 :386–428, 2020.

[RI-66] R. C. Avohou, J. Ben Geloun et N. DUB. On the counting of $O(N)$ tensor invariants. *Advances in Theoretical and Mathematical Physics*, 24(4) :821–878, 2020.

[**RI-65**] F. Bassino, M. Bouvel, V. Féray, L. Gerin, M. Maazoun et A. Pierrot. Universal limits of substitution-closed permutation classes. *Journal of the European Mathematical Society*, 22(11) :3565–3639, 2020.

[RI-64] N. Bédaride et T. Fernique. Canonical projection tilings defined by patterns. *Geometriae Dedicata*, 208 :157–175, 2020.

[RI-63] O. Bodini, A. Genitrini, B. Gittenberger et S. Wagner. On the number of increasing trees with label repetitions. *Discrete Mathematics*, 343(8) :111722 (12 pages), 2020.

[RI-62] F. Bonomo-Braberman, M. C. Dourado, M. Valencia Pabon et J. C. Vera Lizcano. A note on homomorphisms of Kneser hypergraphs. *Applied Mathematics and Computation*, 366 :124764 (5 pages), 2020.

[RI-61] B. Brešar et M. Valencia Pabon. On the P_3-hull number of Hamming graphs. *Discrete Applied Mathematics*, 282 :48–52, 2020.

[RI-60] S. Caracciolo, M. D'Achille, V. Erba et A. Sportiello. The Dyck bound in the concave 1-dimensional random assignment model. *Journal of Physics A*, 53(6) :064001 (27 pages), 2020.

[**RI-59**] J. David, L. Pournin et R. Rakotonarivo. Elementary moves on lattice polytopes. *Journal of Combinatorial Theory, Series A*, 172 :105200 (35 pages), 2020.

[RI-58] A. Deza, A. Deza, Z. Guan et L. Pournin. Distance between vertices of lattice polytopes. *Optimization Letters*, 14 :309–326, 2020.

[**RI-57**] A. Deza, L. Pournin et N. Sukegawa. The diameter of lattice zonotopes. *Proceedings of the American Mathematical Society*, 148(8) :3507–3516, 2020.

[RI-56] A. Leivaditis, A. Singh, G. Stamoulis, D. M. Thilikos, K. Tsatsanis et V. Velona. Minor-obstructions for apex sub-unicyclic graphs. *Discrete Applied Mathematics*, 284 :538–555, 2020.

[RI-55] L. Poinsot. Hilbertian (function) algebras. *Communications in Algebra*, 48(3) :961–991, 2020.

2019 [RI-54] A. Asinowski, C. Banderier, S. Billey, B. Hackl et S. Linusson. Pop-stack sorting and its image : Permutations with overlapping runs. *Acta Mathematica Univerversitatis Comenianae*, LXXXVIII(3) :395–402, 2019.

[RI-53] C. Banderier, C. Krattenthaler, A. Krinik, D. Kruchinin, V. Kruchinin, D. Nguyen et M. Wallner. Explicit formulas for enumeration of lattice paths : basketball and the kernel method. *Lattice Path Combinatorics and Applications*, volume 58 de *Developments in Mathematics*, pages 78–118. Springer, 2019.

[RI-52] C. Banderier et F. Luca. On the period mod m of polynomially-recursive sequences : a case study. *Journal of Integer Sequences*, 22 :Article 19.5.2 (14 pages), 2019.

[**RI-51**] C. Banderier, P. Marchal et M. Wallner. Periodic Pólya Urns, the Density Method, and Asymptotics of Young Tableaux. *Annals of Probability*, 48(4) :1921–1965, 2019.

[RI-50] C. Banderier et M. Wallner. The kernel method for lattice paths below a line of rational slope. *Lattice Path Combinatorics and Applications*, volume 58 de *Developments in Mathematics*, pages 119–154. Springer, 2019.

[RI-49] M. Bendkowski, O. Bodini et S. Dovgal. Statistical Properties of Lambda Terms. *Electronic Journal of Combinatorics*, 26(4) :P4.1 (70 pages), 2019.

[RI-48] V. Bonzom, V. Nador et A. Tanasa. Diagrammatic proof of the large N melonic dominance in the SYK model. *Letters in Mathematical Physics*, 109(12) :2611–2624, 2019.

[RI-47] V. Bonzom, V. Nador et A. Tanasa. Diagrammatics of the quartic $O(N)^3$-invariant Sachdev-Ye-Kitaev-like tensor model. *Journal of Mathematical Physics*, 69(7) :072302 (25 pages), 2019.

[RI-46] B. Brešar et M. Valencia Pabon. Independence number of products of Kneser graphs. *Discrete Mathematics*, 342 :1017–1027, 2019.

[RI-45] F. Colomo, A. G. Pronko et A. Sportiello. Arctic curve of the free-fermion six-vertex model in an L-shaped domain. *Journal of Statistical Physics*, 174(1) :1–27, 2019.

[RI-44] A. Deza et L. Pournin. Diameter, decomposability, and Minkowski sums of polytopes. *Canadian Mathematical Bulletin*, 62(4) :741–755, 2019.

[RI-43] G. H. Duchamp, V. Hoang Ngoc Minh et Q. H. Ngô. Kleene stars of the plane, polylogarithms and symmetries. *Theoretical Computer Science*, 800 :52–72, 2019.

[**RI-42**] T. Fernique et M. Sablik. Weak colored local rules for planar tilings. *Ergodic Theory and Dynamical Systems*, 39(12) :3322–3346, 2019. 30 pages.

[RI-41] L. Poinsot. Topological rigidity as a monoidal equivalence. *Communications in Algebra*, 47(9) :3457–3480, 2019.

[RI-40] L. Pournin. Eccentricities in the flip-graphs of convex polygons. *Journal of Graph Theory*, 92(2) :111–129, 2019.

2018 [RI-39] L. Alcon, F. Bonomo, G. Duran, M. Gutierrez, M. P. Mazzoleni, B. Ries et M. Valencia Pabon. On the Bend number of circular-arc graphs as edge intersection graphs of paths on a grid. *Discrete Applied Mathematics*, 234 :12–21, 2018.

[**RI-38**] F. Bassino, M. Bouvel, V. Féray, L. Gerin et A. Pierrot. The Brownian limit of separable permutations. *Annals of Probability*, 46(4) :2134–2189, 2018.

[RI-37] J. Ben Geloun, A. Kegeles et A. Pithis. Minimizers of the dynamical Boulatov model. *The European Physical Journal C*, 78 :Article 996 (12 pages), 2018.

[**RI-36**] J. Ben Geloun, T. A. Koslowski, D. Oriti et A. D. Pereira. Functional Renormalization Group analysis of rank 3 tensorial group field theory : The full quartic invariant truncation. *Physical Review D*, 97 :126018 (12 pages), 2018.

[RI-35] J. Ben Geloun et V. Rivasseau. A Renormalizable SYK-type Tensor Field Theory. *Annales de l'Institut Henri Poincaré D : Combinatorics, Physics and Their Interactions*, 19 :3357–3395, 2018.

[RI-34] J. Ben Geloun et R. Toriumi. Renormalizable enhanced tensor field theory : The quartic melonic case. *Journal of Mathematical Physics*, 59(11) :112303 (52 pages), 2018.

[RI-33] O. Bodini, D. Gardy, B. Gittenberger et Z. Gołębiewski. On the Number of Unary-Binary Tree-Like Structures with Restrictions on the Unary Height. *Annals of Combinatorics*, 22 :45–91, 2018.

[RI-32] O. Bodini, B. Gittenberger et Z. Gołębiewski. Enumerating lambda terms by weighted length of their De Bruijn representation. *Discrete Applied Mathematics*, 239 :45–61, 2018.

[RI-31] F. Bonomo, G. Duran, I. Koch et M. Valencia Pabon. On the (k, i)-coloring of cacti and complete graphs. *Ars Combinatoria*, 137 :317–333, 2018.

[RI-30] F. Bonomo, I. Koch, P. Torres et M. Valencia Pabon. k-tuple colorings of the Cartesian product of graphs. *Discrete Applied Mathematics*, 245 :177–182, 2018.

[RI-29] V. Bonzom et S. Dartois. Blobbed topological recursion for the quartic melonic tensor model. *Journal of Physics A : Mathematical and Theoretical*, 51(32) :325201 (45 pages), 2018.

[RI-28] V. C. Bui, G. H. Duchamp, Q. H. Ngô, V. Hoang Ngoc Minh et C. Tollu. (Pure) transcendence bases in φ-deformed bialgebra. *Séminaire Lotharingien de Combinatoire*, 74 :B74f (22 pages), 2018.

[RI-27] A. Deza et L. Pournin. Improved bounds on the diameter of lattice polytopes. *Acta Mathematica Hungarica*, 154(2) :457–469, 2018.

[RI-26] C. Lavault. Integral representations and asymptotic behaviour of a Mittag-Leffler type function of two variables. *Advances in Operator Theory*, 3(2) :365–373, 2018.

[RI-25] H. Parlier et L. Pournin. Modular flip-graphs of one holed surfaces. *European Journal of Combinatorics*, 67 :158–173, 2018.

[RI-24] H. Parlier et L. Pournin. Once punctured disks, non-convex polygons, and pointihedra. *Annals of Combinatorics*, 22(3) :619–640, 2018.

[RI-23] L. Poinsot. The solution to the embedding problem of a (differential) Lie algebra into its Wronskian envelope. *Communications in Algebra*, 46(4) :1641–1667, 2018.

2017 [**RI-22**] A. Bacher, O. Bodini, H.-K. Hwang et T.-H. Tsai. Generating Random Permutations by Coin-Tossing : Classical Algorithms, New Analysis and Modern Implementation. *ACM Transactions on Algorithms*, 13(2) :Article 24 (43 pages), 2017.

[RI-21] A. Bacher, O. Bodini et A. Jacquot. Efficient random sampling of binary and unary-binary trees via holonomic equations. *Theoretical Computer Science*, 695 :42–53, 2017.

[RI-20] C. Banderier, J.-L. Baril et C. Moreira Dos Santos. Right-jumps and pattern avoiding permutations. *Discrete Mathematics & Theoretical Computer Science*, 18(2) :12 (17 pages), 2017.

[RI-19] C. Banderier et M. Wallner. Lattice paths with catastrophes. *Discrete Mathematics & Theoretical Computer Science*, 19(1) :23 (32 pages), 2017.

[RI-18] F. Bassino, M. Bouvel, A. Pierrot, C. Pivoteau et D. Rossin. An algorithm for computing combinatorial specifications of permutation classes. *Discrete Applied Mathematics*, 224 :16–44, 2017.

[**RI-17**] N. Bédaride et T. Fernique. Weak local rules for planar octagonal tilings. *Israel Journal of Mathematics*, 222 :63–89, 2017.

[RI-16] J. Ben Geloun et F. Caravelli. On an ordering-dependent generalization of Tutte polynomial. *Journal of Statistical Physics*, 168 :1105–1124, 2017.

[**RI-15**] J. Ben Geloun et S. Ramgoolam. Tensor Models, Kronecker coefficients and Permutation Centralizer Algebras. *Journal of High Energy Physics*, 2017 :Article 92 (75 pages), 2017.

[RI-14] O. Bodini, A. Genitrini et N. Rolin. Extended boxed product and application to synchronized trees. *Electronic Notes Discrete Mathematics*, 59 :189–202, 2017.

[RI-13] R. Bondesan, S. Caracciolo et A. Sportiello. Critical behaviour of spanning forests on random planar graphs. *Journal of Physics A : Mathematical and Theoretical*, 50(7) :074003 (51 pages), 2017.

[RI-12] V. Bonzom et L. Lionni. Counting gluings of octahedra. *The Electronic Journal of Combinatorics*, 24(3) :P3.36 (29 pages), 2017.

[RI-11] V. Bonzom, L. Lionni et V. Rivasseau. Colored Triangulations of Arbitrary Dimensions are Stuffed Walsh Maps. *Electronic Journal of Combinatorics*, 24(1) :P1.56 (46 pages), 2017.

[RI-10] V. Bonzom, L. Lionni et A. Tanasa. Diagrammatics of a colored SYK model and of an SYK-like tensor model, leading and next-to-leading orders. *Journal of Mathematical Physics*, 58(5) :052301 (21 pages), 2017.

[RI-9] V. C. Bui, G. H. Duchamp et V. Hoang Ngoc Minh. Structure of Polyzetas and Explicit Representation on Transcendence Bases of Shuffle and Stuffle Algebras. *Journal of Symbolic Computation*, 83 :93–111, 2017.

[RI-8] S. Caracciolo, A. D. Sokal et A. Sportiello. Spanning forests and $OSP(N|2M)$-invariant σ-models. *Journal of Physics A : Mathematical and Theoretical*, 50(11) :114001 (52 pages), 2017.

[RI-7] G. H. Duchamp, J.-Y. Enjalbert, V. Hoang Ngoc Minh et C. Tollu. The mechanics of shuffle products and their siblings. *Discrete Mathematics*, 340 :2286–2300, 2017.

[RI-6] G. H. Duchamp, V. Hoang Ngoc Minh et Q. H. Ngo. Harmonic sums and polylogarithms at non-positive multiple-indices. *Journal of Symbolic Computation*, 83 :166–186, 2017.

[RI-5] G. H. Duchamp, V. Hoang Ngoc Minh, Q. H. Ngo, K. Penson et P. Simonnet. Mathematical renormalization in quantum electrodynamics via noncommutative generating series. *Applications of Computer Algebra*, volume 198 de *Springer Proceedings in Mathematics & Statistics (PROMS)*, pages 59–100, 2017.

[**RI-4**] H. Parlier et L. Pournin. Flip-graph moduli spaces of filling surfaces. *Journal of the European Mathematical Society*, 19(9) :2697–2737, 2017.

[RI-3] L. Poinsot. Lipschitz groups and Lipschitz maps. *International Journal of Group Theory*, 6(1) :9–16, 2017.

[**RI-2**] L. Pournin. The asymptotic diameter of cyclohedra. *Israel Journal of Mathematics*, 219(2) :609–635, 2017.

[RI-1] P. Torres et M. Valencia Pabon. Shifts of the stable Kneser graphs and hom-idempotence. *European Journal of Combinatorics*, 62 :50–57, 2017.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2022 [CI-31] A. Akhavi, F. Paccaut et B. Vallée. Building Sources of Zero Entropy : Rescaling and Inserting Delays. *The 33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022), Philadelphia, PA, United States, June 20-24, 2022*, volume 225 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 1 :1–1 :28, 2022.

[CI-30] O. Bodini, A. Genitrini et M. Naima. A Combinatorial Link Between Labelled Graphs and Increasingly Labelled Schröder Trees. *The 15th Latin American Symposium in Theoretical Informatics (LATIN 2022), Guanajuato, Mexico, November 7-11, 2022*, volume 13568 de *Lecture Notes in Computer Science*, pages 493–509, 2022.

[CI-29] T. Fernique et C. Porrier. A General approach to Ammann bars for aperiodic tilings. *The 15th Latin American Symposium on Theoretical Informatics (LATIN 2022), Guanjuato, Mexico, November 7-11, 2022*, volume 13568 de *Lecture Notes in Computer Science*, pages 574–589, 2022.

[CI-28] D. Pchelina, N. Schabanel, S. Seki et G. Theyssier. Oritatami systems assemble shapes no less complex than tile assembly model (ATAM). *The 39th International Symposium on Theoretical Aspects of Computer Science (STACS 2022), Marseille, France, March 15-18, 2022*, volume 219 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 51 :1–51 :23, 2022.

2021 [CI-27] D. F. Andriantsiory, J. Ben Geloun et M. Lebbah. Multi-Slice Clustering for 3-order Tensor Data. *The 20th IEEE International Conference on Machine Learning and Applications (ICMLA 2021), Pasadena, CA, USA, December 13-16, 2021*, pages 173–178, 2021.

[**CI-26**] C. Banderier et M. Wallner. Young tableaux with periodic walls : counting with the density method. *The 33rd Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2021), Ramat-Gan, Israel, January 10–13, 2022*, volume 85B de *Séminaire Lotharingien de Combinatoire*, pages 47 : 1–12, 2021.

[CI-25] J. Cardinal, L. Pournin et M. Valencia Pabon. Bounds on the diameter of graph associahedra. *The XI Latin and American Algorithms, Graphs and Optimization Symposium (LAGOS 2021), São Paulo, Brazil, May 17-21 2021*, volume 195 de *Procedia Computer Science*, pages 239–247, 2021.

[CI-24] D. Castonguay, C. M. de Figueiredo, L. A. Kowada, C. S. Patrão, D. Sasaki et M. Valencia Pabon. On total coloring the direct product of complete graphs. *The XI Latin and American Algorithms, Graphs and Optimization Symposium (LAGOS 2021), São Paulo, Brazil, May 17-21 2021*, volume 195 de *Procedia Computer Science*, pages 306–314, 2021.

[CI-23] V. Lutfalla. An Effective Construction for Cut-And-Project Rhombus Tilings with Global n-Fold Rotational Symmetry. *27th IFIP WG 1.5 International Workshop on Cellular Automata and Discrete Complex Systems (AUTOMATA 2021), Aix-Marseille, France, July 12-14, 2021*, volume 90 de *Open Access Series in Informatics (OASIcs)*, pages 9 :1–9 :12, 2021.

[CI-22] A. Pastine, P. Torres et M. Valencia Pabon. On the diameter of Schrijver graphs. *The XI Latin and American Algorithms, Graphs and Optimization Symposium (LAGOS 2021), São Paulo, Brazil, May 17-21 2021*, volume 195 de *Procedia Computer Science*, pages 266–274, 2021.

[CI-21] D. Pchelina, N. Schabanel, S. Seki et Y. Ubukata. Simple Intrinsic Simulation of Cellular Automata in Oritatami Molecular Folding Model. *The 14th Latin American Symposium on Theoretical Informatics (LATIN 2020), São Paulo, Brazil, January 5-8, 2021*, volume 12118 de *Lecture Notes in Computer Science*, pages 425–436, 2021.

2020 [CI-20] C. Banderier et A. Asinowski. On Lattice Paths with Marked Patterns : Generating Functions and Multivariate Gaussian Distribution. *The 31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2020), Klagenfurt, Austria, June 15–19, 2020*, volume 159 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 1 :1–16, 2020.

[**CI-19**] C. Banderier, A. Asinowski et V. Roitner. Generating functions for lattice paths with several forbidden patterns. *The 32nd Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2020), Online July 6-24, 2020*, volume 84B de *Seminaire Lotharingien de Combinatoire*, pages 95 : 1–12, 2020.

[CI-18] C. Banderier et M. Goldwurm. Number of Prefixes in Trace Monoids : Clique Polynomials and Dependency Graphs. *The 16th Conference on Computability in Europe (CiE 2020), Salerno, Italy, June 29-July 2, 2020*, volume 12098 de *Lecture Notes in Computer Science*, pages 251–263. Springer, 2020.

[CI-17] F. Bassino, T. Rakotoarimalala et A. Sportiello. The Complexity of the Approximate Multiple Pattern Matching Problem for Random Strings. *The 31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2020), Klagenfurt, Austria, June 15-19, 2020*, volume 159 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 3 :1–15, 2020.

[CI-16] J. Ben Geloun, C. Coti et A. D. Malony. On-the-fly Optimization of Parallel Computation of Symbolic Symplectic Invariants. *The 19th International Symposium on Parallel and Distributed Computing (ISPDC 2020), Warsaw, Poland, July 5-8, 2020*, pages 102–109, 2020.

[CI-15] O. Bodini, A. Genitrini, M. Naima et A. Singh. Families of Monotonic Trees : Combinatorial Enumeration and Asymptotics. *The 15th International Computer Science Symposium in Russia (CSR 2020), Yekaterinburg, Russia, June 29 - July 3, 2020*, volume 12159 de *Lecture Notes in Computer Science*, pages 155–168, 2020.

[CI-14] T. Fernique et I. Galanov. Local growth of planar rhombus tilings. *The 14th International Conference on Quasicrystals (ICQ14), Kranjska Gora, Slovenia, 26–30 May 2019*, volume 1458 de *Journal of Physics : Conference Series*, page 012001 (4 pages), 2020.

2019 [CI-13] O. Bodini, M. Dien, A. Genitrini et F. Peschanski. The Combinatorics of Barrier Synchronization. *The 40th International Conference on Application and Theory of Petri Nets and Concurrency (PETRI NETS 2019), Aachen, Germany, June 23-28, 2019*, volume 11522 de *Lecture Notes in Computer Science*, pages 386–405, 2019.

[CI-12] T. Fernique, A. Hashemi et O. Sizova. Compact packings of the plane with three sizes of discs. *The 21th IAPR International Conference on Discrete Geometry for Computer Imagery (DGCI 2019), Marne-la-Vallée, France, March 26-28, 2019*, volume 11414 de *Lecture Notes in Computer Science*, pages 420–431, 2019.

2018 [CI-11] A. Asinowski, A. Bacher, C. Banderier et B. Gittenberger. Analytic Combinatorics of Lattice Paths with Forbidden Patterns : Asymptotics Aspects and Borges's Theorem. *The 29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018), Uppsala, Sweden, June 25-29, 2018*, volume 110 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10 :1–13, 2018.

[CI-10] A. Asinowski, A. Bacher, C. Banderier et B. Gittenberger. Analytic Combinatorics of Lattice Paths with Forbidden Patterns : Enumerative Aspects. *The 12th International Conference on Language and Automata Theory and Applications (LATA 2018), Bar-Ilan, Israel, April 9-11, 2018*, volume 10792 de *Lecture Notes in Computer Science*, pages 195–206, 2018.

[CI-9] C. Banderier, P. Marchal et M. Wallner. Periodic Pólya urns and an application to Young tableaux. *The 29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018), , Uppsala, Sweden, June 25-29, 2018*, volume 110 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 11 :1–13, 2018.

[CI-8] O. Bodini, J. Courtiel, S. Dovgal et H.-K. Hwang. Asymptotic Distribution of Parameters in Random Maps. *The 29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018), Uppsala, Sweden, June 25-29, 2018*, volume 110 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 13 :1–13 :12, 2018.

[CI-7] O. Bodini, M. Dien, A. Genitrini et A. Viola. Beyond Series-Parallel Concurrent Systems : The Case of Arch Processes. *The 29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018), Uppsala, Sweden, June 25-29, 2018, Uppsala, Sweden*, volume 110 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 14 :1–14 :14, 2018.

[CI-6] J. David, R. Grappe, M. Lacroix et E. Traversi. Self-sufficient sets in smartgrids. *The Joint EURO/ALIO International Conference 2018 on Applied Combinatorial Optimization (EURO/ALIO 2018), Viña del Mar, Chile, April 11–13, 2022*, volume 69 de *Electronic Notes in Discrete Mathematics*, pages 301–308, 2018.

[CI-5] G. H. Duchamp, V. Hoang Ngoc Minh et K. A. Penson. About Some Drinfel'd Associators. *The 20th International Workshop Computer Algebra in Scientific Computing (CASC 2018), Lille, France, September 17-21, 2018*, volume 11077 de *Lecture Notes in Computer Science*, pages 146–163, 2018.

2017 [CI-4] O. Bodini, M. Dien, A. Genitrini et F. Peschanski. Entropic Uniform Sampling of Linear Extensions in Series-Parallel Posets. *The 12th International Computer Science Symposium in Russia (CSR 2017), Kazan, Russia, June 8-12, 2017*, volume 10304 de *Lecture Notes in Computer Science*, pages 71–84, 2017.

[CI-3] O. Bodini et P. Tarau. On Uniquely Closable and Uniquely Typable Skeletons of Lambda Terms. *The 27th International Symposium on Logic-Based Program Synthesis and Transformation (LOPSTR 2017), Namur, Belgium, October 10-12, 2017*, volume 10855 de *Lecture Notes in Computer Science*, pages 252–268, 2017.

[CI-2] T. Monteil. A universal oracle for signal machines. *The 13th Conference on Computability in Europe 2017 (CiE 2017), Turku, Finland, June 12-16, 2017*, volume 10307 de *Lecture Notes in Computer Science*, pages 315–326, 2017.

[CI-1] S. M. Sedjelmaci. Two fast parallel GCD algorithms of many integers. *The 42nd ACM International Symposium on Symbolic and Algebraic Computation (ISSAC 2017), Kaiserslautern, Germany, July 25-28, 2017*, pages 397–404, 2017.

Autres articles dans des conférences ou workshops à comité de lecture

2022 [CO-15] O. Bodini. On the strange kinetic aesthetic of rectangular shape partitions. *The 12th edition of the conference GASCom on random generation of combinatorial structures, (GASCom 2022), Varese, Italy, June 13-15 2022*, volume 30 de *Pure Mathematics and Applications*, pages 37–44, 2022.

[CO-14] O. Bodini, A. Singh et N. Zeilberger. Sampling β-normal linear λ-terms. *The 12th edition of the conference GASCom on random generation of combinatorial structures (GASCom 2022), Varese, Italy, June 13-15 2022*, volume 30 de *Pure Mathematics and Applications*, pages 45–55, 2022.

2021 [CO-13] T. Monteil et K. Nurligareev. Asymptotics for Connected Graphs and Irreducible Tournaments. *The 2021 European Conference on Combinatorics, Graph Theory and Applications (EuroComb 2021), Barcelona, Spain, September 6-10, 2021*, volume 14 de *Trends in Mathematics*, pages 823–828, 2021.

2020 [CO-12] J. Ben Geloun. On the counting tensor model observables as $U(N)$ and $O(N)$ classical invariants. *The Corfu Summer Institute 2019 School and Workshops on Elementary Particle Physics and Gravity (CORFU2019), Corfù, Greece, August 31-September 25, 2020*, volume 376 de *Proceedings of Science*, pages 175–175, 2020.

2019 [CO-11] O. Bodini, A. Genitrini et M. Naima. Ranked Schröder Trees. *The 16th SIAM Workshop on Analytic Algorithmics and Combinatorics (ANALCO 2019), San Diego, CA, USA, January 6, 2019*, pages 13–26, 2019.

2018 [CO-10] A. Bacher, O. Bodini, A. Hollender et J. Lumbroso. MergeShuffle : a very fast, parallel random permutation algorithm. *The 11th International Conference on Random and Exhaustive Generation of Combinatorial Structures (GASCom 2018), Athens, Greece, June 18-20, 2018*, volume 2113 de *CEUR Workshop Proceedings*, pages 43–52, 2018.

[CO-9] C. Banderier, P. Marchal et M. Wallner. Rectangular Young tableaux with local decreases and the density method for uniform random generation (short version). *The 11th International Conference on Random and Exhaustive Generation of Combinatorial Structures (GASCom 2018), Athens, Greece, June 18-20, 2018*, volume 2113 de *CEUR Workshop Proceedings*, pages 60–68, 2018.

[CO-8] C. Banderier et M. Wallner. Local time for lattice paths and the associated limit laws. *The 11th International Conference on Random and Exhaustive Generation of Combinatorial Structures (GASCom 2018), Athens, Greece, June 18-20, 2018*, volume 2113 de *CEUR Workshop Proceedings*, pages 69–78, 2018.

[CO-7] F. Bassino, T. Rakotoarimalala et A. Sportiello. The complexity of the Multiple Pattern Matching Problem for random strings. *The 15th SIAM Workshop on Analytic Algorithmics and Combinatorics (ANALCO 2018), New Orleans, LA, USA, January 8-9, 2018*, pages 40–53, 2018.

[CO-6] F. Bassino et A. Sportiello. Linear-time exact sampling of sum-constrained random variables. *The 11th International Conference on Random and Exhaustive Generation of Combinatorial Structures (GASCom 2018), Athens, Greece, June 18-10, 2018*, volume 2113 de *CEUR Workshop Proceedings*, pages 94–105, 2018.

[CO-5] M. Bendkowski, O. Bodini et S. Dovgal. Polynomial tuning of multiparametric combinatorial samplers. *The 15th SIAM Workshop on Analytic Algorithmics and Combinatorics (ANALCO 2018), New Orleans, LA, USA, January 8-9, 2018*, pages 92–106, 2018.

[CO-4] J. David, L. Pournin et R. Rakotonarivo. A Markov chain for lattice polytopes. *The 11th International Conference on Random and Exhaustive Generation of Combinatorial Structures (GASCom 2018), Athens, Greece, June 18-20, 2018*, volume 2113 de *CEUR Workshop Proceedings*, pages 132–139, Athens, Greece, 2018.

2017 [CO-3] C. Banderier et M. Wallner. Lattice paths with catastrophes. *The 10th International Conference on Random and Exhaustive Generation of Combinatorial Structures (GASCom'16), La Marana, Corsica, France, June 1- 4, 2016*, volume 59 de *Electronic Notes in Discrete Mathematics*, pages 131–146, 2017.

[CO-2] O. Bodini, M. Dien, A. Genitrini et F. Peschanski. The Ordered and Colored Products in Analytic Combinatorics : Application to the Quantitative Study of Synchronizations in Concurrent Processes. C. Martínez et M. D. Ward, éditeurs, *The 14th SIAM Workshop on Analytic Algorithmics and Combinatorics (ANALCO 2017), Barcelona, Spain, January 16-17, 2017*, pages 16–30, 2017.

[CO-1] A. Deza et L. Pournin. A survey on the diameter of lattice polytopes. *Discrete Geometry, Oberwolfach, Germany, April 9-15, 2017*, volume 19/2017 de *Oberwolfach Reports*, pages 1176–1179, 2017.

Direction d'ouvrages (DO)

Livres

2022 [**LI-2**] N. H. Mustafa. *Sampling in Combinatorial and Geometric Set Systems*, volume 265 de *Mathematical Surveys and Monographs*. American Mathematical Society, 2022.

2020 [LI-1] F. Bassino, I. Kapovich, M. Lohrey, A. Miasnikov, C. Nicaud, A. Nikolaev, I. Rivin, V. Shpilrain, A. Ushakov et P. Weil. *Complexity and Randomness in Group Theory*. de Gruyter, 2020.

Édition d'ouvrages collectifs

[ED-2] F. Bassino, F. Bonomo-Braberman, L. Pournin et M. Valencia Pabon. *Selected papers from LAGOS'17 : IX Latin and American Algorithms, Graphs and Optimization Symposium, Marseille, France, September 11-15, 2017*, volume 281 de *Discrete Applied Mathematics*. Elsevier, 2020.

2017 [ED-1] F. Bassino, F. Bonomo, L. Pournin, M. Valencia Pabon et J. VERA. *The IX Latin-American Algorithms, Graphs, and Optimization Symposium (LAGOS 2017), Marseille, France, September 11-15, 2017*, volume 62 de *Electronic Notes in Discrete Mathematics*. Elsevier, 2017.

Logiciels et Ressources (LO)

2019 [LO-1] M. Bendkowski, O. Bodini et S. Dovgal. Boltzmann-Brain : A standalone application for random generation of combinatorial structures. https://github.com/maciej-bendkowski/boltzmann-brain, 2019.

Thèses et Habilitations (TH)

Thèses

2023 [TH-12] C. Chinaud-Chaix. Assemblage de supercristaux par approche prédictive. Thèse, Toulouse 3, 2023.

2022 [TH-11] K. Nurligareev. Irreducibility of combinatorial objects : asymptotic probability and interpretation. Thèse, Université Sorbonne Paris Nord, 2022.

[TH-10] A. Singh. Une approche unifiée de la combinatoire du lambda-calcul et des cartes : bijections et propriétés limites. Thèse, Université Sorbonne Paris Nord, 2022.

2021 [TH-9] N. Dub. Enumeration of triangulations modulo symmetries and of rooted triangulations counted by their number of (d - 2)-simplices in dimension $d \geq 2$. Thèse, Université Sorbonne Paris Nord, 2021.

[TH-8] V. Lutfalla. Plans discrets substitutifs. Thèse, Université Sorbonne Paris Nord, 2021.

2020 [TH-7] M. Naima. Quantitative Analysis of Directed Acyclic Graphs Modelizing Concurrency Theory or Boolean Functions. Thèse, Université Sorbonne Paris Nord, 2020.

[TH-6] N. Pardal. Caractérisation structurelle de quelques problèmes dans les graphes de cordes et d'intervalles. Thèse, Université Sorbonne Paris Nord, 2020.

2019 [TH-5] S. Dovgal. Interdisciplinary image of Analytic Combinatorics. Thèse, Université Sorbonne Paris Nord, 2019.

[TH-4] I. Galanov. Sur l"auto-assemblage de pavages octogonaux plans de type fini. Thèse, Université Sorbonne Paris Nord, 2019.

[TH-3] T. Rakotoarimalala. Complexité de la recherche de motifs dans un texte aléatoire. Thèse, Université Sorbonne Paris Nord, 2019.

[TH-2] R. Rakotonarivo. Sur la combinatoire des polytopes en nombres entiers, génération aléatoire et enumération. Thèse, Université Sorbonne Paris Nord, 2019.

2017 [TH-1] Q. De Mourgues. Une approche combinatoire pour les dynamiques de type Rauzy. Thèse, Université Sorbonne Paris Cité, 2017.

Habilitations

2023 [HR-3] J. Ben Geloun. Tensor Field Theories : Renormalization, Combinatorics and Algebras. HdR, Université Sorbonne Paris Nord, 2023.

[HR-2] V. Bonzom. Structural and Enumerative Aspects of Discrete Surfaces and PL-Manifolds. HdR, Université Sorbonne Paris Nord, 2023.

2017 [HR-1] A. Bostan. Calcul Formel pour la Combinatoire des Marches. HdR, Université Sorbonne Paris Cité, 2017.

Publications antérieures à l'arrivée au LIPN

Les références comportent « ↑ » et correspondent à des publications antérieures à l'arrivée au LIPN, non référencées par ailleurs.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2022 [↑RI-13] B. Jartoux et N. H. Mustafa. A Tight Analysis of Geometric Local Search. *Discrete & Computational Geometry*, 67(2) :361–379, 2022.

2021 [↑RI-12] J. De Loera, T. Hogan, F. Meunier et N. Mustafa. Tverberg theorems over discrete sets of points. *Polytopes and Discrete Geometry*, 764 :57–70, 2021.

2020 [↑RI-11] K. A. Adiprasito, I. Bárány, N. Mustafa et T. Terpai. Theorems of Carathéodory, Helly, and Tverberg Without Dimension. *Discrete & Computational Geometry*, 64(2) :233–258, 2020.

[↑RI-10] I. Bárány et N. Mustafa. An application of the universality theorem for Tverberg partitions to data depth and hitting convex sets. *Computational Geometry*, 90 :101649 (5 pages), 2020.

2019 [↑RI-9] M. Csikós, N. Mustafa et A. Kupavskii. Tight Lower Bounds on the VC-dimension of Geometric Set Systems. *Journal of Machine Learning Research*, 20 :81 :1–81 :8, 2019.

[↑RI-8] J. De Loera, X. Goaoc, F. Meunier et N. Mustafa. The Discrete Yet Ubiquitous Theorems of Carathéodory, Helly, Sperner, Tucker and Tverberg. *Bulletin of the American Mathematical Society*, 56 :415–511, 2019.

[↑RI-7] K. Dutta, A. Ghosh, B. Jartoux et N. Mustafa. Shallow Packings, Semialgebraic Set Systems, Macbeath Regions, and Polynomial Partitioning. *Discrete & Computational Geometry*, 61(4) :756–777, 2019.

[↑RI-6] A. Kupavskii, N. Mustafa et K. Swanepoel. Bounding the Size of an Almost-Equidistant Set in Euclidean Space. *Combinatorics, Probability and Computing*, 28(2) :280–286, 2019.

[↑RI-5] N. Mustafa et S. Ray. On a Problem of Danzer. *Combinatorics, Probability and Computing*, 28(3) :473–482, 2019.

2018 [↑RI-4] N. Bus, N. Mustafa et S. Ray. Practical and efficient algorithms for the geometric hitting set problem. *Discrete Applied Mathematics*, 240 :25–32, 2018.

[↑RI-3] N. Mustafa, K. Dutta et A. Ghosh. A Simple Proof of Optimal Epsilon Nets. *Combinatorica*, 38(5) :1269–1277, 2018.

2017 [↑RI-2] N. Bus, S. Garg, N. Mustafa et S. Ray. Limits of Local Search : Quality and Efficiency. *Discrete & Computational Geometry*, 57(3) :607–624, 2017.

[↑RI-1] N. Mustafa et S. Ray. ε-Mnets : Hitting Geometric Set Systems with Subsets. *Discrete & Computational Geometry*, 57(3) :625–640, 2017.

Chapitres de livre

[↑CL-1] N. Mustafa et K. Varadarajan. *Epsilon-approximations and Epsilon-nets*, *Handbook of Discrete Computional Geometry*, pages 1241–1267. CRC Press LLC, 2017.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2021 [↑CI-10] M. Csikós et N. Mustafa. Escaping the Curse of Spatial Partitioning : Matchings with Low Crossing Numbers and Their Applications. *37th International Symposium on Computational Geometry (SoCG 2021), Buffalo, NY, USA, June 7-11, 2021*, volume 189 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 28 :1–28 :17, 2021.

2019 [↑CI-9] K. A. Adiprasito, I. Bárány et N. Mustafa. Theorems of Carathéodory, Helly, and Tverberg without Dimension. *13th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2019), San Diego, California, USA, January 6-9, 2019*, pages 2350–2360, 2019.

[↑CI-8] A. Akhavi, J. Clément, D. Darthenay, L. Lhote et B. Vallée. Dichotomic Selection on Words : A Probabilistic Analysis. *30th Annual Symposium on Combinatorial Pattern Matching (CPM 2019), Pisa, Italy, June 18-20, 2019*, volume 128 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 19 :1–19 :19, 2019.

[↑CI-7] C.-C. Huang, M. Mari, C. Mathieu, J. S. B. Mitchell et N. Mustafa. Maximizing Covered Area in the Euclidean Plane with Connectivity Constraint. *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, (APPROX/RANDOM 2019), Cambridge, MA, USA, September 20-22, 2019,*, volume 145 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 32 :1–32 :21, 2019.

[↑CI-6] N. Mustafa. Computing Optimal Epsilon-Nets Is as Easy as Finding an Unhit Set. *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019), Patras, Greece, 8-12 July 2019*, volume 132 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 87 :1–87 :12, 2019.

2018 [↑CI-5] B. Jartoux et N. Mustafa. Optimality of Geometric Local Search. *34th International Symposium on Computational Geometry (SoCG 2018), Budapest, Hungary, June 11-14, 2018*, volume 99 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 48 :1–48 :15, 2018.

[↑CI-4] M. Maria, N. Mustafa, T. Bardoux, J. Defaye et V. Biri. Visibility based WSPD for Global Illumination. *13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018), Funchal, Madeira - portugal, January 27-29 , 2018*, pages 81–90. SciTePress, 2018.

[↑CI-3] N. Mustafa et S. Ray. On a Problem of Danzer. *26th European Symposium on Algorithms (ESA 2018), Helsinki, Finland, August 20–22, 2018*, volume 112 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 64 :1–64 :8, 2018.

2017 [↑CI-2] D. Antunes, C. Mathieu et N. Mustafa. Combinatorics of Local Search : An Optimal 4-Local Hall's Theorem for Planar Graphs. *25th Annual European Symposium on Algorithms (ESA 2017), Vienna, Austria, September 4-6, 2017*, volume 87 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 8 :1–8 :13, 2017.

[↑CI-1] K. Dutta, A. Ghosh, B. Jartoux et N. Mustafa. Shallow Packings, Semialgebraic Set Systems, Macbeath Regions and Polynomial Partitioning. *33rd International Symposium on Computational Geometry (SoCG 2017) , Brisbane, Australia, July 4-7, 2017*, volume 77 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 38 :1–38 :15, 2017.

Édition d'ouvrages collectifs

2019 [↑ED-1] M. Couprie, J. Cousty, Y. Kenmochi et N. Mustafa. *21st IAPR International Conference on Discrete Geometry for Computer Imagery (DGCI 2019)*, volume 11414 de *Lecture Notes in Computer Science*. Springer, 2019.

3.3 Domaine 4. Inscription des activités de recherche dans la société

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique.

Les recherches menées par l'équipe sont de nature fondamentale et académique et ne se prêtent pas vraiment aux interactions avec le monde non-académique. Notons toutefois que la méthode de Boltzmann tend à se propager dans le milieu des ingénieurs, leur permettant d'avoir des outils de white-box testing (des bugs dans certains compilateurs ont ainsi été observés via une telle approche de génération aléatoire de code).

Sur un autre plan d'interactions avec la société, mentionnons l'activité d'art numérique de Olivier Bodini. Ses travaux ouvrent une porte vers la société et l'art via l'art génératif dont il est un représentant reconnu : Olivier a ainsi été sélectionné par le stand FxHash pour représenter l'art génératif à Art Basel Paris en 2023 et ses tableaux, générés aléatoirement, sont exposés dans le monde entier (Beijing Art Exhibition "Colours of Joy", NFTShow Europe à Valence, etc.).

3.3.2 Référence 3. L'équipe partage ses connaissances avec le grand public et intervient dans des débats de société.

Les membres de l'équipe CALIN participent régulièrement à des action de partage de connaissance et de médiation scientifique. Nous avons par exemple donné des exposés généralistes à la journée de la recherche de l'institut Galilée qui a rassemblé tous les laboratoires de l'institut Galilée le 22 novembre 2018 ou à l'université de Klagenfurt le 30 juillet 2019 ou lors d'une conférence à destination du grand public au CIRM à Marseille le 5 novembre 2019. Nous avons également participé à l'accueil et à l'accompagnement de stages de découverte de la recherche d'une semaine à destination des collégiens en novembre 2021 et janvier/février 2023. Dans le cadre des journées "Savante Banlieue" organisées chaque année par l'université Paris 13, et qui accueillent environ 4000 visiteurs (collégiens/lycéens/étudiants) sur 2 jours, nous avons coanimé le stand de l'association Science Ouverte. Nous y avons donné une conférence de vulgarisation scientifique en octobre 2019. Nous avons aussi coanimé des journées de présentation de stages Maths Pour Tous (pour le niveau collège–lycée) et Hippocampe (pour le niveau licence) à l'université de Marseille en 2018, 2019 et 2022.

Figure 3.2: Écoliers et étudiants jouant avec des puzzles inspirés de la théorie des pavages apériodiques.

Depuis fin 2013 nous participons régulièrement aux journées "Filles et maths : une équation lumineuse" organisées par les associations *femmes & mathématiques* et *Animath* plusieurs fois par an à l'USPN ou dans d'autres universités et destinées à des filles en fin de collège et au lycée. Ces journées ont pour but d'encourager les filles à s'orienter vers des études scientifiques et techniques.

Chapitre 4
Trajectoire de l'équipe CALIN

Sommaire

4.1 Atouts, faiblesses, opportunités, menaces 217
4.2 Positionnement par rapport au projet proposé en 2017 219
4.3 Projet 219
4.3.1 Analyse d'algorithmes et de structures combinatoires 219
4.3.2 Combinatoire, physique, géométrie 222

Nous présentons ci-dessous une analyse AFOM[1], qui met en exergue les *atouts, faiblesses, opportunités, menaces* pour l'équipe.

4.1 Atouts, faiblesses, opportunités, menaces

- **Forces et Atouts**

  — *Richesse des thématiques* : L'équipe réunit de manière rare tant au niveau national qu'international des compétences dans les divers domaines de la combinatoire

1. Nous boycottons l'horrible anglicisme anglicisme "SWOT" (strengths, weaknesses, opportunities, threats).

(analytique, géométrique, algébrique et bijective) avec une expertise particulière en analyse d'algorithmes et dans les problèmes à l'interface de la physique, mais aussi une transversalité des thèmes géométriques.

— *Attractivité* : L'équipe CALIN a recruté Nabil Mustafa, comme professeur en 2021. Les doctorants proviennent d'horizons internationaux (Moscou, Russie ; Antananarivo, Madagascar ; Hanoï, Vietnam) et de formations variées (informatique, mathématique et physique). Chaque année des candidats CNRS demandent leur affectation dans l'équipe (Florent Koechlin a été classé premier sur liste complémentaire en 2023).

— *Contrats de recherche* : CALIN a été partenaire de plusieurs projets ANR, dont 2 projets ANR internationaux avec Taïwan et le Luxembourg, et est fortement impliquée dans des collaborations avec l'Amérique du Sud, le Canada, le Luxembourg, Taïwan, l'Autriche et la Pologne.

— *Publications et édition* : L'équipe a un très bon niveau de publication (2 articles en revues internationales par équivalent permanent temps plein et par an, voir section 3.2) dans les meilleures revues internationales et participe au comité éditorial de deux journaux de combinatoire : les *Annales de l'Institut Henri Poincaré D : Combinatorics, Physics and their Interactions* (co-fondé par des membres de l'équipe) et *Annals of Combinatorics*.

— *Implication dans la formation par la recherche* : L'équipe a participe aux masters MPRI, MISA (Madagascar) et à l'EUR M&CS et a mis en place plusieurs groupes de travail ciblés en plus du séminaire d'équipe hebdomadaire.

— *Habilitations à diriger des recherches* : 2 habilitations ont été soutenues tout début 2023. L'un des nouveaux habilités est recruté PU à l'Université Gustave Eiffel.

Trajectoire de l'équipe CALIN

- **Faiblesses**

— *Absence de projet européen* : CALIN participe par l'intermédiaire de Nabil Mustafa à un projet ERC mais aucun membre de l'équipe n'est porteur principal d'un projet européen. V. Bonzom a passé l'oral de l'ERC Starting Grant en 2017, mais son projet n'a pas été financé

— *Développement logiciel trop réduit* : L'équipe fait peu de développement logiciel malgré la contribution importante à SageMath de l'un de ses membres.

— *Équilibre hommes-femmes* : L'équipe est très déséquilibrée de ce point de vue avec une seule femme parmi ses membres permanents.

- **Opportunités**

— *Pluridisciplinarité* : Les expertises variées que compte l'équipe permettent d'envisager la résolution de problèmes combinatoires (à l'interface entre informatique, mathématiques et physique) parmi les plus difficiles.

— *Recrutement de doctorants* : La participation au MPRI, à un master à l'Université d'Antananarivo (Madagascar), et à des écoles d'été en Russie offre des possibilités de recrutements de doctorants de haut niveau.

— *École Universitaire de Recherche à l'USPN* : La mise en place de l'EUR M&CS permet d'envisager des recrutement de doctorants de haut niveau.

— *Fédération MathSTIC à l'USPN* : Le développement de cette structure contribue à renforcer les collaborations avec certaines équipes du LAGA (Probabilités-Statistiques, Topologie Algébrique, Systèmes Dynamiques et Théorie Ergodique).

- **Menaces**

— *Responsabilités administratives* : L'augmentation des charges administratives tous azimuts est une menace importante pour le niveau d'activité scientifique de l'équipe.

En outre, deux des professeurs de l'équipe sont impliqués dans des responsabilités administratives lourdes (direction de l'école doctorale, direction du laboratoire).

— *Pérennité des financements* : Les ressources propres de l'équipe viennent principalement de projets ANR, la rendant dépendante de la politique scientifique menée par l'ANR.

4.2 Positionnement par rapport au projet proposé en 2017

Les thèmes de recherche que nous proposons restent volontairement assez ouverts afin de nous permettre d'orienter notre recherche de façon efficace et judicieuse. Le projet proposé en 2017 se déclinait, comme le projet ci-dessous, selon les axes de l'équipe. Nous avons travaillé sur tous les sous-thèmes proposés en 2017 et obtenus des résultats significatifs. Nous n'avons pas résolu exactement toutes les questions précises que nous avions proposées, mais orienté nos efforts de la façon la plus féconde possible en travaillant sur ces questions afin d'obtenir, de façon plus globale, des résultats qui ont dépassé nos attentes nous avions formulées dans chacun de ces thèmes (voir portfolio de l'équipe). Ceci a été possible car le projet proposé en 2017, structuré et cohérent, était aussi suffisamment flexible pour accommoder le caractère fondamental de nos thèmes de recherche :

La proposition ci-dessous se place dans le même esprit.

4.3 Projet

4.3.1 Analyse d'algorithmes et de structures combinatoires

Nous proposons de continuer d'étudier les questions autour de la génération aléatoire (en particulier dans l'idée d'améliorer la complexité des générateurs de Botzmann), de l'analyse de structures combinatoire (λ-termes, limites de graphes), mais aussi de profiter de la montée en puissance des thèmes plus géométriques avec des projets de livres à l'interface de la combinatoire, de la géométrie algorithmique et de la science des données, ou l'étude de certains graphes géométriques qui apparaissent dans des problèmes de génération aléatoire et en topologie algorithmique. Ces questions permettent de bénéficier de l'expertise de l'équipe dans son ensemble puisqu'elles portent sur des objets traités par les deux axes (typiquement les triangulations, les graphes, ou les polytopes).

- **Génération aléatoire**

Les générateurs Boltzmann ont révolutionné le domaine de la génération aléatoire en offrant l'accès à la génération de structures de très grande taille (typiquement $n = 10^6$, en quelques secondes). Cependant, ils reposent intrinsèquement sur une méthode de rejet lorsque l'objet engendré n'a pas la taille souhaitée. Ce rejet représente une perte nette considérable à la fois en temps et en termes de bits aléatoires consommés. Notre objectif est d'éviter ce rejet brutal de l'objet généré dès lors qu'il n'a pas la taille souhaitée, et de le réutiliser intelligemment pour la construction d'un objet ayant la bonne taille. Pour relever ce défi d'un générateur de Boltzmann exact en temps linéaire, nous allons travailler sur le paradigme que l'on peut qualifier de "rattrapage anticipé", que nous avons introduit [CO-10]. Cette idée est aussi l'un des ingrédients que nous avons utilisé [25] afin d'obtenir un algorithme linéaire de génération d'objets décrits par des grammaires non contextuelles ("context-free") irréductibles.

Nous nous attaquerons aussi aux extensions naturelles suivantes : gérer la classe des grammaires non irréductibles et/ou présentant des phénomènes de périodicité, gérer le

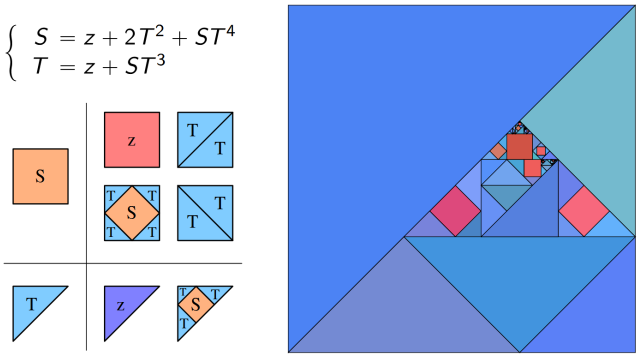


Figure 4.1: Un pavage récursif du carré par un jeu de triangles et de carrés peut être décrit par une grammaire context-free. La méthode de Boltzmann avec rattrapage anticipé permet de générer ces objets en temps linéaire. Nous souhaitons travailler sur l'étude de cette méthode pour des structures plus complexes.

cas multivarié (suivre des paramètres, similairement à [RI-98]), généraliser aux différents constructeurs admissibles de la méthode symbolique de Flajolet et Sedgewick.

- **Analyse de structures combinatoires**

Trajectoire de l'équipe CALIN

L'étude combinatoire des λ-termes reste un domaine largement ouvert. Très peu de travaux ont cherché à analyser le λ-calcul typé. De plus, les recherches sur la dynamique du λ-calcul en sont encore à leurs balbutiements, et une étude plus approfondie permettrait de mieux comprendre les mécanismes combinatoires de la réduction β. Enfin, la génération efficace de λ-termes à l'aide de paramètres de contrôle ouvrirait la porte à un nouveau type de tests logiciels et permettrait la découverte et la mise en évidence de propriétés difficiles à observer sur de petits λ-termes. En nous basant sur les résultats récemment obtenus par l'équipe sur l'énumération asymptotique des λ-termes linéaires [4], nous proposons de continuer notre analyse combinatoire des λ-termes.

L'analogue de la notion de motifs dans les permutations sont les sous-graphes induits et celui des permutations séparables sont les cographes, *i.e.*, les graphes ne contenant pas de P_4 (le chemin à 4 sommets) induits. Nous avons établi que la limite au sens des graphons [21] (*i.e.*, des matrices d'adjacence renormalisées) est un objet brownien, appelé *le cographon brownien* [**RI-101**]. Dans [**RI-102**], nous avons montré qu'un cographe aléatoire uniforme ne contient ni clique, ni ensemble indépendant de taille linéaire. En d'autres termes, le graphe P_4 n'a pas la propriété d'Erdős–Hajnal linéaire asymptotique [17].

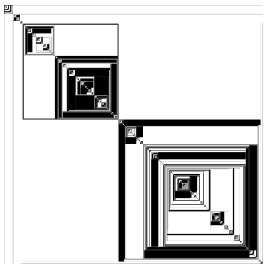


Figure 4.2: La matrice d'adjacence d'un cographe uniforme de taille 4482.

On s'attend à un phénomène d'universalité, et on aimerait montrer que les classes de graphes caractérisés par leur décomposition modulaire converge dans la plupart

des cas vers le *cographon brownien*. De premiers résultats dans cette direction sont obtenus dans [19]. Une autre décomposition permet de transformer des graphes en arbres, *la split décomposition*. Cette fois c'est la convergence après renormalisation en tant qu'espace métrique qui s'avère adapté pour étudier la limite des classes de graphes. Nous avons montré récemment [2] que dans certains cas la limite d'échelle de tels graphes aléatoires uniformes est *l'arbre aléatoire continu brownien d'Aldous*. L'extension de ces résultats à d'autres classes reste toutefois un challenge.

- **Combinatoire et science des données**

L'exploitation des données, qui sont souvent de nature géométrique et qui deviennent de plus en plus massives avec l'émergence de l'intelligence artificielle, pose un sérieux défi à la communauté scientifique, notamment sur le plan algorithmique. Nous proposons de travailler sur l'approximation rapide et optimale de ce type de données.

Pour cela, nous proposons de construire, pour un paramètre d'approximation ϵ donné, un sous ensemble minimal S des données et de l'étendre en une $(1+\epsilon)$-approximation du problème original. Ce projet qui regroupe les chercheurs de l'équipe intéressés par la combinatoire géométrique est divisé en trois parties, et nécessite une expertise en statistiques, en géométrie discrète, en apprentissage, en combinatoire et en algorithmique. Nous allons d'abord dégager les propriétés géométriques pertinentes des données qui seront transférées à S lors de sa construction. Deuxièmement, nous allons étudier les complexités (en temps et en mémoire) de la construction de S lorsque la dimension est grande, en se basant sur une compréhension des phénomènes géométriques et combinatoires sous-jacents. Enfin, nous souhaitons utiliser S pour améliorer la précision, et la rapidité des algorithmes d'optimisation.

Trajectoire de l'équipe CALIN

En plus des publications dans des revues scientifiques et des conférences, deux livres sont en préparation :

- *Modern Combinatorial Algorithms in Geometry*. Travail en commun János Pach (Budapest).
  Ce livre présentera les progrès récents sur les algorithmes d'optimisation utilisant des méthodes combinatoires et géométriques modernes, y compris : principes d'inclusion-exclusion, recherche locale, partions de l'espace, théorie de la séparation, algorithmes probabilistes.
- *The Multiplicative Weights Update Technique in AI, Geometry and Optimisation*. Travail en commun avec Wolfgang Mulzer (Berlin).
  La technique de mise à jour des poids multiplicatifs est un puissant outil algorithmique utilisé dans de nombreux domaines tels que l'intelligence artificielle, l'optimisation, la combinatoire, la géométrie et l'algorithmique.

- **Topologie algorithmique et géométrie des graphes**

Une triangulation d'une surface orientable $\mathcal{S}$ de type fini munie d'un ensemble $\mathcal{P}$ de points marqués est une décomposition de $\mathcal{S}$ en triangles dont les sommets appartiennent à $\mathcal{P}$. Le graphe des flips correspondant est celui dont les sommets sont ces triangulations et dont les arêtes joignent deux triangulations qui diffèrent d'exactement deux triangles. Ce type de graphe apparaît tant en informatique, lorsque la surface est un disque et que les triangulations sont duales aux arbres binaires [24] qu'en mathématiques où ils sont un modèle combinatoire du *mapping class group* de la surface [10, **RI-4**].

La constante de Cheeger d'un graphe permet d'identifier si celui-ci admet des goulots d'étranglement (et est donc liée au temps de mélange d'une chaine de Markov mise en oeuvre sur ce graphe). Cette quantité a été introduite par Jeff Cheeger dans le contexte de la géométrie riemannienne. Récemment, Frischberg et Eppstein ont montré que la constante de Cheeger du graphe des flips d'un polygone tend vers 0 lorsque le nombre de sommets du polygone tend vers l'infini [12], résolvant un problème ouvert posé dans plusieurs communautés. Le pendant de cette question existe et reste ouvert

dans le cas des surfaces (orientables et de type fini). Le résultat de Frischberg et Eppstein suggère une construction similaire dans ce cas que nous nous proposons d'étudier en nous appuyant sur notre expertise des aspects métriques des graphes de flips des surfaces topologiques [RI-4, RI-24, RI-25, 23]. Ceci permettrait d'étudier certaines questions de topologie algorithmique liées à la constante de Cheeger.

4.3.2 Combinatoire, physique, géométrie

Notre équipe continuera son travail important en physique combinatoire, et ses interactions avec la géométrie, notamment via les aspects suivants.

- **Cartes et intégrabilité en physique**

Dans les cartes combinatoires, nous étudierons les aspects algébriques et bijectifs. L'aspect algébrique est celui à l'origine des formules de récurrence à la Goulden–Jackson [14], à travers ce qu'on appelle l'*intégrabilité*. C'est une structure très profonde qui reste mal comprise et sous-exploitée en combinatoire. Nous souhaitons en particulier systématiser son utilisation pour obtenir des formules de récurrence. Il faut pour cela aussi être capable de reconnaître si des objets combinatoires satisfont l'intégrabilité, ce qui nécessitera un travail de fond à base de théorie des fonctions symétriques.

Les récurrences à la Goulden–Jackson sont des formules d'une simplicité incroyable qui suggère qu'elles peuvent être comprises directement sur les cartes en termes bijectifs. Ce problème est encore ouvert mais on sait [9, 20] qu'il s'agirait de *bijections d'un type nouveau*, ne reposant pas sur les distances à la manière de la célèbre bijection de Cori–Vauquelin–Schaeffer et ses descendants.

Trajectoire de l'équipe CALIN

En dimension 3, il existe plusieurs modèles de triangulations de sphères définis en termes d'arbres [11, 6, 3] dont certains assez récents, définis par des physiciens ou des mathématiciens, et qui n'ont pas encore été soumis à l'analyse combinatoire de notre école. Nous espérons obtenir des formules d'énumération ou des asymptotiques. Nous sommes aussi intéressés par des questions algorithmiques de décidabilité, comme décider si un triangulation donnée appartient à tel ensemble comme celui des localement constructibles de [11], ou encore s'il est possible de toutes les générer.

- **Combinatoire algébrique et classes de complexité**

Les coefficients de Kronecker qui apparaissent naturellement dans l'énumération de triangulations de dimension 3 [RI-15] sont des objets modèles en théorie géométrique de la complexité. Nous approfondirons les nombreuses questions de combinatoire algébrique associées, comme le calcul des multiplicités dans les décompositions de produits tensoriels en représentations irréductibles pour un groupe quelconque sujet à des contraintes minimales. Les résultats obtenus dans [RI-118] induisent l'existence d'une nouvelle classe d'algèbres de barycentres, dont la combinatoire reste à déterminer. L'étude du groupe $GL(n)$ et des coefficients Littlewood–Richardson qui jouent le rôle des coefficients de multiplicités pour ce groupe [13, 18] serait un cas intéressant à étudier.

Il est connu que le calcul des coefficients de Kronecker est dans la classe GapP et qu'il est #P-dur [22]. La classe de complexité de l'algorithme fourni en [RI-118] est exponentielle et donc ne peut fournir des réponses sur l'appartenance ou non à #P. Par contre, cet algorithme peut fournir des réponses à questions plus globales telles que : "calculer tous les coefficients de Kronecker ", ou "calculer la somme des coefficients de Kronecker". Il existe donc d'autres classes de problèmes intéressants à définir dont les classes de complexité seront à déterminer. Un récent papier a en outre introduit un algorithme quantique pour le calcul des coefficients de Kronecker et le place dans la classe #QMA [5]. Une question naturelle serait de se demander si la procédure trouvée

plus haut aurait un algorithme quantique de la forme [16], permettant d'accélérer son calcul et faire ainsi tomber son coût exponentiel.

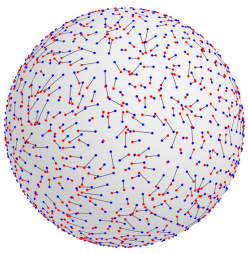
- **Pavages et empilements de sphères**

Un des objectifs généraux de l'équipe est de montrer que la notion d'ordre apériodique, bien que complexe, peut apparaître relativement naturellement, en particulier dans le contexte des sciences des matériaux (quasi-cristaux modélisés par des pavages apériodiques). Un premier objectif concret est d'achever le projet de caractérisation complète des pavages par coupe et projection qui peuvent être assemblés en respectant de simples contraintes locales. Une telle caractérisation pourrait alors servir de sorte de "feuille de route" pour la recherche de nouveau quasi-cristaux (un peu comme le tableau périodique des éléments, à peine à moitié plein lorsqu'il a été proposé par Mendeleïev en 1869, a stimulé la recherche d'éléments qui en rempliraient les cases vides). L'idée est la même dans le cas des arrangements de disques ou de sphères, encore plus naturels d'un point de vue sciences des matériaux. Notamment, la découverte de tailles explicites de disques ou de sphères dont l'arrangement le plus dense serait apériodique créerait un lien spectaculaire avec la théorie des pavages apériodiques et offrirait, par sa simplicité, un cadre motivant pour tenter une réalisation pratique (par exemple avec des nanoparticules dont on sait relativement bien contrôler la taille). Dans l'immédiat, l'objectif est d'explorer les arrangements potentiellement denses et d'affiner les techniques de preuve de densité maximale (un problème qui peut être très difficile, comme l'atteste la complexité de la preuve de la conjecture de Kepler [15]). Par exemple, Thomas Fernique travaille déjà sur le cas de deux tailles des sphères de ratio $\sqrt{2}-1$, dont l'arrangement conjecturé le plus dense a la structure du cristal de sel.

Trajectoire de l'équipe CALIN

- **Transport optimal entre points aléatoires sur une surface**

La théorie du transport optimal considère des problèmes d'optimisation formulés par Monge en 1781, et revisités par Kantorovich en 1942. C'est d'un domaine de recherche très actif avec de nombreuses applications (économie, logistique, reconnaissance d'images) sur lequel ont notamment travaillé deux médaillés Fields (Cédric Villani et Alessio Figalli), ainsi qu'un prix Nobel de physique (Giorgio Parisi). On s'intéresse ici à la version aléatoire 2-dimensionnelle du problème du transport optimal, consistant à trouver une bijection qui minimise le coût entre deux ensembles de n points aléatoires sur une surface (par exemple le carré, la sphère, un tore, une bouteille de Klein, ...), où le coût est défini comme la somme des carrés des distances euclidiennes entre les points appariés (tout comme le coût dans la méthode des moindres carrés).



Un exemple d'appariement entre points rouges et bleus choisis uniformément sur la sphère, qui minimise le coût (somme des carrés des distances géodésiques). Le challenge est de comprendre l'asymptotique de ce coût quand le nombre de points tend vers l'infini.

Ce coût est une fonction de n et dépend a priori de la géométrie de la surface dans laquelle sont choisis les points. Or, il a été conjecturé en 2014 [7, 8], et prouvé en 2017 [1] qu'en dimension 2 ce coût est asymptotiquement $\frac{1}{2\pi}\ln n/n$: le premier terme asymptotique ne dépend donc pas de la surface !

Nous [RI-78] avons récemment proposé un ansatz sur les termes asymptotiques suivants leur permettant de conjecturer que la forme de la surface n'intervient qu'à l'ordre $1/n$, avec un coefficient multiplicatif lié à la fonction zêta spectrale du laplacien associé à cette surface, $\sum_{k\geq 1}\lambda_k^{-s}$ (pour $s=1+\epsilon$), offrant ainsi une connexion jusqu'ici insoupçonnée avec la théorie spectrale. Ils ont notamment montré que cette fonction est le

premier terme d'une série perturbative (dont les coefficients sont des sommes de diagrammes de Feynman) de manière analogue à ce qui se fait en théorie quantique des champs. Ils espèrent développer une notion de diagrammes de Feynman "habillés" pour cette série, qui permettrait de contourner les méthodes d'analyse fonctionnelle et l'introduction du noyau de la chaleur (responsable dans [1] d'une perte de précision dans les termes d'erreur). Ceci est une piste prometteuse pour réussir à déterminer les premiers termes asymptotiques du coût optimal moyen et prouver notamment la conjecture de [RI-78].

Références complémentaires

[1] L. Ambrosio, F. Stra, and D. Trevisan. A pde approach to a 2-dimensional matching problem. *Probability Theory and Related Fields*, 173(1) :433–477, 2019.

[2] F. Bassino, M. Bouvel, V. Féray, L. Gerin, and A. Pierrot. Scaling limit of graph classes through split decomposition. *arXiv :2207.12253*, 2022.

[3] B. Benedetti and M. Pavelka. 2-LC triangulated manifolds are exponentially many. *Annales de l"Institut Henri Poincaré D*, 2021.

[4] O. Bodini, A. Singh, and N. Zeilberger. Asymptotic distribution of parameters in trivalent maps and linear lambda terms. *arXiv :2106.08291*, 2022.

[5] S. Bravyi, A. Chowdhury, D. Gosset, V. Havlíček, and G. Zhu. Quantum complexity of the Kronecker coefficients. *arXiv :2302.11454*, 2023.

[6] T. Budd and L. Lionni. A family of triangulated 3-spheres constructed from trees. *arXiv :2203.16105*, 2022.

[7] S. Caracciolo, C. Lucibello, G. Parisi, and G. Sicuro. Scaling hypothesis for the euclidean bipartite matching problem. *Physical Review E*, 90 :012118, Jul 2014.

[8] S. Caracciolo and G. Sicuro. Quadratic stochastic euclidean bipartite matching problem. *Physical Review Letters*, 115 :230601, Dec 2015.

[9] G. Chapuy, V. Féray, and E. Fusy. A simple model of trees for unicellular maps. *Journal of Combinatorial Theory, Series A*, 120 :2064–2092, 2013.

[10] V. Disarlo and H. Parlier. The geometry of flip graphs and mapping class groups. *Transactions of the American Mathematical Society*, 372 :3809–3844, 2019.

[11] B. Durhuus and T. Jonsson. Remarks on the entropy of 3-manifolds. *Nuclear Physics B*, 445 :182–192, 1995.

[12] D. Eppstein and D. Frishberg. Improved mixing for the convex polygon triangulation flip walk. *arXiv :2207.09972*, 2022.

[13] W. Fulton. *Young Tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, 1997.

[14] I. P. Goulden and D. M. Jackson. The KP hierarchy, branched covers, and triangulations. *Advances in Mathematics*, 219 :932–951, 2008.

[15] T. C. Hales. A proof of the Kepler conjecture. *Annals of Mathematics*, 162 :1065–1185, 2005.

[16] A. W. Harrow, A. Hassidim, and S. Lloyd. Quantum algorithm for linear systems of equations. *Physical Review Letters*, 103 :150502, 2009.

[17] R. J. Kang, C. McDiarmid, B. A. Reed, and A. D. Scott. For most graphs H, most H-free graphs have a linear homogeneous set. *Random Structures and Algorithms*, 45(3) :343–361, 2014.

[18] R. C. King, C. Tollu, and F. Toumazet. Factorisation of littlewood–richardson coefficients. *Journal of Combinatorial Theory, Series A*, 116 :314–333, 2009.

[19] T. Lenoir. Graph classes with few P_4's : Universality and Brownian graphon limits. *arXiv :2301.13607*, 2023.

[20] B. Louf. A new family of bijections for planar maps. *Journal of Combinatorial Theory, Series A*, 168 :374–395, 2019.

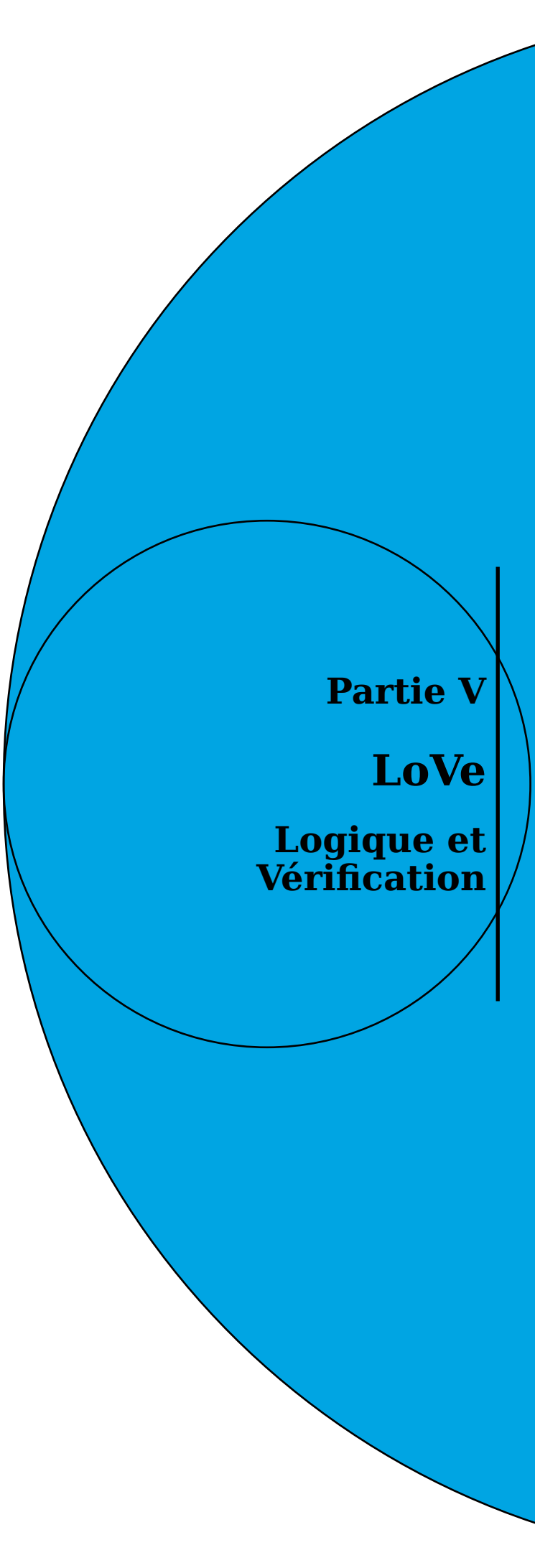
[21] L. Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012.

[22] I. Pak and G. Panova. On the complexity of computing Kronecker coefficients. *Computational complexity*, 26 :1–36, 2017.

[23] H. Parlier and L. Pournin. Getting lost in flip-graphs. *in preparation*, 2023.

[24] D. Sleator, R. Tarjan, and W. Thurston. Rotation distance, triangulations, and hyperbolic geometry. *Journal of the American Mathematical Society*, 1 :647–681, 1988.

[25] A. Sportiello. Boltzmann sampling of irreducible context-free structures in linear time. *arXiv :2105.12881*, 2021.



Partie V

LoVe

Logique et Vérification

Sommaire

Informations générales sur l'équipe LoVe **229**
1.1 Membres de l'équipe LoVe . . . 230
1.2 Les thématiques scientifiques et leurs enjeux . . . 232
1.3 Profil d'activités liées à la recherche . . . 239
1.4 Prise en compte des recommandations du précédent rapport . . . 239

Introduction du portfolio **243**

Auto-évaluation de l'équipe LoVe **245**
3.1 Domaine 2. Attractivité . . . 246
3.2 Domaine 3. Production scientifique . . . 251
3.3 Domaine 4. Inscription des activités de recherche dans la société . . . 273

Trajectoire de l'équipe LoVe **275**
4.1 SWOT de l'équipe . . . 275
4.2 Trajectoire . . . 276

Chapitre 1

Informations générales sur l'équipe LoVe

Sommaire

1.1 Membres de l'équipe LoVe 230
1.1.1 Membres permanents 230
1.1.2 Membres non-permanents au 01/01/2023 231
1.1.3 Évolution de l'équipe (membres permanents) 232
1.2 Les thématiques scientifiques et leurs enjeux 232
1.2.1 Logique, théorie de la programmation et complexité 233
1.2.2 Vérification 236
1.3 Profil d'activités liées à la recherche 239
1.4 Prise en compte des recommandations du précédent rapport . . . 239

Nom de l'équipe : Logique et Vérification (LoVe)

Responsable de l'équipe (depuis 2020) : Damiano Mazza

Responsable de l'axe « vérification » : Kaïs Klai

Responsable de l’équipe (2017–2020) : Christophe Fouqueré

1.1 Membres de l’équipe LoVe

Tableau synthétique de l’évolution des membres de l’équipe (nombre de membres au 1er janvier de chaque année) :

Corps	2017	2018	2019	2020	2021	2022	(2023)
DR	1	1	1	1	2	2	(2)
CR	2	3	3	3	3	3	(3)
PU	6	6	6	6	5	5	(5)
MCF	10	9	8	8	8	9	(10)
IR	0	0	1	1	1	1	(1)
Total permanents	*19*	*19*	*19*	*19*	*19*	*20*	*(21)*
Post-docs et ATER	0	0	0	0	3	1	(2)
Doctorants	8	6	8	9	10	9	(12)
Total	*27*	*25*	*27*	*28*	*32*	*30*	*(35)*

1.1.1 Membres permanents

(membres ou, *en italique, anciens membres*, et, **en gras, responsable de l’équipe** au 01/01/2023)

	Nom	*Prénom*	*Situation*	*Institution*
	André	Étienne	PU	IG
	Arias	Jaime	IR	CNRS
	Boudes	Pierre	MCF	IG
	Breuvart	Flavien	MCF	IG
	Choppy	*Christine*	PU	IG
	Coupey	*Pascal*	PU	IUTV
	Evangelista	Sami	MCF	IUTV
	Fouqueré	Christophe	PU	IG
	Guerrini	Stefano	PU	IG
	Jacobé de Naurois	Paulin	CR	CNRS
	Kerjean	Marie	CR	CNRS
	Klai	Kaïs	MCF HDR	IG

Informations générales sur l’équipe LoVe

	Loddo	Jean-Vincent	MCF	IUTV
	Manzonetto	Giulio	MCF HDR	IUTV
	Mayero	Micaela	MCF HDR	IUTV
	Mazza	**Damiano**	DR	CNRS
	Mogbil	Virgile	MCF HDR	IG
	Moyen	*Jean-Yves*	MCF	IG
	Olarte	Carlos	MCF	IG
	Petrucci	Laure	PU	IUTV
	Schwer	Sylviane	PU	IG
	Seiller	Thomas	CR	CNRS
	Touili	Tayssir	DR	CNRS
	Youcef	Samir	MCF	IG

1.1.2 Membres non-permanents au 01/01/2023

	Nom	*Prénom*	*Situation*	*Financement*
	Chanus	Baptiste	doctorant	CD USPN
	Dufour	Aloÿs	doctorant	CD USPN
	Eng	Boris	doctorant/ATER	CD USPN
	Fievet	Baptiste	doctorant	CD CNRS
	Kerinec	Axel	doctorant	CDSN
	Léchine	Ulysse	doctorant	CD USPN
	Mirwasser	Simon	doctorant	CD USPN
	Munnich	Nicolas	doctorant	CD ANR
	Paquet	Hugo	post-doctorant	COFUND UE/IdF

Informations générales sur l'équipe LoVe

Ragot	Adrien	doctorant	bourse UFI
Rogers	Morgan	post-doctorant	MathSTIC
Souid	Nour	doctorante	CD Erasmus+

1.1.3 Évolution de l'équipe (membres permanents)

La taille de l'équipe est restée essentiellement stable sur la période, avec une croissance vers la fin : les 4 départs (1 promotion MCF, 2 mises en disponibilité de MCF et PU, 1 départ à la retraite de PU) ont été compensés par 6 recrutements (2 MCF dont un en mutation, 1 PU, 2 CR et 1 IR).

Plus en détail, l'axe vérification, qui s'était déjà affaibli en 2016 par le départ (en disponibilité en entreprise) de l'un de ses MCF, a subi le départ d'un MCF en 2018 (recruté PU à Nancy) et le départ à la retraite d'une PU en 2020. Ces départs ont été compensés par :

— le recrutement d'un IR CNRS en 2018 ;
— le recrutement d'un MCF en 2021 ;
— le recrutement d'un PU en 2022 (il s'agit en fait du MCF ayant quitté l'équipe en 2018) ;
— l'arrivée d'un MCF en mutation en 2022.

L'axe vérification s'est donc renforcé sur la période.

Les effectifs de l'axe logique sont restés stables. Deux membres sont actuellement en disponibilité : un MCF depuis 2017 (emploi en entreprise) et un PU depuis 2022 (ce dernier n'était plus actif depuis plusieurs années, consacrées à des tâches d'administration). Nous signalons aussi un MCF qui, depuis septembre 2021, est à temps partiel (50%). Ces départs ont été compensés par le recrutement de 2 CR CNRS, en 2018 et 2020. L'axe a aussi été renforcé par la promotion d'un CR en DR.

Départs

2017 J.-Y. Moyen, MCF IG (logique, disponibilité entreprise)

2018 É. André, MCF IUTV (vérification, promu PU à Nancy)

2020 C. Choppy, PU IG (vérification, retraite)

2020 D. Mazza, CR CNRS (logique, promu DR)

2022 P. Coupey, PU IUTV (logique, disponibilité académique)

Arrivées

2018 J. Arias, IR CNRS (vérification, thèse LaBRI Bordeaux)

2018 T. Seiller, CR CNRS (logique, thèse IML Marseille)

2020 M. Kerjean, CR CNRS (logique, thèse IRIF Paris 7)

2020 D. Mazza, DR CNRS (logique, CR LIPN)

2021 C. Olarte, MCF IG (vérification, prof. associé Universidade Federal do Rio Grande do Norte, Brazil)

2022 É. André, PU IG (vérification, PU Nancy)

2022 S. Youcef, MCF IG (vérification, mutation MCF Nancy)

Informations générales sur l'équipe LoVe

1.2 Les thématiques scientifiques et leurs enjeux

Les recherches de l'équipe LoVe s'articulent sur deux axes thématiques, menant leurs activités de façon essentiellement indépendante :

— **logique, théorie de la programmation et complexité** ;
— **vérification**.

La coexistence de ces deux axes dans la même équipe est justifiée par des raisons historiques plus que thématiques. Cependant, les interactions ne sont pas inexistantes :

— des collaborations, quoiqu'en nombre faible, se sont produites dans les années passées, surtout pour ce qui concerne l'interaction entre vérification et preuves formelles.

— Sur la période d'évaluation, l'équipe a porté un projet de collaboration international CNRS (projet TrAVAIL) impliquant plusieurs universités danoises (Copenhague, Aarhus) et des membres des deux axes de l'équipe. Après son achèvement, une demande de création d'IRN CNRS a été faite, afin de stabiliser ces coopérations dans un cadre formel. Elle a reçu l'avis favorable de l'INS2I et la mise en place de l'IRN se fera fin 2023/début 2024.

— Les deux axes interagissent également au niveau de l'enseignement à l'Institut Galilée, notamment dans le cadre du Master 2 « Programmation et Logiciels Sûrs » et de l'EUR « Mathématique et Informatique », qui offrent des modules à la fois en logique et programmation et en spécification et vérification.

Il est néanmoins plus simple et cohérent de présenter les activités en deux sections séparées, ce que nous ferons par la suite.

1.2.1 Logique, théorie de la programmation et complexité

Historiquement, le noyau de cet axe est la *sémantique des languages de programmation* et ses liens avec la logique et la théorie de la démonstration, surtout par le biais de la *correspondence de Curry-Howard* et de la *logique linéaire*. La spécificité de l'équipe consiste en une sensibilité pour les *aspects quantitatifs* du calcul et des programmes, avec un intérêt particulier pour la *complexité*. Une expertise en *preuves formelles* est aussi présente dans l'équipe, ainsi que des recherches complémentaires en linguistique formelle.

Ces thèmes s'inscrivent, en France, dans le contexte du GDR Informatique Mathématique, et plus précisément des groupes de travail SCALP (Structures Formelles pour les Preuves et les Programmes) et LHC (Logique, Homotopie, Catégories) et également, en partie, dans le contexte du GDR Génie de la Programmation et du Logiciel, groupe de travail LTP (Langages, Types et Preuves).

Au niveau international, les membres de l'axe émargent aux communautés représentées par EACSL en Europe (*European Association for Computer Science Logic*), et ACM SIGLOG et SIGPLAN en Amérique du Nord (*ACM Special Interest Group on Logic and Computation* et *ACM Special Interest Group on Programming Languages*). Les principales conférences de ces domaines sont LICS (Logic in Computer Science) et similaires (ICALP, FoSSaCS, CSL, FSCD. . .), pour les aspects liés à la logique et la complexité, et POPL (Principles of Programming Languages) et similaires (ESOP, ICFP, PPDP. . .), pour les aspects plus strictement liés à la théorie des langages de programmation, ainsi que des conférences orientées aux méthodes formelles (FM, CPP. . .).

Informations générales sur l'équipe LoVe

La vie scientifique de cet axe de l'équipe est centrée autour d'un séminaire, tendentiellement hebdomadaire (mais avec des irrégularités), et de groupes de travail thématiques qui peuvent évoluer en fonction des besoins. Depuis le début de la pandémie de Covid-19, toutes les séances de ces séminaires/groupes de travail sont en modalité hybride et sont enregistrées pour consultation ultérieure. Actuellement, nous avons :

séminaire « Logique » : il est principalement consacré à des intervenants externes, sur des sujets d'intérêt général pour les membres de l'axe. Nous avons 44 séances enregistrées de mi-avril 2020 à décembre 2022.

séminaire « Complexités » : c'est essentiellement un groupe de travail interne au laboratoire, le lieu d'animation de l'axe transversal « Complexités » du LIPN. Plusieurs membres de notre équipe y participent régulièrement. De fréquence tendentiellement hebdomadaire, il porte sur des sujets de complexité algorithmique. Nous avons 9 séances enregistrées de mi-septembre à décembre 2022.

groupe de travail « Catégories » : il s'agit d'un groupe de travail commun avec l'équipe de Topologie Algébrique du LAGA (le laboratoire de mathématiques de USPN), dans le cadre de la Fédération MathSTIC et de son axe « Catégories, Calcul et Topologie », de création récente (printemps 2022). Les séances sont consacrées à des thèmes spécifiques ou à des relectures d'articles, et sont souvent organisées par cycles thématiques. La fréquence est tendiellement hebdomadaire. En 2022, il y a eu 24 séances, par 10 orateurs internes (LIPN ou LAGA) et 5 invités externes.

Au cours de la période d'évaluation nous avons également eu un groupe de travail « Sémantique de la programmation » (2012–2019) et un groupe de travail « Mathématiques du calcul » (2019–2021). Par ailleurs, le séminaire « Logique » adhère au « macro-séminaire » *Chocola* (Curry-Howard : Calcul et Logique), qui est organisé une fois par mois à Lyon et tient lieu de séminaire d'équipe quand il a lieu.

Il est important de souligner que, pour les membres de cet axe, ces moments d'interaction définissent l'appartenance à l'équipe. En effet, de par la nature de nos thématiques scientifiques, le plus souvent il n'existe pas de projet commun auquel chaque membre contribue en développant une partie ou un aspect, comme il pourrait se produire avec des thématiques de nature plus appliquée. L'appartenance à l'équipe, donc, ne signifie pas tellement ici le fait de travailler au même projet mais, plutôt, de parler la même langue, partager ses connaissances, ses avancées, et s'enrichir mutuellement grâce aux moyens d'interaction fournis par l'équipe. Ces séminaires et groupes de travail sont essentiels dans ce sens.

Les résultats majeurs obtenus dans cet axe au cours de la période d'évaluation s'inscrivent dans trois thèmes principaux :

— **sémantique de la programmation** ;
— **complexité** ;
— **preuves formelles**.

Avant de présenter une synthèse de ces résultats, mentionnons également la présence de recherches appartenant à des thèmes complémentaires, à savoir :

linguistique formelle : les travaux sur la mathématique du dialogue [LI-2] ont permis d'expliciter la manière dont les modèles récents en théorie de la démonstration (en particulier issus de la logique linéaire) peuvent rendre compte des interactions dans un dialogue entre humains, aussi dans des contextes à priori difficiles à modéliser, comme ceux relevant de patients schizophrènes [CL-9].

Nous avons également une collaboration avec le SCRIME (Bordeaux) dans le domaine de l'informatique musicale, en particulier représentation et raisonnement temporel qualitatif dans le cadre de la théorie des langages (modèle des s-langages) [CO-35].

Informations générales sur l'équipe LoVe

philosophie de l'informatique : nous étudions l'impact de résultats récents en informatique théorique sur des questions philosophiques [RI-24], notamment dans le cadre d'un projet ANR avec l'Université Paris 1, porté par Alberto Naibo.

- **Sémantique de la programmation**

Un des thèmes unifiants de l'axe est la théorie des *approximations de programmes*, en particulier (mais pas seulement) les approximations *linéaires*. L'idée de fond est que les programmes arbitraires peuvent être approximés par des programmes « simples » (typiquement, des programmes ne contenant aucune boucle, qui terminent toujours en temps linéaire en leur taille), et que tout programme est la limite des ses approximations « simples ». Cette idée est déjà contenue dans le papier original de Girard introduisant la logique linéaire [6] et a ensuite été formalisée par Ehrhard et Regnier [4, 5]. Les travaux de l'équipe ont dévoilé le lien profond entre approximations et *types intersection* [**CI-37**], permettant ainsi d'étendre ces systèmes de types, qui sont capables de capturer très précisément le comportement des programmes, à des contextes complètement nouveaux, comme la programmation probabiliste [CI-47] ou concurrente [**CI-63**].

La notion d'approximation a aussi été étendue au cadre plus général du $\lambda\mu$-calcul [**CI-112**], et il a été montré comment elle permet de retrouver, avec des preuves plus simples, un nombre impressionant de résultats fondamentaux (voire nouveaux !) sur le λ-calcul [**CI-82**] (*voir le portfolio pour plus de détails*).

L'équipe s'est aussi penchée sur des sujets particulièrement d'actualité en théorie de la programmation, comme la programmation probabiliste [**RI-28**] et la *programmation différentiable*. Cette dernière s'occupe de langages de programmation intégrant l'algorithme de *rétropropagation* permettant de calculer efficacement le gradient d'une fonction spécifiée par un programme écrit dans le langage lui-même. La programmation différentiable est d'importance capitale, entre autre, en apprentissage profond. L'équipe est parmi les plus avancées au monde sur cette question, fournissant la première généralisation prouvablement correcte de la rétropropagation aux langages fonctionnels [**CI-81**], ainsi qu'une analyse logique fine de la rétropropagation, la reliant à la « transformation Dialectica » de Gödel (travail en préparation [7]), et, pour finir, la première analyse formelle des erreurs que la rétropropagation introduit dans les langages de programmation utilisés en pratique [**CI-88**]. Incidentalement, la notion d'approximation mentionnée ci-dessus joue un rôle clé dans ce résultat (*voir le portfolio pour plus de détails*).

L'équipe produit également des résultats dans la théorie du λ-calcul pur, où une conjecture vieille de plus de 40 ans a été refutée [CI-13, **RI-17**]. La co-publication avec Barendregt du livre *A Lambda-Calculus Satellite* [**LI-3**], une sorte de « mise à jour » de la célèbre référence historique sur le λ-calcul publiée dans les années 1980 par Barendregt lui-même, est aussi un belle réalisation pour l'axe (*voir le portfolio pour plus de détails*).

- **Complexité**

L'équipe possède une expertise historique et internationalement reconnue en *complexité implicite*, c'est-à-dire la caractérisation de classes de complexité à travers des systèmes logiques ou des langages de programmation (souvent basés sur de tels systèmes). Dans cette veine, des nouvelles caractérisation de « petites » classes (LOGTIME, ALOGTIME, L, NC, P) ont été obtenues à l'aide d'une extension des techniques de « *tiering* » à la classe des fonctions récursives (à la Klenee) appliquée au calcul sur les pointeurs [CI-61]. Des caractérisations existantes pour les classes de complexité séquentielles (P) ont été étendues aux classes parallèles (PAR, temps polynomial parallèle) [CI-108] ou aux fonctions PTIME sur les nombres réels [CI-80].

Dans une direction différente, l'équipe a introduit une perspective nouvelle sur la complexité implicite, obtenant des caractérisations logiques ou dans des langages fonctionnels (variantes linéaires du λ-calculs simplement typé) de classes de fonctions/-problèmes calculées/décidés par différents types d'automates [**RI-6** , **CI-87**, **CI-78**, **CI-55**]. Ces résultats sont réellement novateurs dans le domaine, à la fois pour les outils employés (actions de monoïdes, graphages, logique linéaire non-commutative. . .) et pour les liens tissés avec la théorie des automates, et constituent de réalisations remarquables pour l'axe.

Un autre résultat intéressant de l'axe a été le transport des techniques de complexité implicite à l'analyse statique de programmes. La première application a été à la simplification de boucles (détection de quasi-invariants de boucle) [CI-11], qui a mené ensuite à un outil, implementé en version prototype, permettant d'analyser automatiquement la complexité de programmes d'un sous-ensemble de C [CI-113], ainsi qu'à une extension de cette méthode permettant de paralleliser automatiquement du code C [CI-119].

- **Preuves formelles**

Concernant les preuves formelles, l'axe s'intéresse principalement à la formalisation de l'analyse réelle (en logique classique) et à ses applications à la programmation, notamment à l'analyse numérique. Le but des travaux en lien avec l'analyse numérique

consiste à utiliser l'assistant à la preuve Coq pour augmenter la confiance dans les programmes de résolution numérique d'équations aux dérivées partielles (EDP) et de bibliothèques d'éléments finis. Ces travaux sont effectués en collaboration avec des chercheurs et chercheuses d'Inria et de l'UTC! [**RI-38**, CI-24] et donnent également lieu au développement de la bibliothèque `coq-num-analysis` [LO-14]. Dans une autre direction, l'axe a contribué au projet MathComp-Analysis (https://github.com/math-comp/analysis) en participant au port de MathComp à Hierarchy Builder (une nouvelle manière de gérer les hiérarchies de structures mathématiques).

Les membres de l'axe ont également combiné leur expertise en complexité implicite avec les preuves formelles, en fournissant une preuve certifiée de la borne polynomiale donnée par les *quasi-interprétations*, une méthode standard en complexité implicite [CI-41].

Dans une direction complètement différente, des membres ont expérimenté l'utilisation d'assistants de preuves dans l'enseignement supérieur (niveau licence) [RE-5]. Ils participent d'ailleurs, depuis 2021, au défi Inria LiberAbaci sur ce sujet (https://www.inria.fr/fr/liberabaci).

1.2.2 Vérification

Les travaux de cet axe s'articulent autour de deux principales thématiques :

1. La conception d'approches et algorithmes de vérification basés sur l'exploration de l'espace des états d'un système (dites de *model checking*).
2. La conception d'approches de vérification formelles assurant la sécurité des systèmes.

Certaines contributions des membres de l'axe sont dédiées à des domaines spécifiques, en exploitant leurs particularités : monitoring [**CI-51**, **CI-52**, ↑RI-14], les processus métiers [CI-23, CI-48], les processus dans le Cloud [CI-40, CI-39, **CI-98**] et dans la Blockchain [**CI-92**, LO-10, CI-109]. Par ailleurs, une caractéristique majeure de l'axe Vérification est l'importance accordée au développement logiciel qui accompagne quasi-systématiquement les contributions proposées. Une description plus détaillée des outils développés dans l'axe est présentée dans le portfolio 6.

Informations générales sur l'équipe LoVe

- **Vérification de systèmes concurrents**

Vérification de systèmes temporisés et/ou paramétrés. La prise en compte de contraintes temporelles est déterminante dans la modélisation de systèmes critiques, et en particulier cyber-physiques. Les automates temporisés et les réseaux de Petri temporels sont des modèles largement étudiés. Toutefois, au moment de la conception ou de la maintenance, les valeurs des bornes temporelles ne sont pas nécessairement connues. Les automates temporisés paramétrés (PTA) pemettent d'exprimer des constantes de temps non connues *a priori*. Il s'agit non seulement de décider si la propriété est valide, mais aussi de synthétiser l'ensemble des valeurs de paramètres pour qu'elle le soit. Les propriétés de ces systèmes étant généralement indécidables, nous développons des heuristiques efficaces qui fournissent une réponse exacte lorsque l'algorithme termine (mais cette terminaison n'est pas garantie) [**CI-26**, **CI-73**, CI-116, **CI-100**, RI-30, CI-36, RI-18, **?**].

Par ailleurs, la logique de réécriture [8] (LR) ainsi que l'outil associé Maude [2] sont dotés de puissantes techniques de raisonnement [3] et une riche théorie équationnelle que les PTA [1] et réseaux de Petri temporels paramétrés (PTPN) [9] ne possèdent pas. Dans [**?**] nous avons proposé une sémantique basée sur la LR pour les PTAs et les PTPNs qui capture le comportement (symbolique) des systèmes. De plus, nous avons utilisé Maude-SMT pour mettre en œuvre des analyses telles que : accessibilité, *model checking* de LTL, et synthèse de paramètres. Nous avons également montré qu'il

est possible d'analyser ces systèmes selon une stratégie d'exécution définie par l'utilisateur sans modifier le modèle sous-jacent. Ces publications ont ouvert la voie pour explorer la possibilité d'étendre ces résultats aux théories de réécriture généralisées et définir des techniques de vérification paramétrée pour la LR.

Vérification parallèle et/ou distribuée. Nous avons attaqué cette thématique pour la vérification parallèle des propriétés temporelles d'un système en distribuant la construction et l'exploration de l'espace d'états sur plusieurs processus chacun pouvant avoir plusieurs *threads*. L'espace d'états est représenté par une abstraction.

Une approche utilise le SOG (*Symbolic Observation Graph*) qui préserve les propriétés temporelles insensibles au bégaiement. Ces travaux [CI-7, CI-32, CI-57, CI-84] ont permis d'aboutir à un model checker parallèle pour LTL, avec des résultats prometteurs. Nous avons également proposé une extension hybride (PMC-SOG) [CO-34].

Une autre approche a permis la conception d'algorithmes de vérification distribués opérant sur des réseaux basés sur la technologie RDMA (*Remote Direct Memory Access*). Cela nécessite de revisiter les algorithmes de vérification distribués existants pour les adapter à ce modèle de communication ou d'en concevoir de nouveaux. Ces travaux ont fait l'objet de plusieurs publications : [**CI-46**, CI-45, CI-110].

Vérification de systèmes multi-agents. La vérification de systèmes multi-agents ajoute une nouvelle dimension : existe-t-il une stratégie d'un agent ou d'un groupe d'agents pour garantir la satisfaction d'une propriété ? Les travaux dans ce cadre ont de nombreuses facettes, en particulier les sémantiques exprimant les connaissances de l'agent ([**RI-16**], voir la fiche portfolio 1), les modèles synchrones ou asynchrones [**CI-120** , CI-115] et les logiques permettant d'exprimer les propriétés [**RI-16** , **CI-120**] qui intègrent stratégies d'agents et contraintes temporelles.

Nous avons aussi étudié les arbres attaque-défense sous l'angle de systèmes d'agents collaborant pour mener à bien une attaque, en proposant une approche de modélisation [CI-83], des réductions d'espace d'états [CI-59], une planification des agents [CI-114], et une implémentation logicielle [**CI-99**].

- **Vérification de propriétés de sécurité**

Détection de malwares. La détection de malware est un sujet d'actualité. En effet, les techniques actuelles implémentées dans les antivirus commerciaux présentent des limitations qui les rendent faciles à contourner par les développeurs de malware. Pour avoir des techniques robustes de détection de malware, il faut des approches permettant d'analyser le code (pas la syntaxe) des exécutables sans les exécuter. C'est ce que permet le model checking.

Dans [CI-8, CI-9, CI-35, CI-58], nous avons proposé de nouvelles logiques temporelles qui permettent de spécifier de manière précise et compacte les comportements néfastes. Dans [CI-10], nous avons proposé un algorithme de model-checking pour les DPNs (réseau d'automates à pile qui modélisent les programmes concurrents) pour les formules CARET. Nous avons développé dans [**CI-15**] un algorithme pour calculer l'ensemble de toutes les configurations accessibles de DPNs avec priorités et verrous. Nous avons proposé dans [CI-4] une approche qui permet de résoudre l'accessibilité des DPNs pour lesquels les processus communiquent par rendez-vous. Par ailleurs, pour pouvoir analyser les codes automodifiants, nous avons proposé dans [CI-1, CI-54, **CI-76**] un nouveau modèle (les automates à pile automodifiants) et des techniques d'analyse pour ces modèles. Enfin, nous avons proposé dans [CI-16, **CI-17**, CI-44, CI-18, **CI-43**, RI-37, CI-62] des approches basées sur des techniques d'apprentissage qui permettent d'extraire de manière automatique les comportements néfastes à partir d'un ensemble de malware et de benware.

Vérification et supervision de l'opacité. La sécurité des systèmes communicants est une thématique cruciale, spécialement quand il s'agit de systèmes critiques. L'*opacité* est une propriété de confidentialité qui formule la capacité d'un système à garder caché un secret pour un attaquant. Les travaux de l'équipe sur ce sujet ont été (1) de concevoir des algorithmes nouveaux et efficaces pour la vérification de l'opacité et ses variantes [CI-20, CI-22, RI-3, CI-64] et, (2) de synthétiser un superviseur optimal afin de contrôler cette propriété pour les systèmes à évènements discrets [**CI-85**, **CI-105**, CI-104, CI-103]. Les travaux sur la vérification de l'opacité ont inspiré des contributions (dont [RE-1, RI-29]) sur la vérification d'une propriété duale à savoir la diagnosticabilité des systèmes à évènements discrets. En outre, nous avons considéré l'opacité dans un cadre temporisé (celui des automates temporisés), avec des extensions à la synthèse de paramètres : nous avons proposé une nouvelle définition de l'opacité, où l'attaquant(e) n'a accès qu'au modèle et au temps d'exécution ; dans ce cas, l'opacité d'un système devient décidable [**CI-65**, ↑RI-16] ; des extensions ont été étudiées pour le contrôle [CO-37] et l'opacité avec une date d'expiration [**?**].

Arbres attaques-défense. Les arbres attaque-défense permettent d'exprimer des problèmes de sécurité. Nous les avons étudiés sous l'angle de systèmes d'agents collaborant pour mener à bien une attaque, en proposant une approche de modélisation [CI-83], des réductions d'espace d'états [CI-59], une planification des agents [CI-114], et une implémentation logicielle [**CI-99**].

- **Développement logiciel**

Les membres de l'axe Vérification sont de grands défenseurs de la science ouverte. Dans ce contexte, les logiciels développés par l'équipe sont open-source, et la majorité de publications sont accompagnées d'un artefact accessible au public (dépôt `git`) pour la reproductibilité des résultats. De plus, le code source des logiciels est archivé dans la plate-forme *Software Heritage*.

En sachant que l'installation et la configuration de certains outils sont compliquées, l'équipe dispose d'un serveur puissant dédié pour mettre à disposition des utilisateurs ses logiciels afin de les expérimenter. De même, plusieurs logiciels sont intégrés à la plateforme CosyVerif qui fournit une interface graphique conviviale pour l'utilisation des nos outils.

Informations générales sur l'équipe LoVe

Références complémentaires

[1] R. Alur, T. A. Henzinger, and M. Y. Vardi. Parametric real-time reasoning. In *STOC*, pages 592–601. ACM, 1993.

[2] M. Clavel, F. Durán, S. Eker, P. Lincoln, N. Martí-Oliet, J. Meseguer, and C. L. Talcott. *All About Maude - A High-Performance Logical Framework, How to Specify, Program and Verify Systems in Rewriting Logic*, volume 4350 of *LNCS*. Springer, 2007.

[3] F. Durán, S. Eker, S. Escobar, N. Martí-Oliet, J. Meseguer, R. Rubio, and C. L. Talcott. Programming and symbolic computation in maude. *J. Log. Algebraic Methods Program.*, 110, 2020.

[4] T. Ehrhard and L. Regnier. Böhm trees, krivine's machine and the taylor expansion of lambda-terms. In *Proceedings of Computability in Europe (CiE)*, volume 3988 of *Lecture Notes in Computer Science*, pages 186–197. Springer, 2006.

[5] T. Ehrhard and L. Regnier. Uniformity and the taylor expansion of ordinary lambda-terms. *Theor. Comput. Sci.*, 403(2-3) :347–372, 2008.

[6] J.-Y. Girard. Linear logic. *Theor. Comput. Sci.*, 50 :1–102, 1987.

[7] M. Kerjean and P.-M. Pédrot. ∂ is for Dialectica. 2021.

[8] J. Meseguer. Conditioned rewriting logic as a united model of concurrency. *Theor. Comput. Sci.*, 96(1) :73–155, 1992.

[9] L.-M. Traonouez, D. Lime, and O. H. Roux. Parametric model-checking of time Petri nets with stopwatches using the state-class graph. In F. Cassez and C. Jard, editors, *Formal Modeling and Analysis of Timed Systems (FORMATS 2008)*, volume 5215 of *LNCS*, pages 280–294. Springer, 2008.

1.3 Profil d'activités liées à la recherche

Activités (répartir 100 points sur ces 7 items)	
Administration et animation de la recherche : pilotage de la recherche (VP, direction d'institut, DAS, par exemple), participation à des instances d'évaluation (CNU, CoNRS, CSS, Hcéres, par exemple), responsabilité de dispositifs Idex ou Isite, direction de projets (ANR, Horizon Europe, ERC, CPER, PIA, France 2030, par exemple), responsabilités éditoriales dans des revues ou collections nationales et internationales.	30
Aide aux politiques publiques et expertise technique : pouvoirs publics aux niveaux européen, national et régional, entreprises, instances internationales comme FAO, OMS, etc.	1
Contribution à l'adossement d'enseignements innovants à la recherche : EUR, SFRI, etc.	10
Dissémination de la recherche : partage de connaissances avec le grand public, médiation scientifique, interface sciences et société	7
Recherche et encadrement de la recherche.	45
Valorisation, transfert, innovation.	7
Autres activités. (à préciser en une ligne maximum).	0

1.4 Prise en compte des recommandations du précédent rapport

Nous commençons par analyser la section « Recommandations à l'équipe » de l'évaluation précédente (p.41 du document).

— « *L'équipe doit penser à faire émerger une équipe originale et plus intégrée ou bien constituer deux équipes distinctes* ».
Cette recommandation avait déjà fait l'objet d'une réflexion pendant la période du rapport précédent, qui avait mené à la décision de séparer l'équipe en deux, à l'horizon 2017–2022. En 2018, après le départ rapproché de deux MCF, l'axe vérification s'était réduit à 5 permanents, dont une PU qui allait partir à la retraite en 2020. Dans ces conditions, la séparation n'était plus envisageable.
L'évolution récente de l'équipe, avec l'arrivée de 4 nouveaux membres sur la thématique vérification, nous place actuellement dans une position similaire à celle de 2016. Après discussion, nous avons renouvelé notre décision de constituer deux équipes, à l'horizon 2025/2026 (voir la trajectoire de l'équipe pour plus de détails).

— « *Impliquer les chercheurs dans les responsabilités scientifiques locales, est souhaitable* ».
Nous supposons qu'on parle ici de chercheurs CNRS. Cette recommandation a été suivie. Les chercheurs et chercheuses, ainsi que l'IR de l'équipe sont très impliqué(e)s dans l'animation scientifique du laboratoire : responsabilité d'équipe, responsabilité d'axe transversal « Complexités » ou axe MathSTIC « Catégories », implication dans les commissions de gouvernance du laboratoire (conseil de laboratoire, correspondante parité, correspondant Europe, commission comm-web. . .), responsabilité des séminaires d'équipe, etc.

— « *Une mise en perspective des enjeux théoriques, de leurs retombées scientifiques potentielles aurait été appréciée. Ne pas se donner le temps ni les moyens de cette réflexion peut fragiliser le projet, compte tenu de la proximité d'unités de recherche comme l'IRIF ou le LSV qui travaillent dans des thématiques proches. Or, le potentiel et la richesse thématiques respectivement portés par chaque axe permettent de penser qu'une identité scientifique forte peut émerger.* »

Nous avons apprécié la pertinence de cette recommandation, que nous avons essayé de mettre en oeuvre. La séparation en deux équipes permettra justement de renforcer l'identité de chaque axe, en lui donnant plus de visibilité. En même temps, chaque axe a veillé à maintenir une spécificité par rapport aux laboratoires géographiquement et thématiquement proches, à travers l'excellence dans des thématiques bien établies (automates temporisés paramétrés, λ-calcul et logique linéaire) et le développement de sujets distinctifs (réseaux de Petri, analyse de malware, vérification de blockchain, programmation différentiable, complexité...). Nous estimons que la réussite de cette stratégie se reflète dans l'attractivité des deux groupes, attestée par les nombreux recrutements sur la période.

— « *[U]ne collaboration plus poussée avec d'autres équipes du LIPN telles qu'A3 (pour Verif) ou CALIN et AOC mériterait d'être menée* ».
Dans la mesure où il est clairement insensé de forcer des collaborations, l'équipe encourage toujours les interactions avec les autres équipes lorsque celles-ci émergent naturellement. C'est le cas, par exemple, de l'équipe CALIN, avec qui nous avons deux projets ANR en commun, un sur la combinatoire des λ-termes et l'autre sur les aspects de théorie des représentations liés à la complexité. Pour ce qui concerne l'équipe AOC, il y avait une collaboration fructueuse avec l'axe vérification (publications cosignées, coencadrement de thèses), et ceci même avant que cette recommendation ne soit faite. Malheureusement, cela concernait une MCF de l'équipe AOC ayant quitté le laboratoire en 2021.
La création de l'axe transversal « Complexités », proposée justement par l'équipe LoVe, ainsi que l'organisation de la dernière journée « Apprentissage artificiel au LIPN » (une initiative de notre équipe), sont davantage de preuves de notre volonté de stimuler les échanges inter-équipes.

Le reste du document d'évaluation contient aussi des remarques assimilables à des recommandations, que nous analysons par la suite.

— p. 38 : « *Les liens avec la fédération MathSTIC sont faibles : les étoffer ne pourrait que renforcer les deux axes de l'équipe* ».
Cette recommandation a été suivie : avec la création, au printemps 2022, du 4ème axe MathSTIC « Catégories : entre Calcul et Topologie », les deux axes de l'équipe sont désormais intégrés dans la Fédération et y participent à part entière (par exemple, l'axe logique de l'équipe a bénéficié d'un post-doc de 16 mois financé par la Fédération MathSTIC).

Informations générales sur l'équipe LoVe

— p. 39 : « *La nature théorique des travaux portés par l'axe [logique] de l'équipe explique le fait que les interactions industrielles soient quasi inexistantes. En revanche, l'axe [vérification] de l'équipe pourrait développer de telles interactions* ».
L'affaiblissement progressif de l'axe vérification de 2016 à 2020 n'a permis de suivre cette recommandation que très partiellement (un projet de maturation SATT ERGANEO). Cependant, les recrutements récents dans l'axe permettront, dans le futur proche, de relancer l'équipe sur cette question.

— p. 40 : « *Il faudra veiller, pour les deux axes, à être en mesure de continuer d'attirer des doctorants sur des thématiques fondamentales complexes à appréhender* ».
Nous avons tenu compte de cette recommandation et nous estimons que la stabilité du nombre de doctorant(e)s dans l'équipe est un indicateur positif vis-à-vis de cette question.

— p. 40 : « *Le ratio F/H concernant les EC et C est clairement en défaveur des collègues féminines. Ce point est à prendre en compte dans le futur : améliorer l'attractivité de l'équipe au sein du LIPN et dans le contexte Paris Nord* ».
Cette recommandation a été suivie autant que possible. Nous faisons attention à l'équilibre F/H dans la recherche de candidat(e)s et avons eu plusieurs candidates

sur la période : 2 candidates CR (2020 et 2021) et 1 candidate PU (2021). Les deux candidates CR ont été admises au concours, mais seulement une a fini par être affectée au LIPN. La candidate PU n'a pas été sélectionnée (le poste a été attribué à un candidat de l'équipe CALIN).
Pour ce qui concerne les doctorant(e)s, l'équilibre F/H est plutôt bon pour l'axe vérification, mais est catastrophique pour l'axe logique : pas une seule candidate les derniers trois ans. Il s'agit d'un problème structurel dont les racines sont profondes, difficile à contrôler au niveau de l'équipe.

Chapitre 2
Introduction du portfolio

Avec notre portfolio nous avons souhaité montrer les différentes facettes du travail de l'équipe : sa production scientifique sous des formes variées (articles, livres, logiciels), en couvrant le plus possible le spectre de ses thématiques, et l'investissement conséquent de ses membres en termes d'organisation de manifestations scientifiques de portée internationale. Cette volonté nous a amené à sélectionner 3 éléments par axe, plus un 7ème élément commun :

- pour l'axe logique, 2 articles et 1 livre montrant l'expertise de niveau international en λ-calcul, logique linéaire et applications à la théorie des langages de programmation ;
- pour l'axe vérification, 2 articles, l'un donnant un exemple de résultat de nature fondamentale, l'autre présentant une application très originale des recherches de l'équipe. Le troisième élément consiste à exposer la production logicielle de l'axe, un point mis en valeur par la contribution d'un ingénieur de recherche.
- L'élément commun est l'organisation d'événements scientifiques internationaux, dans lesquels toute l'équipe est fréquemment impliquée : conférences internationales (FSCD/IJCAR 2020, PetriNets 2020, PetriNets 2021, Microservices 2022, ETAPS 2023), écoles de recherche (la série Caleidoscope 2019–2021, l'école sur la Théorie de la Complexité Géométrique en 2022, l'EPIT 2023).

Chapitre 3

Auto-évaluation de l'équipe LoVe

Sommaire

3.1 Domaine 2. Attractivité 246

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche 246

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels 248

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs 250

3.2 Domaine 3. Production scientifique 251

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité. 251

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels. 252

3.2.3 Publications 253

3.3 Domaine 4. Inscription des activités de recherche dans la société . 273

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique. 273

3.3.2 Référence 2. L'équipe développe des produits à destination du monde culturel, économique et social. 273
3.3.3 Référence 3. L'équipe partage ses connaissances avec le grand public et intervient dans des débats de société. 273

@

3.1 Domaine 2. Attractivité

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche

L'équipe accueille des chercheurs et chercheuses bien intégré(e)s et reconnu(e)s dans leurs communautés scientifiques, comme en témoignent plusieurs invitations à des conférences et workshops, participations à des comités de programmes et/ou de pilotage de conférences et de nombreuses responsabilités dans des instances de pilotage ou d'évaluation de la recherche.

Particulièrement reconnues, au niveau international, sont l'expertise en applications de la logique linéaire aux langages de programmation et en complexité implicite (pour l'axe logique) et l'expertise en model checking et spécialement en vérification d'automates temporisés paramétrés (pour l'axe vérification).

Dans la suite, nous allons apporter les éléments permettant de montrer le rayonnement de l'équipe, en termes d'invitations à des colloques et congrès, de participation à des instances de pilotage ou évaluation de la recherche et de prix et distinctions. Les aspects liés à l'organisation de manifestations scientifiques sont détaillés dans le portfolio.

Invitations à des colloques/congrès. Les membres de l'équipe sont invité(e)s régulièrement à des manifestations scientifiques, le plus souvent à caractère international, liées à leurs thématiques de recherche. Nous donnons ici une sélection de ces invitations.

Auto-évaluation de l'équipe LoVe

Exposés invités à des conférences internationales :

— *International Forum on Cybersecurity* (FIC 2018), Lille.

Exposés invités à des workshop internationaux :

— *Workshop on Logic, Language, Information and Computation* (WoLLIC 2022), Iasi, Roumanie.
— *Forum Blockchain* affilié à la conférence *Business Process Management* (BPM 2022), Munster, Allemagne.
— *Workshop AIPLAN, affilié à Neurips 2021*, en ligne.
— *International Workshop on Trends in Linear Logic* (TLLA 2021), en ligne.
— *Workshop BIRS "Tangent Categories and their Applications"* 2021, en ligne.
— *Workshop Lean 2021*, en ligne.
— *International Workshop on Higher-Order Rewriting* (HOR 2019), Dortmund, Allemagne.
— *International Workshop on Functional High-Performance and Numerical Computing* (FHPNC 2019), Berlin, Allemagne.
— *Joint International Workshop on Structures and Deduction and Trends in Linear Logic* (SD/TLLA 2019), Dortmund, Allemagne.
— *International Workshop on Logic and Computational Complexity* (LCC 2019), Patras, Grèce.
— *Cologne-Twente workshop on Graphs and Combinatorial Optimization*, 2018, Paris.

— *International Workshop on Intersection Types and Related Systems* (ITRS 2018), Oxford, Royaume-Uni.
— tutoriel invité à CRISIS 2018, Arcachon.
— *Langages et Outils pour la Fiabilité logicielle* (OSIS 2017), Paris.
— *International Workshop on Logic and Computational Complexity* (LCC 2017), Rejkyavik, Islande.

Participation sur invitation à workshop thématiques internationaux :
— *New Challenges in Program Semantics*, 2022, Lorentz Center, Leyde, Pays-Bas.
— (2 membres invités) *Shonan Meeting on Higher Order Complexity and Applications*, 2019, Shonan Village, Japon.
— *Analysis of Autonomous Mobile Collectives in Complex Physical Environments*, 2019, Dagstuhl, Allemagne.

Cours/exposés invités à des écoles de recherche :
— (3 membres invités) *Linear Logic Winter School*, 2022, Marseille.

Les membres de l'équipes ont aussi été invité(e)s à plusieurs manifestations à caractère national (écoles, workshops, journées. . .) dont on ne donne pas le détail ici.

Responsabilités éditoriales. L'équipe est régulièrement impliquée dans les comités de programme de conférences ou de workshop internationaux :
— 1 comité éditorial de journal international (ToPNoC : *Transactions on Petri Nets and Other Models of Concurrency*)
— Édition d'un numéro spécial de la revue internationale *Mathematical Structures in Computer Science* [ED-2].
— 2 PC chair de conférences internationales avec actes (FORMATS 2019 [ED-4], PetriNets 2022 [ED-5]).
— 32 PC de conférences internationales (en tant que membres).
— 2 PC chair de workshop internationaux avec actes (FMICS-AVoCS 2017 [ED-1], DICE/FOPARA 2019 [ED-3]).
— Édition du track « *Specification, Analysis and Verification of Systems, Security* » de la newsletter de l'*Association for Logic Programming*.

Participation à des instances de pilotage ou évaluation de la recherche. L'équipe est très investie au niveau du pilotage de la recherche, avec plusieurs membres ayant des responsabilités administratives importantes.

Administration de la recherche :
— présidence de l'Université Sorbonne Paris Nord (depuis décembre 2020) ;
— vice-présidence formation et vie universitaire, USPN (2020–2021) ;
— vice-présidence systèmes d'information, USPN (2016–2020) ;
— direction du LIPN (2012–2017) ;
— CoNRS, section 06 (membre et secrétaire scientifique, depuis janvier 2021) et CID 53 (membre et membre de bureau, depuis septembre 2021) ;
— Adjointe au Directeur Scientifique Référent (ADSR) du CNRS pour le site Normandie (depuis décembre 2019) ;
— Conseil Scientifique de l'INS2I du CNRS (membre élu, 2018–2020) ;
— CNU section 27 (2 membres, depuis 2016 et 2020)
— Conseil Académique USPN (membre, depuis 2016)
— Conseil d'Institut IG (membre, 2016–2020)
— direction du DIM RFSI (2018–2022)
— direction de l'Université Numérique Nationale IEL (IUTEnLigne, 2017–2020)
— direction IREM Paris Nord (sur toute la période)

Participation à comités de pilotage de conférences ou workshop :
— 5 SC de conférences internationales (VECOS 2017–2022, FSCD 2019–2022, FSCD 2020–2023, PetriNets 2022–présent, ETAPS 2022–présent) ;

— 2 SC de workshop internationaux (FSCD 2017–présent, Syncop 2017–présent) ;
— 1 SC de conférence nationale (JFLA 2013–présent).

Participation à des comités d'expertise ou d'évaluation de la recherche :
— 3 comités HCERES (1 présidence, 1 vice-présidence, 1 membre) ;
— 1 comité d'évaluation Inria, équipes affiliées au thème "Security and Confidentiality" (membre, 2019) ;
— 3 comités d'évaluation de laboratoires en Tunisie (2019) ;
— 2 comités ANR (membre CE48, appels 2020–2021 et 2021–2022) ;
— 4 jurys IUF (membre, junior 2020 et 2021, senior 2019 et 2022) ;
— jury de l'agrégation externe d'informatique (membre, depuis 2022) ;
— instance représentante du CNRS à l'ECSO (European Cyber Security Organisation, membre, 2017–2022).

Un membre de l'équipe a été co-responsable du groupe de travail « Structures Formelles pour le Calcul et les Preuves » (SCALP) du GDR Informatique Mathématique (2018–2023).

Prix et distinctions.
— prix « Paolo Gentilini » pour la logique mathématique et ses applications à l'informatique, attribué par l'Association Italienne de la Logique et ses Applications (AILA), 2020.
— *Distinguished Paper* POPL 2020 (CORE rang A*).
— *Best Paper Award* ICECCS 2018 (CORE rang A).
— *Best Paper Award* FORTE 2017 (CORE rang A).
— *Best Paper Award* FMICS-AVoCS 2017.

Organisation de manifestations scientifiques. L'équipe s'investit beaucoup dans l'organisation de manifestations scientifiques, y compris d'événements internationaux de taille relativement importante. Cet aspect est détaillé dans le portfolio.

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels

Auto-évaluation de l'équipe LoVe

L'équipe a su être attractive à la fois sur les postes permanents et non-permanents. Sur la période d'évaluation, nous avons :
— 1 recrutement de PU ;
— 1 recrutement de MCF ;
— 2 recrutements de CR CNRS ;
— 1 recrutement d'IR CNRS ;
— 1 mutation entrante de MCF ;
— 4 recrutements de post-docs (plus un 5ème en janvier 2023) ;
— 15 doctorant(e)s ont commencé leur thèse, dont 5 ont déjà soutenu.

L'équipe est particulièrement proactive vis-à-vis de la recherche de candidats. Pour les postes CR CNRS, nous contactons chaque année les candidat(e)s pressenti(e)s et les incitons à déposer un dossier en mentionnant le LIPN comme unité d'accueil. Nous mettons à leur disposition une archive de dossiers « modèles » de candidature CR CNRS, déposés par des personnes qui ont été recrutées dans le passé récent. Nous proposons ensuite une relecture des dossiers : chaque dossier est relu par au moins deux membres de l'équipe et les commentaires sont envoyés aux candidat(e)s. Pour les candidat(e)s auditionné(e)s, nous proposons une simulation d'audition.

Le nombre de candidat(e)s est très variable d'une année sur l'autre (entre 3 et 8), mais le mécanisme donne des bons résultats : sur les quatre campagnes de recrutement qui ont eu lieu depuis sa mise en place en 2020, nous avons recruté une CR en 2020, une deuxième CR en 2021 qui a été affectée à une autre unité, et nous avons eu des candidats auditionnés à toutes les campagnes sauf celle de 2022.

Post-doctorants recrutés par l'équipe. L'approche proactive est adoptée aussi dans le cas des post-docs, pour lesquels l'équipe a des difficultés à trouver des financements réguliers. Nous cherchons ainsi des candidats potentiels et leur proposons de soumettre des dossiers à des appels très compétitifs (par exemple les bourses Marie-Sklodowska Curie ou similaire), qui ont l'avantage d'être totalement financés par des sources externes (comme l'Union Européenne, la Région, etc.) et indépendantes du laboratoire, contrairement aux projets de type ANR, etc. Nous proposons un mécanisme de relecture des dossiers, similaire à celui mentionné plus haut pour les candidatures CNRS. Le taux de succès ici est plus faible, mais nous avons réussi à recruter un post-doc de 24 mois sur financement COFUND (UE/Région IdF), qui a commencé en janvier 2023 (hors période d'évaluation donc).

La liste des post-doctorants que l'équipe a accueilli sur la période est donnée ci-dessous, avec la source de financement :

— Maxime Lucas (2020, projet CoHOp, DIM RFSI)
— Federico Olimpieri (2020–2021, projet ANR CoGITARe)
— Morgan Rogers (2021–2023, Fédération MathSTIC)
— Thomas Rubiano (2021, projet CoHOp, DIM RFSI)

Doctorants encadrés par l'équipe. L'équipe bénéficie des allocations de l'École Doctorale Galilée (USPN) pour recruter des doctorant(e)s, et attire aussi régulièrement des étudiants normaliens (provenant, par exemple, du MPRI), avec l'avantage de pouvoir utiliser les financements de l'ENS.

Nous cherchons également des financements de thèse en dehors des sources habituelles et essayons de répondre aux appels que nous estimons pertinents. Par exemple, nous avons obtenu une bourse de thèse Vinci, financée par l'Université Franco-Italienne, pour une cotutelle avec l'Université Roma Tre.

La liste des doctorant(e)s que l'équipe a accueilli sur la période est donnée ci-dessous, avec la source de financement, la durée et, si disponible, le devenir du doctorant/de la doctorante :

— Baptiste Chanus (allocation USPN, depuis 2022).
— Baptiste Fiévet (bourse internationale CNRS, co-direction Aarhus, Danemark, depuis 2022).
— Simon Mirwasser (bourse ENS, depuis 2022).
— Adrien Ragot (bourse Vinci, cotutelle avec Università Roma Tre, Italie, depuis 2022).
— Nour Elhouda Souid (allocation USPN, depuis 2021).
— Aloÿs Dufour (allocation USPN, depuis 2021).
— Ulysse Léchine (allocation USPN, depuis 2021).
— Nicolas Münnich, (ANR CoGITARe, depuis 2020).
— Boris Eng (allocation USPN + ATER, depuis 2019)
— Ikram Garfatta (2019–2022, 45 mois, ATER USPN).
— Axel Kerinec, (bourse ENS, depuis 2019).
— Davide Barbarossa (allocation USPN, cotutelle avec Università Roma Tre, Italie, 2018–2021, 39 mois, post-doc Università di Bologna).
— Jawher Jerray (2018–2021, 39 mois, post-doc Inria).
— Nguyen Le Thanh Dung (bourse ENS, 2018–2021, 39 mois, post-doc ENS Lyon).
— Daniel Torres (allocation USPN, 2018–2021, 39 mois, IR Inria).
— Hiba Ouni (allocation USPN, 2017–2019, 27 mois).
— Mathias Ramparison (allocation USPN, 2016–2019, 35 mois, MCF Université de Grenoble).
— Antoine Kaszczyc (allocation USPN + ATER, 2015-2019, 51 mois)
— Hoang Gia Nguyen (allocation USPN, 2015–2018, 35 mois, postdoc Inria).
— Thi Thanh Huyen Nguyen (bourse Vietnam, 2015–2018, 44 mois (maternité pendant la thèse), enseignante au Vietnam).
— Alice Pavaux (allocation USPN, 2014–2017, 39 mois)

— Luc Pellissier (bourse ENS, 2014–2017, 39 mois, MCF UPEC).

Au total, 12 thèses ont été soutenues pendant la période d'évaluation.

On peut constater que la durée moyenne des thèses dans l'équipe est d'environ 40 mois, ce qui est proche des 39 mois visés par l'équipe, correspondant à un début en septembre de l'année N et une soutenance en décembre de l'année $N+3$.

Les membres de l'équipe participent aussi au co-encadrement de doctorant(e)s non inscrit(e)s à USPN, dans le cadre de co-tutelles ou, plus souvent, de co-directions de thèses. Nous avons actuellement 6 doctorant(e)s dans cette situation (1 cotutelle avec Melbourne, Australie, et 5 co-directions avec des établissements français). Ces doctorant(e)s ne sont pas mentionné(e)s dans la liste ci-dessus mais ils/elles bénéficient du statut de membre associé de l'équipe et visitent régulièrement leurs encadrants au LIPN (par exemple, le doctorant en cotutelle avec Melbourne est au LIPN depuis 2 ans).

Invitations de chercheurs/chercheuses. L'équipe invite régulièrement des chercheurs provenant de laboratoires nationaux ou internationaux, afin d'animer la vie scientifique de ses membres et de développer des collaborations scientifiques. Ces invitations sont financées par plusieurs dispositifs : Prof/MCF invités de l'Université Sorbonne Paris Nord, financements sur projet, Fédération MathSTIC, ressources propres LIPN... La liste ci-dessous est limitée aux chercheurs invités au minimum deux semaines sur la période, et ne concerne que des invités internationaux (d'autres chercheurs/chercheuses, typiquement en provenance de laboratoires français, ont été invité(e)s sur des périodes plus courtes) :

— Wojciech Penczek, 2022, Prof invité USPN + projet MoSART (5 semaines)
— Teofil Sidoruk, 2022, projet MoSART (4 semaines)
— Jaco van de Pol, 2022, Prof invité USPN (5 semaines)
— Lars M. Kristensen, 2021, Prof invité USPN (1 mois)
— Jean-Simon Lemay, 2021, LIPN (2 semaines)
— Jaco van de Pol, 2021, projet TrAVAIL (1 semaine)
— Simona Ronchi Della Rocca, 2020, Prof invitée USPN (2 semaines)
— Wojciech Penczek, 2019-2021, projet PARTIES (2 semaines)
— Christian Ikenmeyer, 2019, Prof invité Math-Stic (2 semaines)
— Michal Knapik, 2019, projet PARTIES (2 semaines)
— Teofil Sidoruk, 2019, projet PARTIES (2 semaines)
— Jun Sun, 2018, MCF invité USPN (2 semaines)
— Tarmo Uustalu, 2019, projet CoGITARe (2 semaines)
— Jaco van de Pol, 2018-2019, projet PAMPAS (4 semaines)
— Wojciech Penczek, 2017, Prof invité USPN (2 semaines)

Auto-évaluation de l'équipe LoVe

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs

L'équipe encourage ses membres à répondre aux appels à projets pertinents vis-à-vis de leur situation (début de carrière, existence d'une thématique émergente, etc.), tout en restant ancrée à la conviction que le financement de la recherche sur ressources propres du laboratoire est primordial (les projets sont « la cerise sur le gâteau », ils ne doivent pas devenir « le gâteau »). Dans ce sens, nous soutenons vivement la mutualisation des fonds sur projet entre les équipes du LIPN.

Les nouveaux recrutés de l'équipe répondent de manière presque systématique aux appels à projets internes aux tutelles (BQR USPN, projets INS2I), avec un excellent taux de réussite (à une exception près, tous les MCF/CR recrutés sur la période ont eu deux projets INS2I, BQR ou similaire, pour un total de 6 projets).

L'étape suivante est la participation aux appels JCJC de l'ANR, sur lesquels l'équipe a aussi un bon taux de réussite (3 propositions sur la période, 2 financées, 1 en liste complémentaire, resoumise à l'appel 2023).

Les membres plus expérimentés et avec des propositions suffisamment solides répondent aux appels multipartenaires de l'ANR (2 projets de ce type se sont achevés pendant la période, tous les deux commencés en 2014), voire aux appels européens (3 soumissions ERC sur la période, Starting, Consolidator et Synergy, les trois n'ont pas passé la première phase). Les membres de l'équipe participent aussi aux projets ANR multipartenaires en tant que responsables de site (1 proposition financée sur la période).

Au niveau international, sur la période d'évaluation l'équipe a obtenu le financement de 5 projets de coopération (PHC, IEA CNRS, INS2I, etc.), dont un a abouti à une demande de création d'IRN CNRS, soumise en 2022 et approuvée en 2023.

L'équipe répond également, avec succès, à des projets régionaux, avec 3 propositions financées par le DIM RFSI (Région Île-de-France). Récemment, dans le contexte du renforcement de ses rapports avec l'industrie, l'équipe a aussi commencé à être porteuse de projets de maturation de startup (SATT ERGANEO).

Le résumé des projets portés par l'équipe sur la période 2017–2022 est donné ci-dessus :

Projets avec financement de personnel (montant supérieur ou égal à 50k euros) :

- 3 ANR multipartenaires ;
- 2 ANR JCJC ;
- 1 DIM RFSI ;
- 1 maturation SATT ERGANEO.

Projets sans financement de personnel (montant inférieur à 50k euros) :

- 7 CNRS/INS2I ;
- 4 BQR ou autre (**MSH!** Paris Nord) ;
- 2 DIM RFSI ;
- 1 PHC ;
- 2 autres (Thomas Jefferson Fund (USA), Institut Français au Danemark).

3.2 Domaine 3. Production scientifique

Nous fournissons ci-dessous un tableau de synthèse des publications de l'équipe, organisées par année et par typologie :

	2017	2018	2019	2020	2021	2022	2023	*Total*
Livres	1	1	0	0	0	1	0	*3*
Chapitres de livre	3	3	1	0	2	1	0	*10*
Revues internationales	5	7	7	4	10	5	3	*41*
Autres revues	0	1	1	1	1	1	0	*5*
Conférences internationales	28	23	24	9	16	16	4	*120*
Autres communications	13	6	7	1	3	7	0	*37*
Total	*50*	*41*	*40*	*15*	*32*	*31*	*7*	*216*

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité.

L'équipe a obtenu des résultats de premier plan dans ses domaines de recherche, mentionnés plus haut dans la section « thématiques scientifiques et leurs enjeux » et, pour certains, détaillés dans le portfolio. Pour l'axe logique, ces recherches s'appuient sur des méthodes d'ordre sémantique (spécialement de type catégorique et/ou reliées à la logique linéaire) pour étudier les programmes et leur complexité, tandis que pour l'axe vérification elle reposent sur différentes approches pour l'analyse d'automates

temporisés paramétrés et de réseaux de Petri (algorithmes de model checking, logiques temporelles, etc.), quasiment toujours accompagnés d'implémentations permettant de reproduire les résultats.

Dans le domaine de recherche de l'équipe LoVe, que ce soit pour l'axe logique ou l'axe vérification, les conférences (avec comité de lecture) sont le moyen de publication privilégié. Dans ces thématiques, un article accepté dans une conférence très sélective équivaut à un article publié dans une revue prestigieuse. Le fait que l'équipe publie régulièrement dans les meilleures conférences est donc un signe clair de la solidité de ses résultats. Cela explique aussi pourquoi l'équipe publie trois fois plus dans des conférences que dans des journaux.

Les conférences « phare » dans le domaine des deux axes de l'équipe sont les suivantes (le nombre d'articles publiés par l'équipe pendant la période d'évaluation est noté entre parenthèses) :

axe logique : LICS (2+1), POPL (5+1) ;

axe vérification : CAV (3), AAMAS (1+1).

Ces conférences sont considérées de rang maximal par tous les classements disponibles (rang A* dans CORE, rang A dans ERA, etc.). Le « +1 » signale des publications parues ou à paraître, au moment de la rédaction de ce rapport, en 2023. Bien que hors période d'évaluation, ces publications reflètent des travaux développés en 2022 et sont indicatives de la qualité et régularité de la production de l'équipe, qui publie systématiquement autour de 2 articles par an dans des conférences de rang maximal.

Nous notons de suite le nombre de publications dans d'autres conférences considérées comme de premier rang par leurs communautés, mais dont le statut est moins consensuel parmi les différents classements (variable entre A et B) :

axe logique : CSL (2), FoSSaCS (1), ICALP (3) ;

axe vérification : ICECCS (13), ICSOC (2), TACAS (2).

Nous signalons aussi les conférences FSCD (5 publications) et CPP (2 publications) pour l'axe logique, et FORMATS (3 publications) et PetriNets (+2 publications) pour l'axe vérification, qui ne sont pas très reconnues dans les classements officiels à cause de leur caractère spécialisé mais qui ont une excellente réputation au sein de leur domaine.

L'axe vérification de l'équipe produit également des logiciels, qui s'appuient sur les résultats théoriques et les valident en les réalisant concrètement. La production logicielle de l'équipe est détaillée dans le portfolio.

Auto-évaluation de l'équipe LoVe

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels.

Le nombre total d'ETPT de l'équipe sur la période, limité au membres permanents et tenant compte de la présence partielle de certains membres, est 67.69. Cela donne

— $39/67.14 = 0.58$ articles de revue internationale par personne par an ;
— $117/67.69 = 1.74$ articles de conférence internationale par personne par an.

Comme nous avons signalé dans la section précédente, la disparité entre nombre d'articles de revue et nombre d'articles de conférence s'explique par l'importance considérable que ces dernières jouent dans les communautés scientifiques dont les membres de l'équipe font partie.

Il est utile de souligner que, si chaque article est compté un nombre de fois égal au nombre de membres de l'équipe l'ayant co-signé, les chiffres deviennent

— 0.74 articles de revue internationale par personne par an ;
— 2.16 articles de conférence internationale par personne par an.

(Ceci dérive de l'existence de 7 papiers de revue et 15 papiers de conférence à double signature, ainsi que 1 papier de revue et 1 papier de conférence à triple signature).

Le fait que l'équipe dispose d'un IR CNRS est un atout important en termes de production scientifique, en particulier pour l'axe vérification, où sa présence permet le développement approfondi des logiciels mettant en pratique les résultats publiés, auxquels d'ailleurs il contribue.

La production scientifique de l'équipe repose aussi sur la contribution de ses doctorant(e)s, qui sont appelé(e)s à y participer de façon active. Sur 12 doctorant(e)s ayant soutenu leur thèse pendant la période d'évaluation, 11 ont publié au moins un article, le plus souvent avec leur directeur ou directrice de thèse et/ou d'autres membres de l'équipe (6 articles de revue et 44 articles de conférence), mais aussi indépendamment, c'est-à-dire seul ou avec un co-auteur hors LIPN (6 articles de conférence, tous d'ailleurs de haute qualité : LICS, 3 ICALP, FSCD, CSL). En total, 20% des papiers sur revue et environ 45% des papiers de conférence ont été (co)signés par un(e) doctorant(e).

À l'exception d'un post-doctorant ayant interrompu son contrat pour être recruté en entreprise, les post-doctorants accueillis par l'équipe ont aussi tous publié au moins un papier avec leur affiliation LIPN (pour cela, nous comptons aussi les publications de notre post-doctorant le plus récent, soumises en 2022 mais parues en 2023).

Un petit nombre de membres de l'équipe (entre 15 et 20% selon le moment auquel on compte les effectifs) n'ont participé que peu, ou pas, à la production de recherche, en dépit d'un rattachement à temps complet au LIPN le long de toute la période d'évaluation. Chaque cas est différent mais ces situations se justifient essentiellement par la prise de responsabilités administratives très lourdes (présidence ou vice-présidence d'université, direction d'instances académiques, etc.) et/ou une volonté de réorienter sa carrière vers des directions alternatives à la recherche (administration, engagement dans les services numériques pour l'université, activités extra-académiques, etc.). Il s'agit en tout cas de choix personnels où l'équipe ne joue pas de rôle : la plupart des membres dans cette situation auront quitté l'équipe, au moins partiellement (disponibilité, temps partiel, etc.), à l'horizon du prochain quinquennal, s'ils ne l'ont pas déjà fait.

L'équipe peut jouer, et joue un rôle dans le cas de membres s'éloignant de la recherche sans avoir d'autre véritable projet professionnel. Un petit nombre de situations de ce type s'est produit. La stratégie de l'équipe dépend du cas spécifique mais pour l'essentiel elle repose sur l'implication des membres en difficulté dans la dynamique du groupe, dans le but de fournir des pistes potentielles de reprise d'activité. Les personnes en question sont actuellement en voie de sortie de leur impasse, et l'équipe continuera à veiller à ce que des situations similaires soient repérées le plus tôt possible.

3.2.3 Publications

Les références
- comportant « * » sont acceptées en 2022 et à paraître.
- indiquées sur fond gris sont les productions phare de chaque catégorie.
- en gras constituent les 20% les meilleures de la production scientifique de chaque catégorie.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2023 [RI-41] C. Olarte, E. Pimentel et C. Rocha. A rewriting logic approach to specification, proof-search, and meta-proofs in sequent systems. *Journal of Logical and Algebraic Methods in Programming*, 130 :100827 (37 pages), 2023.

[RI-40] M. Rogers. Toposes of topological monoid actions. *Compositionality*, 5(1) :49 pages, 2023.

[RI-39*] M. Rogers et R. Hemelaer. Geometric morphisms between toposes of monoid actions : factorization systems. *Theory and Applications of Categories*, page 41 pages, 2023.

2022 [**RI-38**] S. Boldo, F. Clément, F. Faissole, V. Martin et M. Mayero. A Coq Formalization of Lebesgue Integration of Nonnegative Functions. *Journal of Automated Reasoning*, 66(2) :175–213, 2022.

[RI-37] K. H. T. Dam et T. Touili. Extracting malicious behaviours. *Int. J. Inf. Comput. Secur. 17(3/4) : 365-404*, 2022.

[RI-36] G. Della Penna, B. Intrigila et G. Manzonetto. Addressing Machines as models of lambda-calculus. *Logical Methods in Computer Science*, 18(3) :10 :1–10 :31, 2022.

[RI-35] T. Touili et X. Ye. LTL Model Checking of Self Modifying Code. *Formal Methods in System Design*, 60 :195–227, 2022.

[RI-34] B. Xavier, C. Olarte et E. Pimentel. A linear logic framework for multimodal logics. *Mathematical Structures in Computer Science*, 29 :29 pages, 2022.

2021 [RI-33] E. André, E. Coquard, L. Fribourg, J. Jerray et D. Lesens. Parametric schedulability analysis of a launcher flight control system under reactivity constraints. *Fundamenta Informaticae*, 182(1) :31–67, 2021.

[RI-32] E. André, D. Lime et M. Ramparison. Parametric updates in parametric timed automata. *Logical Methods in Computer Science*, 17(2) :Article 13 (67 pages), 2021.

[RI-31] E. André, D. Lime, M. Ramparison et M. Stoelinga. Parametric analyses of attack-fault trees. *Fundamenta Informaticae*, 182(1) :69–94, 2021.

[RI-30] E. André, H. G. Nguyen, L. Petrucci et J. Sun. Distributed parametric model checking timed automata under non-Zenoness assumption. *Formal Methods in System Design*, 59(1) :253–290, 2021.

[RI-29] A. Boussif, M. Ghazel et K. Klai. A semi-symbolic diagnoser for fault diagnosis of bounded labeled petri nets. *Asian Journal of Control*, 23(2) :648–660, 2021.

[**RI-28**] F. Breuvart, U. Dal Lago et A. Herrou. On Higher-Order Probabilistic Subrecursion. *Logical Methods in Computer Science*, 17(4) :Article 25 (35 pages), 2021.

[**RI-27**] C. Coti, L. Petrucci, C. Rodriguez et M. Sousa. Quasi-Optimal Partial Order Reduction. *Formal Methods in System Design*, 57 :3–33, 2021.

[RI-26] A. Felty, C. Olarte et B. Xavier. A focused linear logical framework and its application to metatheory of object logics. *Mathematical Structures in Computer Science*, 31(3) :312–340, 2021.

[RI-25] S. Guerrini et A. Masini. Proof Nets for Classical Logic. *Notre Dame Journal of Formal Logic*, 62(2) :303 – 343, 2021.

[RI-24] J.-B. Joinet et T. Seiller. From abstraction and indiscernibility to classification and types : revisiting hermann weyl's theory of ideal elements. *Kagaku Tetsugaku*, 53(2) :65–93, 2021.

2020 [RI-23] E. André, B. Delahaye et P. Fournier. Consistency in Parametric Interval Probabilistic Timed Automata. *Journal of Logical and Algebraic Methods in Programming*, 110 :100459 (45 pages), 2020.

[RI-22] A. Kerinec, G. Manzonetto et M. Pagani. Revisiting Call-by-value Böhm trees in light of their Taylor expansion. *Logical Methods in Computer Science*, 16(3) :6 :1–26, 2020.

[RI-21] L. T. D. Nguyen. Unique perfect matchings, forbidden transitions and proof nets for linear logic with Mix. *Logical Methods in Computer Science*, 16(1) :27 :1–31, 2020.

[RI-20] A. Pommellet et T. Touili. LTL Model Checking for Communicating Concurrent Programs. *Innovations in Systems and Software Engineering*, 16 :161–179, 2020.

2019 [RI-19] E. André. What's decidable about parametric timed automata ? *International Journal on Software Tools for Technology Transfer*, 21(2) :203–219, 2019.

[RI-18] E. André, M. Knapik, D. Lime, W. Penczek et L. Petrucci. Parametric Verification : An Introduction. *Transactions on Petri Nets and Other Models of Concurrency*, 14 :64–100, 2019.

[**RI-17**] B. Intrigila, G. Manzonetto et A. Polonsky. Degrees of extensionality in the theory of Böhm trees and Sallé's conjecture. *Logical Methods in Computer Science*, 15(1) :Article 6 (36 pages), 2019.

[**RI-16**] M. J. Knapik, E. André, L. Petrucci, W. Jamroga et W. Penczek. Timed ATL : Forget Memory, Just Count. *Journal of Artificial Intelligence Research*, 66 :197–223, 2019.

[RI-15] G. Manzonetto, M. Pagani et S. Ronchi Della Rocca. New Semantical Insights Into Call-by-Value λ-Calculus. *Fundamenta Informaticae*, 170(1-3) :241–265, 2019.

[**RI-14**] G. Manzonetto, A. Polonsky, A. Saurin et J. Grue Simonsen. The fixed point property and a technique to harness double fixed point combinators. *Journal of Logic and Computation*, 29(5) :831–880, 2019.

[**RI-13**] T. Seiller. Interaction graphs : exponentials. *Logical Methods in Computer Science*, 15(3) :25 :1–58, 2019.

2018 [RI-12] E. André et S.-W. Lin. The language preservation problem is undecidable for parametric event-recording automata. *Information Processing Letters*, 136 :17–20, 2018.

[**RI-11**] F. Breuvart, G. Manzonetto et D. Ruoppolo. Relational Graph Models at Work. *Logical Methods in Computer Science*, 14(3) :Article 2 (43 pages), 2018.

[RI-10] C. Choppy, J. Desel et L. Petrucci. Specialisation and Generalisation of Processes. *Enterprise Modelling and Information Systems Architectures*, 13(1) :35–46, 2018.

[RI-9] C. Fouqueré et M. Quatrini. Study of Behaviours via Visitable Paths. *Logical Methods in Computer Science*, 14(2) :Article 7 (47 pages), 2018.

[RI-8] D. Mazza. The True Concurrency of Differential Interaction Nets. *Mathematical Structures in Computer Science*, 28(7) :1097–1125, 2018.

[RI-7] T. Seiller. A correspondence between maximal abelian sub-algebras and linear logic fragments. *Mathematical Structures in Computer Science*, 28(1) :77–139, 2018.

[**RI-6**] T. Seiller. Interaction graphs : nondeterministic automata. *ACM Transactions on Computational Logic*, 19(3) :Article 21 (24 pages), 2018.

2017 [RI-5] E. André, T. Chatain et C. Rodriguez. Preserving Partial-Order Runs in Parametric Time Petri Nets. *ACM Transactions on Embedded Computing Systems*, 16(2) :43 :1–26, 2017.

[RI-4] J. E. Avery, J. Y. Moyen, P. Růžička et J. G. Simonsen. Chains, Antichains and Complements in Infinite Partition Lattice. *Algebra Universalis*, 79 :Article 37 (21 pages), 2017.

[RI-3] A. Bourouis, K. Klai, N. Ben Hadj-Alouane et Y. El Touati. On the Verification of Opacity in Web Services and Their Composition. *IEEE Transactions on Services Computing*, 10(1) :66–79, 2017.

[RI-2] L. Fippo Fitime, O. Roux, C. Guziolowski et L. Paulevé. Identification of bifurcation transitions in biological regulatory networks using Answer-Set Programming. *Algorithms for Molecular Biology*, 12(12 :19) :12 :19 (14 pages), 2017.

[RI-1] D. Mazza. Infinitary Affine Proofs. *Mathematical Structures in Computer Science*, 27(5) :581–602, 2017.

Articles dans des revues nationales ou internationales

2022 [RE-5] M. Kerjean, F. Le Roux, P. Massot, M. Mayero, Z. Mesnil, S. Modeste, J. Narboux et P. Rousselin. Utilisation des assistants de preuves pour l'enseignement en L1 ; retours d'expériences. *Gazette de la SMF*, 174 :22 pages, 2022.

2021 [RE-4] J. Arias, B. Barbot, F. Hulin-hubard, F. Kordon et L. Petrucci. CosyVerif : an online verification platform. *Petri Net Newsletter*, 87 :5–7, 2021.

2020 [RE-3] S. Guerrini, S. Martini et A. Masini. Quantum Turing Machines : Computations and Measurements. *Applied Sciences*, 10(16) :Article 551 (33 pages), 2020.

2019 [RE-2] S. Alayrangues, E. Beffara, S. Daniel, C. Declercq, A. Heam, J.-V. Loddo, P. Marquet, J.-C. Masseron, A. Meyer, M. More, F. Neny, V. Pantaloni, G. Perrin, C. Prouteau, G. Saliba, S. R. Schwer, F. Tarissan, C. Ubera, J.-M. Vincent et E. Volte. Une analyse des exercices d'algorithmique et de programmation du brevet 2017. *Repères IREM*, 116 :47–81, 2019.

2018 [RE-1] A. Boussif, M. Ghazel et K. Klai. Fault Diagnosis of Discrete-Event systems Based on the Symbolic Observation Graph. *International Journal of Critical Computer-Based Systems*, 8(2) :141–168, 2018.

Chapitres de livre

2022 [CL-10] S. Schwer. *La Géométrographie : vers un langage de programmation des algorithmes de tracés des figures géométriques*, N. Chevalarias, éditeur, *Géométries d'hier à demain : pratiques, méthodes, enseignement*, Chapitre 12, pages 223–240. Actes du 23e Colloque de la commission Inter-IREM Epistémologie et Histoire des mathématiques, Poitiers, France, 23-25 mai 2019. IREM&S de Poitiers, 2022.

2021 [CL-9] C. Fouqueré, J.-J. Pinto et M. Quatrini. Language Interaction and Schizophrenia. M. Amblard, M. Musiol et M. Rebuschi, éditeurs, *(In)coherence of Discourse - Formal and Conceptual Issues of Pathological Discourse*, volume 10 de *Language, Cognition, and Mind*, pages 91–115. Springer, 2021.

[CL-8] S. Schwer. *Quelques réflexions sur l'évolution des systèmes de numération en langue : l'exemple de quelques parlers marchois*, L. Esher, M. Guérin, N. Quint et M. Russo, éditeurs, *Le Croissant linguistique : entre oc, oïl et francoprovençal. Des mots à la grammaire, des parlers aux aires*, pages 313–334. L'Harmattan, 2021.

2019 [CL-7] S. Schwer. *Revenir aux mathématiques par les récréations : l'exemple de Henri Auguste Delannoy (1833-1915)*, N. Chevalarias, M. Gandit, M. Morales et D. Tournès, éditeurs, *Mathématiques récréatives. Eclairages historiques et épistémologiques*, pages 85–112. EDP Sciences, 2019.

2018 [CL-6] J. Arias, R. Marczak et M. Desainte-Catherine. *Timed Automata for Video Games and Interaction*, N. Lee, éditeur, *Encyclopedia of Computer Graphics and Games*, pages 1–9. Springer, 2018.

[CL-5] J. Barnat, V. Bloemen, A. Duret-Lutz, A. Laarman, L. Petrucci, J. van de Pol et E. Renault. *Parallel Model Checking Algorithms for Linear-Time Temporal Logic*, L. Sais et Y. Hamadi, éditeurs, *Handbook of Parallel Constraint Reasoning*, Chapitre 12, pages 457–507. Springer, 2018.

[CL-4] A. Gupta, V. Kahlon, S. Qadeer et T. Touili. *Model Checking Concurrent Programs*, E. M. Clarke, T. A. Henzinger, H. Veith et R. Bloem, éditeurs, *Handbook of Model Checking*, pages 573–611. Springer, 2018.

2017 [CL-3] C. Fouqueré. *Edition en ligne de fac-similés et de transcriptions*, P.-M. de Biasi et A. Herschberg Pierrot, éditeurs, *L'oeuvre comme processus*, pages 495–502. CNRS Editions, 2017.

[CL-2] A. Gupta, V. Kahlon, S. Qadeer et T. Touili. *Model Checking Concurrent Programs*, E. Clarke, T. Henzinger et H. Veith, éditeurs, *The Handbook of Model Checking*, pages 573–611. Springer, 2017.

[CL-1] S. R. Schwer. *Usage des échiquiers arithmétiques dans la résolution de problèmes combinatoires et applications actuelles*, E. Barbin, C. Goldstein, M. Moyon, S. Schwer et S. Vinatier, éditeurs, *Les travaux combinatoires en France (1870-1914) et leur actualité : un hommage à Henri Delannoy*, pages 195–220. Collection « Savoirs scientifiques et pratiques d'enseignement. PULIM, 2017.

Tutoriaux dans des conférences internationales

2018 [TU-2] E. André et G. Lipari. Tutorial on Schedulability Analysis under Uncertainty using Formal Methods. *Embedded Systems Week 2018 (ESWEEK 2018), Torino, Italy, 30 September – 5 October, 2018*, 2018. https ://www.imitator.fr/tutorials/ESWEEK18/.

2017 [TU-1] E. André, D. Lime, W. Penczek et L. Petrucci. Parametric Verification. *The 38th International Conference on Applications and Theory of Petri Nets and Concurrency (Petri Nets 2017), Zaragoza, Spain, June 25-30, 2017 - 1-day tutorial*, 2017.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2023 [**CI-120**] J. Arias, W. Jamroga, W. Penczek, L. Petrucci et T. Sidoruk. Strategic (Timed) Computation Tree Logic. *The 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, UK, 29 May-2 June, 2023*, pages 382–390, 2023.

[CI-119] C. Aubert, T. Rubiano, N. Rusch et T. Seiller. Distributing and Parallelizing Non-canonical Loops. *The 24th International Conference on Verification, Model Checking, and Abstract Interpretation (VMCAI 2023), Boston, MA, USA, January 16-17, 2023*, volume 13881 de *Lecture Notes in Computer Science*, pages 1–24, 2023.

[**CI-118**] S. Boldo, F. Clément, V. Martin, M. Mayero et H. Mouhcine. A Coq Formalization of Lebesgue Induction Principle and Tonelli's Theorem. *The 25th International Symposium on Formal Methods (FM 2023), Lübeck, Germany, March 6-10, 2023*, volume 14000 de *Lecture Notes in Computer Science*, pages 39–55, 2023.

[**CI-117**] A. Kerinec, G. Manzonetto et F. Olimpieri. Why Are Proofs Relevant in Proof-Relevant Models ? *The 50th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2023), Boston, Massachusetts, United States, January 18-20, 2023*, volume 7 de *Proceedings of the ACM on Programming Languages*, pages 8 :218– 8 :248, 2023.

2022

[CI-116] E. André, D. Marinho, L. Petrucci et J. van de Pol. Efficient Convex Zone Merging in Parametric Timed Automata. *The 20th International Conference on Formal Modeling and Analysis of Timed Systems (FORMATS 2022), Warsaw, Poland, September 13-15, 2022*, volume 13465 de *Lecture Notes in Computer Science*, pages 200–218, 2022.

[CI-115] J. Arias, M. Knapik, W. Penczek et L. Petrucci. Modular Analysis of Tree-Topology Models. *The 23rd International Conference on Formal Engineering Methods (ICFEM'22), Madrid, Spain, October 24-27, 2022*, volume 13478 de *Lecture Notes in Computer Science*, pages 36–53, 2022.

[CI-114] J. Arias, Ł. Maśko, W. Penczek, L. Petrucci et T. Sidoruk. Minimal Schedule with Minimal Number of Agents in Attack-Defence Trees. *The 26th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'22), Hiroshima City, Japan, March 26-30, 2022*, pages 1–10, 2022.

[CI-113] C. Aubert, T. Rubiano, N. Rusch et T. Seiller. mwp-Analysis Improvement and Implementation : Realizing Implicit Computational Complexity. *The 7th International Conference on Formal Structures for Computation and Deduction (FSCD 2022), Haifa, Israel, August 2-5, 2022*, volume 228 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26 :1–26 :23, 2022.

[**CI-112**] D. Barbarossa. Resource approximation for the $\lambda\mu$-calculus. *The 37th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS'22), Haifa, Israel, August 2 - 5, 2022*, pages 27 :1–12, 2022.

[CI-111] C. Coti, L. Petrucci et D. A. Torres Gonzalez. A Formal Model for Fault Tolerant Parallel Matrix Factorization. *The 26th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'22), Hiroshima City, Japan, March 26-30, 2022*, pages 62–70, 2022.

[CI-110] S. Evangelista, L. Kristensen et L. Petrucci. Distributed Explicit State Space Exploration with State Reconstruction for RDMA Networks. *26th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'22), Hiroshima City, Japan, March 26-30, 2022*, pages 107–116. IEEE, 2022.

[CI-109] I. Garfatta, K. Klai, M. Graïet et W. Gaaloul. Model checking of vulnerabilities in smart contracts : a solidity-to-CPN approach. *The 37th ACM/SIGAPP Symposium on Applied Computing (SAC '22), Virtual Event, April 25 - 29, 2022*, pages 316–325, 2022.

[CI-108] P. Jacobé de Naurois. Parallelism in Soft Linear Logic. *The 30th EACSL Annual Conference on Computer Science Logic (CSL 2022), Göttingen, Germany, February 14-19, 2022*, volume 216 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26 :1–26 :16, 2022.

[CI-107] J.-V. Loddo et R. Kanawati. Mariotel : A Web-Based Virtual Remote Computer Science Lab. *The 20th International Conference on Information Technology Based Higher Education and Training (ITHET'22), Antalya, Turkey, November 7-9, 2022*, pages 1–6, 2022.

[CI-106] H. Ochi et K. Klai. Checking Composition-Aware Service Substitutability. *The 20th IEEE International Conference on Parallel & Distributed Processing with Applications (ISPA 2022), Melbourne, Australia, December 17-19, 2022*, pages 1–8, 2022.

[**CI-105**] N. Souid, K. Klai, C. Ameur Abid et S. Ben Ahmed. At Design-Time Approach for Supervisory Control of Opacity. *The 28th International Conference on Cooperative Information Systems (CoopIS 2022), Bozen-Bolzano, Italy, October 4-7, 2022*, volume 13591 de *Lecture Notes in Computer Science*, pages 37–54, 2022.

[CI-104] N. Souid, K. Klai, C. Ameur Abid et S. Ben Ahmed. Hyper Symbolic Observation Graph to Enforce Opacity of Discrete Event Systems using Supervisory Control. *The 8th International Conference on Control, Decision and Information Technologies (CoDIT 2022), Istanbul, Turkey, May 17-20, 2022*, pages 1397–1402, 2022.

[CI-103] N. Souid, K. Klai, C. Ameur Abid et S. Ben Ahmed. Optimal Supervisory Control of Opacity for Modular Systems. *The 20th IEEE International Conference on Parallel & Distributed Processing with Applications (ISPA 2022), Melbourne, Australia, December 17-19, 2022*, pages 1–8, 2022.

[CI-102] T. Touili. Register Automata for Malware Specification. *The 17th International Conference on Availability, Reliability and Security (ARES 2022), Vienna, Austria, August 23 - 26, 2022*, pages 147 :1–147 :7, Vienna, Austria, 2022.

[CI-101] T. Touili et X. Ye. SMODIC : A Model Checker for Self-modifying Code. *The 17th International Conference on Availability, Reliability and Security (ARES 2022), Vienna, Austria, August 23 - 26, 2022*, pages 20 :1–20 :6, 2022.

2021 [**CI-100**] E. André, J. Arias, L. Petrucci et J. van de Pol. Iterative Bounded Synthesis for Efficient Cycle Detection in Parametric Timed Automata. *The 27th International Conference Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2021), Luxembourg City, Luxembourg, March 27 - April 1, 2021*, volume 12651 de *Lecture Notes in Computer Science*, pages 311–329, 2021.

[**CI-99**] J. Arias, W. Penczek, L. Petrucci et T. Sidoruk. ADT2AMAS : Managing Agents in Attack-Defence Scenarios. *The 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'21), London, UK, May 3-7, 2021*, pages 1749–1751, 2021. (Demonstration Track).

[**CI-98**] R. Ben Halima, K. Klai, M. Sellami et Z. Maamar. Formal Modeling and Verification of Property-based Resource Consumption Cycles. *The 2021 IEEE International Conference on Services Computing (SCC 2021), Chicago, IL, USA, September 5-10, 2021*, pages 370–375, 2021.

[**CI-97**] C. Coti, L. Petrucci et D. A. Torres Gonzalez. Fault-tolerant LU factorisation is low cost. *The 27th International European Conference on Parallel and Distributed Computing (Euro-Par'21), Lisbon, Portugal, September 1-3, 2021*, volume 12820 de *Lecture Notes in Computer Science*, pages 536–549, 2021.

[CI-96] K. H. T. Dam et T. Touili. MADLIRA - a tool for Android malware detection. *The 7th International Conference on Information Systems Security and Privacy (ICISSP 2021), Lisbon, Portugal, February 11-13, 2021*, pages 670–675, 2021.

[CI-95] E. Finster, S. Mimram, M. Lucas et T. Seiller. A cartesian bicategory of polynomial functors in homotopy type theory. *37th Conference on the Mathematical Foundations of Programming Semantics (MFPS 2021), Salzburg, Austria, August 30–September 2, 2021*, volume 352 de *Electronic Proceedings in Theoretical Computer Science*, pages 67–83, 2021.

[CI-94] I. Garfatta, K. Klai, W. Gaaloul et M. Graiet. A Survey on Formal Verification for Solidity Smart Contracts. *The 2021 Australasian Computer Science Week Multiconference (ACSW '21), Dunedin, New Zealand, February 1-5, 2021*, pages 3 :1–3 :10, 2021.

[CI-93] I. Garfatta, K. Klai, M. Graïet et W. Gaaloul. A Solidity-to-CPN Approach Towards Formal Verification of Smart Contracts. *The 30th IEEE International Conference on Enabling Technologies : Infrastructure for Collaborative Enterprises (WETICE 2021), Bayonne, France, October 27-29, 2021*, pages 69–74. IEEE, 2021.

[**CI-92**] I. Garfatta, K. Klai, M. Graïet et W. Gaaloul. Model Checking of Solidity Smart Contracts Adopted for Business Processes. *The 19th International Conference on Service-Oriented Computing (ICSOC 2021), Virtual Event, November 22-25, 2021*, volume 13121 de *Lecture Notes in Computer Science*, pages 116–132, 2021.

[CI-91] J. Jerray, L. Fribourg et E. André. An Approximation of Minimax Control using Random Sampling and Symbolic Computation. *The 7th IFAC Conference on Analysis and Design of Hybrid Systems (ADHS 2021), Brussels, Belgium, July 7-9, 2021*, volume 54 (5) de *IFAC-Papers On Line*, pages 265–270, 2021.

[**CI-90**] J. Jerray, L. Fribourg et E. André. Robust optimal periodic control using guaranteed Euler's method. *The 2021 American Control Conference (ACC 2021), New Orleans, LA, USA, May 25-28, 2021*, pages 986–991, 2021.

[CI-89] A. Kerinec, G. Manzonetto et S. Ronchi Della Rocca. Call-By-Value, Again ! *6th International Conference on Formal Structures for Computation and Deduction (FSCD 2021), Buenos Aires, Argentina, July 17-24, 2021*, volume 195 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7 :1–7 :18, 2021.

[**CI-88**] D. Mazza et M. Pagani. Automatic differentiation in PCF. *The 48th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2021), Copenhagen, DEnmark, January 20-22, 2021*, volume 5 de *Proceedings of the ACM on Programming Languages*, pages 28 :1–27, 2021.

[**CI-87**] C. Noûs, P. Pradic et L. T. D. Nguyen. Comparison-Free Polyregular Functions. *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021), Glasgow, Scotland, July 12-16, 2021*, volume 198 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 139 :1–20, 2021.

[**CI-86**] F. Olimpieri. Intersection Type Distributors. *36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2021), Rome, Italy, June 29 - July 2, 2021*, pages 1–15, 2021.

[**CI-85**] N. Souid et K. Klai. A Novel Approach for Supervisor Synthesis to Enforce Opacity of Discrete Event Systems. *The 23rd International Conference on Information and Communications Security (ICICS 2021), Chongqing, China, November 19-21, 2021*, volume 12919 de *Lecture Notes in Computer Science*, pages 210–227, 2021.

2020 [CI-84] C. Ameur Abid, K. Klai, J. Arias et H. Ouni. SOG-Based Multi-Core LTL Model Checking. *The 18th IEEE International Conference on Parallel & Distributed Processing with Applications (ISPA 2020), Exeter, United Kingdom, December 17-19, 2020,*, pages 9–17, 2020.

[CI-83] J. Arias, W. Penczek, L. Petrucci, T. Sidoruk et M. Stoelinga. Hackers vs. Security : Attack-Defence Trees as Asynchronous Multi-agent Systems. *The 22nd International Conference on Formal Engineering Methods (ICFEM 2020), Singapore, Singapore, March 1-3, 2021*, volume 12531 de *Lecture Notes in Computer Science*, pages 3–19, 2020.

[**CI-82**] D. Barbarossa et G. Manzonetto. Taylor subsumes Scott, Berry, Kahn and Plotkin. *The 47th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2020), New Orleans, Louisiana, United States, January 22-24, 2020*, volume 4 de *Proceedings of the ACM on Programming Languages*, pages Article 1 :1–23, 2020.

[**CI-81**] A. Brunel, D. Mazza et M. Pagani. Backpropagation in the Simply Typed Lambda-Calculus with Linear Negation. *The 47th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2020), New Orleans, Louisiana, United States, January 22-24, 2020*, volume 4 de *Proceedings of the ACM on Programming Languages*, pages 64 :1–64 :27, 2020.

[CI-80] E. Hainry, D. Mazza et R. Péchoux. Polynomial Time over the Reals with Parsimony. *The 15th International Symposium on Functional and Logic Programming (FLOPS 2020), Akita, Japan, September 14–16, 2020*, volume 12073 de *Lecture Notes in Computer Science*, pages 50–65, 2020.

[CI-79] J. Jerray, L. Fribourg et E. André. Guaranteed phase synchronization of hybrid oscillators using symbolic Euler's method (verification challenge). *The 7th International Workshop on Applied Verification of Continuous and Hybrid Systems (ARCH20), Berlin, Germany, July 12, 2020*, volume 74 de *EPiC Series in Computing*, pages 197–208, 2020.

[**CI-78**] P. Pradic et L. T. D. Nguyen. Implicit Automata in Typed λ-Calculi I : Aperiodicity in a Non-Commutative Logic. *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020), Saarbrücken, Germany, July 8-11, 2020*, volume 168 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 135 :1–20, 2020.

[**CI-77**] D. Schemmel, J. Büning, C. Rodriguez, D. Laprell et K. Wehrle. Symbolic Partial-Order Execution for Testing Multi-Threaded Programs. *The 32nd International Conference on Computer Aided Verification (CAV 2020), Los Angeles, CA, USA, July 21-24, 2020*, volume 12224 de *Lecture Notes in Computer Science*, pages 376–400, 2020.

[**CI-76**] T. Touili et X. Ye. CTL Model Checking of Self Modifying Code. *The 25th International Conference on Engineering of Complex Computer Systems (ICECCS 2020), Singapore, October 28-31, 2020*, pages 11–20, 2020.

2019

[CI-75] E. André. Formalizing Time4sys using parametric timed automata. *The 2019 International Symposium on Theoretical Aspects of Software Engineering (TASE 2019), Guilin, China, July 29-31, 2019*, pages 176–183, 2019.

[CI-74] E. André, P. Arcaini, A. Gargantini et M. Radavelli. Repairing Timed Automata Clock Guards through Abstraction and Testing. *The 13th International Conference on Tests and Proofs (TAP 2019), Porto, Portugal, October 9-11, 2019*, volume 11823 de *Lecture Notes in Computer Science*, pages 129–146, 2019.

[**CI-73**] E. André, V. Bloemen, L. Petrucci et J. van de Pol. Minimal-Time Synthesis for Parametric Timed Automata. *The 25th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2019), Prague, Czech Republic, April 6-11, 2019*, volume 11428 de *Lecture Notes in Computer Science*, pages 211–228, 2019.

[CI-72] E. André, E. Coquard, L. Fribourg, J. Jerray et D. Lesens. Parametric Schedulability Analysis of a Launcher Flight Control System Under Reactivity Constraints. *The 19th International Conference on Application of Concurrency to System Design (ACSD 2019), Aachen, Germany, June 23-28, 2019*, pages 13–22, 2019.

[CI-71] E. André, B. Delahaye, P. Fournier et D. Lime. Parametric Timed Broadcast Protocols. *The 20th International Conference on Verification, Model Checking, and Abstract Interpretation (VMCAI 2019), Cascais, Portugal, January 13-15, 2019*, volume 11388 de *Lecture Notes in Computer Science*, pages 491–512, 2019.

[**CI-70**] E. André, L. Fribourg, J.-M. Mota et R. Soulat. Verification of an industrial asynchronous leader election algorithm using abstractions and parametric model checking. *The 20th International Conference on Verification, Model Checking, and Abstract Interpretation (VMCAI 2019), Cascais, Portugal, January 13-15, 2019*, volume 11388 de *Lecture Notes in Computer Science*, pages 409–424, 2019.

[CI-69] E. André, J. Jerray et S. Mhiri. Time4sys2imi : A tool to formalize real-time system models under uncertainty. *The 16th International Colloquium on Theoretical Aspects of Computing (ICTAC 2019), Hammamet, Tunisia, October 31 - November 4, 2019*, volume 11884 de *Lecture Notes in Computer Science*, pages 113–123, 2019.

[CI-68] E. André, D. Lime et M. Ramparison. On the expressive power of invariants in parametric timed automata. *The 24th International Conference on Engineering of Complex Computer Systems (ICECCS 2019), Guangzhou, China, November 10-13, 2019*, pages 87–96, 2019.

[**CI-67**] E. André, D. Lime et M. Ramparison. Parametric updates in parametric timed automata. *39th International Conference on Formal Techniques for Distributed Objects, Components, and Systems (FORTE 2019), Kongens Lyngby, Denmark, June 17-21, 2019*, volume 11535 de *Lecture Notes in Computer Science*, pages 39–56, 2019.

[CI-66] E. André, D. Lime, M. Ramparison et M. Stoelinga. Parametric analyses of attack-fault trees. *The 19th International Conference on Application of Concurrency to System Design (ACSD 2019), Aachen, Germany, June 23-28, 2019*, pages 33–42, 2019.

[**CI-65**] E. André et J. Sun. Parametric Timed Model Checking for Guaranteeing Timed Opacity. *The 17th International Symposium on Automated Technology for Verification and Analysis (ATVA 2019), Taipei, Taiwan, October 28-31, 2019*, volume 11781 de *Lecture Notes in Computer Science*, pages 115–130, 2019.

[CI-64] A. Bourouis, K. Klai et N. Ben. Measuring Opacity for Non-Probabilistic DES : a SOG-Based Approach. *The 24th International Conference on Engineering of Complex Computer Systems (ICECCS 2019), Guangzhou, China, November 10-13, 2019*, pages 242–247, 2019.

[**CI-63**] U. Dal Lago, M. de Visme, D. Mazza et A. Yoshimizu. Intersection Types and Runtime Errors in the Pi-Calculus. *The 46th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2019), Cascais, Portugal, January 16-18, 2019*, volume 3 de *Proceedings of the ACM on Programming Languages*, pages 7 :1–7 :29, 2019.

[CI-62] K. H. T. Dam et T. Touili. STAMAD : a STAtic MAlware Detector. *The 14th International Conference on Availability, Reliability and Security (ARES 2019), Canterbury, UK, August 26-29, 2019*, pages 25 :1–25 :6, 2019.

[CI-61] P. Jacobé de Naurois. Pointers in Recursion : Exploring the Tropics. *The 4th International Conference on Formal Structures for Computation and Deduction (FSCD 2019),*

Dortmund, Germany, June 24-30, 2019, volume 131 de *Leibniz International Proceedings in Informatics (LIPIcs)*, page Article 29 (18 pages), 2019.

[CI-60] R. Jain, K. Klai et S. Tata. Formal Modeling and Verification of Scalable Process-Aware Distributed IoT Applications. *The 17th IEEE International Conference on Parallel & Distributed Processing with Applications (ISPA 2019), Xiamen, China, December 16-18, 2019*, pages 263–270, 2019.

[CI-59] M. Knapik, W. Penczek, L. Petrucci et T. Sidoruk. Squeezing State Spaces of (Attack-Defence) Trees. *The 24th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'19), Guangzhou, China, November 10-13, 2019.*, pages 71–80, 2019.

[CI-58] H. V. Nguyen et T. Touili. BCARET Model Checking for Malware Detection. *16th International Colloquium on Theoretical Aspects of Computing (ICTAC 2019), Hammamet, Tunisia, October 31 - November 4*, volume 11884 de *Lecture Notes in Computer Science*, pages 273–291, Hammamet, Tunisia, 2019.

[CI-57] H. Ouni, K. Klai, C. Ameur Abid et B. Zouari. Towards Parallel Verification of Concurrent Systems Using the Symbolic Observation Graph. *19th International Conference on Application of Concurrency to System Design (ACSD 2019), Aachen, Germany, June 23-28, 2019*, pages 23–32, 2019.

[CI-56] H. Ouni, K. Klai et B. Zouari. Parallel construction of the Symbolic Observation Graph. *17th International Conference on High Performance Computing & Simulation (HPCS 2019), Dublin, Ireland, July 15-19, 2019*, pages 1011–1013, 2019.

[**CI-55**] P. Pradic et L. T. D. Nguyen. From Normal Functors to Logarithmic Space Queries. *The 46th International Colloquium on Automata, Languages, and Programming (ICALP 2019), July 9-12, 2019, Patras, Greece*, volume 132 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 123 :1–15, 2019.

[CI-54] T. Touili et X. Ye. LTL Model Checking of Self Modifying Code. *The 24th International Conference on Engineering of Complex Computer Systems (ICECCS 2019), Guangzhou, China, November 10-13, 2019*, pages 1–10, Guangzhou, China, 2019.

[CI-53] M. Waga et E. André. Online Parametric Timed Pattern Matching with Automata-Based Skipping. *The 11th International Symposium NASA Formal Methods (NFM 2019), Houston, TX, USA, May 7-9, 2019*, volume 11460 de *Lecture Notes in Computer Science*, pages 371–389, 2019.

[**CI-52**] M. Waga, E. André et I. Hasuo. Symbolic Monitoring against Specifications Parametric in Time and Data. *The 31st International Conference on Computer Aided Verification (CAV 2019), New York City, NY, USA, July 15-18, 2019*, volume 11561 de *Lecture Notes in Computer Science*, pages 520–539, 2019.

2018 [**CI-51**] E. André, I. Hasuo et M. Waga. Offline Timed Pattern Matching under Uncertainty. *The 23rd International Conference on Engineering of Complex Computer Systems (ICECCS 2018), Melbourne, Australia, December 12-14, 2018*, pages 10–20, 2018.

[CI-50] E. André, D. Lime et M. Ramparison. TCTL model checking lower/upper-bound parametric timed automata without invariants. *The 16th International Conference on Formal Modeling and Analysis of Timed Systems (FORMATS 2018), Beijing, China, September 4-6, 2018*, volume 11022 de *Lecture Notes in Computer Science*, pages 37–52, 2018.

[CI-49] E. André, D. Lime et M. Ramparison. Timed automata with parametric updates. *The 18th International Conference on Application of Concurrency to System Design (ACSD 2018), Bratislava, Slovakia, June 25-29, 2018*, pages 21–29, 2018.

[CI-48] S. Boubaker, K. Klai, H. Kortas et W. Gaaloul. A Formal Model for Business Process Configuration Verification Supporting OR-Join Semantics. *On the Move to Meaningful Internet Systems. OTM 2018 Conferences - Confederated International Conferences : CoopIS, C&TC, and ODBASE 2018, Valletta, Malta, October 22-26, 2018*, volume 11229 de *Lecture Notes in Computer Science*, pages 623–642, 2018.

[CI-47] F. Breuvart et U. Dal Lago. On Intersection Types and Probabilistic Lambda Calculi. *The 20th International Symposium on Principles and Practice of Declarative Programming (PPDP 2018), Frankfurt am Main, Germany, September 03-05, 2018*, pages 8 :1–13. Association for Computing Machinery, 2018.

[**CI-46**] C. Coti, S. Evangelista et L. Petrucci. One-Sided Communications for more Efficient Parallel State Space Exploration over RDMA Clusters. *The 24th International European Conference on Parallel and Distributed Computing (Euro-Par'18), Torino, Italy, August 27-31, 2018*, volume 11014 de *Lecture Notes in Computer Science*, pages 432–446, 2018.

[CI-45] C. Coti, S. Evangelista et L. Petrucci. State Compression Based on One-Sided Communications for Distributed Model Checking. *The 23rd IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'18), Melbourne, Australia, December 12-14, 2018*, pages 41–50, 2018.

[CI-44] K. H. T. Dam et T. Touili. Learning Malware Using Generalized Graph Kernels. *The 13th International Conference on Availability, Reliability and Security (ARES 2018), Hamburg, Germany, August 27-30, 2018*, pages 28 :1–28 :6, 2018.

[**CI-43**] K. H. T. Dam et T. Touili. Precise Extraction of Malicious Behaviors. *The 42nd IEEE Annual Computer Software and Applications Conference (COMPSAC 2018), Tokyo, Japan, July 23-27, 2018*, pages 229–234, 2018.

[CI-42] M. Diaz et T. Touili. Model Checking Dynamic Pushdown Networks with Locks and Priorities. *The 6th International Conference on Networked Systems (NETYS 2018), Essaouira, Morocco, May 9-11, 2018, Revised Selected Papers*, volume 11028 de *Lecture Notes in Computer Science*, pages 240–251, 2018.

[CI-41] H. Férée, S. Hym, M. Mayero, J.-Y. Moyen et D. Nowak. Formal proof of polynomial-time complexity with quasi-interpretations. *The 7th ACM SIGPLAN International Conference on Certified Programs and Proofs CPP 2018), Los Angeles, CA, USA, January 8-9, 2018*, pages 146–157, 2018.

[CI-40] I. Garfatta, K. Klai, M. Graiet et W. Gaaloul. Formal Modelling and Verification of Cloud Resource Allocation in Business Processes. *On the Move to Meaningful Internet Systems. OTM 2018 Conferences - Confederated International Conferences : CoopIS, C&TC, and ODBASE 2018, Valletta, Malta, October 22-26, 2018*, volume 11229 de *Lecture Notes in Computer Science*, pages 552–567, 2018.

[CI-39] I. Garfatta, K. Klai, M. Graiet et W. Gaaloul. [WiP] Formal Modelling of IT Resource Allocation in Business Processes. *11th IEEE Conference on Service-Oriented Computing and Applications (SOCA 2018), Paris, France, November 20-22, 2018*, pages 227–232, 2018.

[CI-38] D. Khanna, S. Sharma, C. Rodriguez et R. Purandare. Dynamic Symbolic Verification of MPI Programs. *The 22nd International Symposium on Formal Methods (FM 2018), Oxford, UK, July 15-17, 2018*, volume 10951 de *Lecture Notes in Computer Science*, pages 466–484, 2018.

[**CI-37**] D. Mazza, L. Pellissier et P. Vial. Polyadic Approximations, Fibrations and Intersection Types. *The 45th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2018), Los Angeles, California, United States, January 10-12, 2018*, volume 2 de *Proceedings of the ACM on Programming Languages*, pages 6 :1–28, 2018.

[CI-36] H. G. Nguyen, L. Petrucci et J. van de Pol. Layered and Collecting NDFS with Subsumption for Parametric Timed Automata. *The 23rd IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'18), Melbourne, Australia, DEcember 12-14, 2018*, pages 1–9, 2018.

[CI-35] H.-V. Nguyen et T. Touili. Branching Temporal Logic of Calls and Returns for Pushdown Systems. *The 14th International Conference on Integrated Formal Methods (IFM 2018), Maynooth, Ireland, September 5-7, 2018*, volume 11023 de *Lecture Notes in Computer Science*, pages 326–345, 2018.

[CI-34] L. T. D. Nguyen. Unique perfect matchings and proof nets. *The 3rd International Conference on Formal Structures for Computation and Deduction (FSCD 2018), Oxford, UK, July 9–12, 2018*, volume 108 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 25 :1– :20, 2018.

[**CI-33**] T. T. H. Nguyen, C. Rodriguez, M. Sousa, C. Coti et L. Petrucci. Quasi-Optimal Partial Order Reduction. *The 30th International Conference on Computer Aided Verification (CAV 2018), Oxford, UK, July 14-17, 2018*, volume 10982 de *Lecture Notes in Computer Science*, pages 354–371, 2018.

[CI-32] H. Ouni, K. Klai, C. Ameur Abid et B. Zouari. Reducing Time and/or Memory Consumption of The SOG construction in a Parallel Context. *The 16th IEEE International Symposium on Parallel and Distributed Processing with Application (ISPA 2018), Melbourne, Australia, December 11-13, 2018*, pages 147–154, 2018.

[**CI-31**] L. Petrucci et J. van de Pol. Parameter Synthesis Algorithms for Parametric Interval Markov Chains. *38th IFIP International Conference on Formal Techniques for Distributed Objects, Components, and Systems (FORTE'18), Madrid, Spain, June 18-21, 2018*, volume 10854 de *Lecture Notes in Computer Science*, pages 121–140, 2018.

[CI-30] A. Pommellet et T. Touili. LTL Model-Checking for Communicating Concurrent Programs. *The 12th International Conference on Verification and Evaluation of Computer and Communication Systems (VECoS 2018), Grenoble, France, September 26-28, 2018*, volume 11181 de *Lecture Notes in Computer Science*, pages 150–165, 2018.

[CI-29] A. Pommellet et T. Touili. Model-Checking HyperLTL for Pushdown Systems. *The 25th International Symposium on Model Checking Software (SPIN 2018), Malaga, Spain, June 20-22, 2018*, volume 10869 de *Lecture Notes in Computer Science*, pages 133–152, 2018.

2017 [CI-28] E. André et D. Lime. Liveness in L/U-Parametric Timed Automata. *The 17th International Conference on Application of Concurrency to System Design (ACSD'17), Zaragoza, Spain*, pages 9–18, 2017.

[**CI-27**] E. André et S.-W. Lin. Learning-based compositional parameter synthesis for event-recording automata. A. Bouajjani et S. Alexandra, éditeurs, *The 37th IFIP WG 6.1 International Conference on Formal Techniques for Distributed Objects, Components, and Systems (FORTE'17), Neuchâtel, Switzerland, June 19-22, 2017*, volume 10321 de *Lecture Notes in Computer Science*, pages 17–32. Springer, 2017. (Best paper award).

[**CI-26**] E. André, H. G. Nguyen et L. Petrucci. Efficient parameter synthesis using optimized state exploration strategies. *The 22nd IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'17), Fukuoka, Japan, November 5-8, 2017*, pages 1–10, 2017.

[CI-25] E. André, H. G. Nguyen, L. Petrucci et J. Sun. Parametric model checking timed automata under non-Zenoness assumption. *The 9th NASA Formal Methods International Symposium (NFM 2017), Moffett Field, CA, USA, May 16-18, 2017*, volume 10227 de *Lecture Notes in Computer Science*, pages 35–51. Springer, 2017.

[CI-24] S. Boldo, F. Clément, F. Faissole, V. Martin et M. Mayero. A Coq formal proof of the LaxMilgram theorem. *The 6th ACM SIGPLAN Conference on Certified Programs and Proofs (CPP 2017), Paris, France*, pages 79–89, 2017.

[CI-23] S. Boubaker, K. Klai, K. Schmitz, M. Graiet et W. Gaaloul. Deadlock-Freeness Verification of Business Process Configuration Using SOG. *The 15th International Conference on Service-Oriented Computing (ICSOC 2017), Malaga, Spain, November 13-16, 2017*, volume 10601 de *Lecture Notes in Computer Science*, pages 96–112, 2017.

[CI-22] A. Bourouis, K. Klai et N. Ben. Measuring opacity in web services. *The 19th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2017), Salzburg, Austria, December 4-6, 2017*, pages 530–534, 2017.

[**CI-21**] F. Breuvart, U. Dal Lago et A. Herrou. On Higher-Order Probabilistic Subrecursion. *The 20th International Conference on Foundations of Software Science and Computation Structures (FOSSACS 2017), Uppsala, Sweden, April 22-29, 2017*, volume 10203 de *Lecture Notes in Computer Science*, pages 370–386, 2017.

[CI-20] A. Burouis, N. Ben et K. Klai. Computing Quantified Opacity for SOG-Abstracted Web Services. *The 2017 IEEE International Conference on Services Computing (SCC 2017), Honolulu, HI, USA, June 25-30, 2017*, pages 362–369, 2017.

[CI-19] C. Coti, J.-V. Loddo et E. Viennet. An Overview of the Options Available for Practical Activities in Distributed Computing. *The 16th International Conference on Information Technology Based Higher Education and Training (ITHET 17), Ohrid, Macedonia, Jumy 10-12, 2017*, pages 1–2, 2017.

[CI-18] K. H. T. Dam et T. Touili. Extracting Android Malicious Behaviors. *The 3rd International Conference on Information Systems Security and Privacy (ICISSP 2017), Porto, Portugal, February 19-21, 2017*, pages 714–723. SCITEPress, 2017.

[**CI-17**] K. H. T. Dam et T. Touili. Learning Android Malware. *The 12th International Conference on Availability, Reliability and Security (ARES 2017), Reggio Calabria, Italy, August 29 - September 1, 2017*, pages 59 :1–59 :9, 2017.

[CI-16] K. H. T. Dam et T. Touili. Malware Detection Based On Graph Classification. *The 3rd International Conference on Information Systems Security and Privacy (ICISSP 2017), Porto, Portugal, February 19-21, 2017*, pages 455–463, 2017.

[**CI-15**] M. Diaz et T. Touili. Dealing with priorities and locks for concurrent programs. *The 15th International Symposium on Automated Technology for Verification and Analysis (ATVA 2017), Pune, India, October 3-6, 2017*, volume 10482 de *Lecture Notes in Computer Science*, pages 208–224, 2017.

[CI-14] S. Guerrini et M. Solieri. Is the optimal implementation inefficient? Elementarily not. *The 2nd International Conference on Formal Structures for Computation and Deduction (FSCD 2017), Oxford, UK, September 3-9, 2017*, volume 84 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 17 :1–16, 2017.

[CI-13] B. Intrigila, G. Manzonetto et A. Polonsky. Refutation of Sallé's Longstanding Conjecture. *The 2nd International Conference on Formal Structures for Computation and Deduction (FSCD 2017), Oxford, UK, September 3–9, 2017*, volume 84 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 20 :1–20 :18, 2017.

[CI-12] J. Li, J. Sun, B. Gao et E. André. Classification based Parameter Synthesis for Parametric Timed Automata. *The 19th International Conference on Formal Engineering Methods (ICFEM 2017), Xi'an, China, November 13-17, 2017*, volume 10610 de *Lecture Notes in Computer Science*, pages 243–261, 2017.

[CI-11] J.-Y. Moyen, T. Rubiano et T. Seiller. Loop Quasi-Invariant Chunk Detection. *15th International Symposium on Automated Technology for Verification and Analysis (ATVA 2017), Pune, India, October 3–6, 2017*, volume 10482 de *Lecture Notes in Computer Science*, pages 91–108, 2017.

[CI-10] H.-V. Nguyen et T. Touili. CARET Analysis of Multithreaded Programs. *The 27th International Symposium Logic-Based Program Synthesis and Transformation (LOPSTR 2017), Namur, Belgium, October 10-12, 2017, Revised Selected Papers*, volume 10855 de *Lecture Notes in Computer Science*, pages 73–90, 2017.

[CI-9] H.-V. Nguyen et T. Touili. CARET Model Checking for Malware Detection. *The 24th ACM SIGSOFT International SPIN Symposium on Model Checking of Software (SPIN 2017), Santa Barbara, CA, USA, July 10-14, 2017*, pages 152–161, 2017.

[CI-8] H.-V. Nguyen et T. Touili. CARET Model Checking For Pushdown Systems. *The 32nd ACM SIGAPP Symposium on Applied Computing (SAC 2017), Marrakech, Morocco, April 3-7, 2017*, pages 1393–1400, 2017.

[CI-7] H. Ouni, K. Klai, C. Ameur Abid et B. Zouari. Parallel Symbolic Observation Graph. *The 15th IEEE International Symposium on Parallel and Distributed Processing with Application (ISPA 2017), Guangzhou, China, December 12-15, 2017*, pages 770–777, 2017.

[CI-6] A. Pavaux. Inductive and Functional Types in Ludics. V. Goranko et M. Dam, éditeurs, *The 26th EACSL Annual Conference on Computer Science Logic (CSL 2017), Stockholm, Sweden, August 20-24, 2017*, volume 82 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 34 :1–20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[CI-5] A. Pommellet, M. Diaz et T. Touili. Reachability Analysis of Pushdown Systems with an Upper Stack. *11th International Conference on Language and Automata Theory and Applications (LATA 2017), Umea, Sweden, March 6-9, 2017*, volume 447–459 de *Lecture Notes in Computer Science*, pages 447–459, 2017.

[CI-4] A. Pommellet et T. Touili. Static Analysis of Multithreaded Recursive Programs Communicating via Rendez-Vous. *The 15th Asian Symposium on Programming Languages and Systems (APLAS 2017), Suzhou, China, November 27-29, 2017*, volume 10695 de *Lecture Notes in Computer Science*, pages 235–254, 2017.

[**CI-3**] M. Sousa, C. Rodriguez, V. D'Silva et D. Kroening. Abstract Interpretation with Unfoldings. *The 29th International Conference on Computer Aided Verification (CAV 2017), Heidelberg, Germany, July 24-28, 2017*, volume 10427 de *Lecture Notes in Computer Science*, pages 197–216, 2017.

[CI-2] S. Tata, K. Klai et R. Jain. Formal Model and Method to Decompose Process-Aware IoT Applications. *On the Move to Meaningful Internet Systems. OTM 2017 Conferences - Confederated International Conferences : CoopIS, C&TC, and ODBASE 2017, Rhodes, Greece, October 23-27, 2017*, volume 10573 de *Lecture Notes in Computer Science*, pages 663–680, 2017.

[CI-1] T. Touili et X. Ye. Reachability Analysis of Self Modifying Code. *22nd International Conference on Engineering of Complex Computer Systems (ICECCS 2017), Fukuoka, Japan, November 5-8, 2017*, pages 120–127, 2017.

Autres articles dans des conférences ou workshops à comité de lecture

2022 [CO-37] E. André, S. Bolat, E. Lefaucheux et D. Marinho. strategFTO : Untimed control for timed opacity. *The 8th ACM SIGPLAN International Workshop on Formal Techniques for Safety-Critical Systems (FTSCS 2022), Auckland, New Zealand, 7 December, 2022*, pages 27–33, 2022.

[CO-36] J. Arias, K. Bae, C. Olarte, P. Ölveczky, L. Petrucci et F. Rømming. Rewriting Logic Semantics and Symbolic Analysis for Parametric Timed Automata. *The 8th ACM SIGPLAN International Workshop on Formal Techniques for Safety-Critical Systems (FTSCS'22), Auckland, New Zealand, December 7, 2022*, pages 3–15, 2022.

[CO-35] M. Desainte-Catherine, I. Durand, J. Haury, B. Serpette et S. Schwer. Algebraic Framework for Design and Qualitative Evaluation of Interactive Musical Systems Temporality. *The 19th Sound and Music Computing Conference (SMC 22), Saint-Étienne, France, June 5-12, 2022*, pages 247–254, 2022. https ://www.youtube.com/watch ?v=IpnBtSVEjjo&t=2216s.

[CO-34] K. Klai, C. A. Abid, J. Arias et S. Evangelista. Hybrid Parallel Model Checking of Hybrid LTL on Hybrid State Space Representation. *14th International Conference on Verification and Evaluation of Computer and Communication Systems (VECoS 2021), Beijing, China, November 22-23, 2021*, volume 13187 de *Lecture Notes in Computer Science*, pages 27–42, 2022.

[CO-33] J.-V. Loddo et R. Kanawati. Mariotel : A virtual remote computer science lab. *The 1st International Conference on Online-Labs in Education, Stuttgart, Germany, March 10-12, 2022*, pages 485–494, 2022.

[CO-32] S. Schwer. Les systèmes de repérage temporel dans le Croissant. *3èmes rencontres sur les parlers du Croissant, Boussac, France, 7-9 octobre 2022*, 2022. https ://www.canal-u.tv/135117.

[CO-31] R. D. Strand, L. Kristensen et L. Petrucci. Formal Specification and Validation of a Data-driven Software System for Fire Risk Prediction. *The 2022 International Workshop on Petri Nets and Software Engineering (PNSE'22), Bergen, Norway, June 21, 2022*, volume 3170 de *CEUR Workshop Proceedings*, pages 1 :1–20, 2022.

2021 [CO-30] B. Eng et T. Seiller. A gentle introduction to girard's transcendental syntax. *The 5th International Workshop on Trends in Linear Logic and Applications (TLLA 2021), Rome, Italy, June 27-28, 2021*, rome (virtual), italy, 2021.

[CO-29] D. Nantes, C. Olarte et D. Ventura. A subexponential view of domains in session types. *The 16th International Workshop on Logical and Semantic Frameworks with Applications (LSFA 2021), Buenos Aires, Argentina, July 23- 24 July, 2021*, volume 357 de *Electronic Proceedings in Theoretical Computer Science*, pages 93–111, 2021.

[CO-28] J. van de Pol et L. Petrucci. On Completeness of Liveness Synthesis for Parametric Timed Automata. *The 25th International Workshop on Algebraic Development Techniques (WADT'20), Dublin, Ireland, April 29, 2021*, volume 12669 de *Lecture Notes in Computer Science*, pages 3–10, 2021.

2020 [CO-27] I. Garfatta, K. Klai, M. Graïet et W. Gaaloul. Blockchain-Based Business Processes : A Solidity-to-CPN Formal Verification Approach. *The 2020 Workshops on Service-Oriented Computing (ICSOC 2020) Dubai, United Arab Emirates, December 14-17, 2020*, volume 12632 de *Lecture Notes in Computer Science*, pages 47–53, 2020.

2019 [CO-26] C. Coti, L. Petrucci et D. A. Torres Gonzalez. Fault-tolerant matrix factorisation : a formal model and proof. *The 6th International Workshop on Synthesis of Complex Parameters (SynCoP'19), Prague, Czech Republic, April 6-7, 2019*, pages 1–4, 2019.

[CO-25] C. Coti, L. Petrucci et D. A. Torres Gonzalez. Process scheduling on volatile nodes for fault-tolerant linear algebra. *The 3rd Workshop on Data Locality (COLOC'19), Göttingen, Germany, August 26, 2019*, 2019. (Poster).

[CO-24] P. Jacobe de Naurois. Pointers in Recursion : Exploring the Tropics. *The 3rd Joint Workshop on Developments in Implicit Computational complExity and Foundational & Practical Aspects of Resource Analysis (DICE & FOPARA 2019), Prague, Czech Republic, April 6-7, 2019*, volume 298 de *Electronic Proceedings in Theoretical Computer Science*, pages 31–46, 2019.

[CO-23] L. T. D. Nguyen. On the Elementary Affine Lambda-Calculus with and Without Fixed Points. *Therd Joint Workshop on Developments in Implicit Computational complExity and Foundational & Practical Aspects of Resource Analysis (DICE-FOPARA@ETAPS 2019), Prague, Czech Republic, April 6-7, 2019*, volume 298 de *Electronic Proceedings in Theoretical Computer Science*, pages 15–29, 2019.

[CO-22] L. T. D. Nguyen et T. Seiller. Coherent interaction graphs. *Joint International Workshop on Linearity & Trends in Linear Logic and Applications (Linearity-TLLA 2018), Oxford, UK, July 7-8, 2018*, volume 292 de *Electronic Proceedings in Theoretical Computer Science*, pages 104–117, 2019.

[CO-21] L. Petrucci et J. van de Pol. Pruning NDFS for Parametric Timed Automata. *The 6th International Workshop on Synthesis of Complex Parameters (SynCoP'19), Prague, Czech Republic, April 6-7, 2019*, pages 1–9, 2019.

[CO-20] L. Petrucci et J. van de Pol. Taming NDFS for Parametric Timed Automata. *The 8th IFIP WG 1.8 Workshop on Trends in Concurrency Theory (TRENDS'19), Amsterdam, The Netherlands, August 31, 2019*, 2019.

2018

[CO-19] E. André. A Benchmark Library for Parametric Timed Model Checking. *The 6th International Workshop on Formal Techniques for Safety-Critical Systems (FTSCS 2018), Gold Coast, Australia, November 16, 2018, Revised Selected Papers*, volume 1008 de *Communications in Computer and Information Science*, pages 75–83, 2018.

[CO-18] C. Fouqueré et M. Quatrini. Ludics for modelling the role of a judge during legal debates. *Workshop on Logic, Law and Legal Reasoning (LOGLEG 2018) at UNILOG'2018, Vichy, France, June 21 - 26, 2018*, page 155, 2018.

[CO-17] G. Guerrieri et G. Manzonetto. The Bang Calculus and the Two Girard's Translations. *2018 Joint International Workshop on Linearity & Trends in Linear Logic and Applications (Linearity & TLLA 2018), Oxford, UK, 7-8 July 2018*, volume 292 de *Electronic Proceedings in Theoretical Computer Science*, pages 15–30, 2018.

[CO-16] L. T. D. Nguyen. On some tractable constraints on paths in graphs and in proofs. *The 16th Cologne-Twente Workshop (CTW18) on Graphs and Combinatorial Optimization, Paris, France, June 18-20, 2018*, 2018.

[CO-15] L. T. D. Nguyen. On the complexity of finding cycles in proof nets. *Developments in Implicit Computational Complexity 2018 (DICE 2018), Thessaloniki, Greece, April 14-15, 2018*, pages 1–5, 2018.

[CO-14] A. Søgaard, C. Braud et A. Kerinec. When does deep multi-task learning work for loosely related document classification tasks ? *The 2018 EMNLP Workshop BlackboxNLP : Analyzing and Interpreting Neural Networks for NLP, Brussels, Belgium, November 1, 2018*, pages 1 – 8, Brussels, Belgium, 2018.

2017

[CO-13] E. André. A Unified Formalism for Monoprocessor Schedulability Analysis under Uncertainty. *Joint 22nd International Workshop on Formal Methods for Industrial Critical Systems and 17th International Workshop on Automated Verification of Critical Systems (FMICS-AVoCS 2017), Turin, Italy, September 18-20, 2017*, volume 10471 de *Lecture Notes in Computer Science*, pages 100–115, 2017.

[CO-12] E. André, W. Jamroga, M. Knapik, W. Penczek et L. Petrucci. Timed ATL : Forget Memory, Just Count. *The 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), São Paulo, Brazil, May 8- 12, 2017*, pages 1460–1462, 2017. (Poster).

[CO-11] A. Boussif, M. Ghazel et K. Klai. DPN-SOG : A Software Tool for Fault Diagnosis of Labeled Petri Nets Using the Semi-Symbolic Diagnoser. *11ème Colloque sur la Modélisation des Systèmes Réactifs (MSR 2017), Marseille, France, 15-17 Novembre 2017*, 2017.

[CO-10] J. Chouquet, G. Guerrieri, L. Pellissier et L. Vaux. Normalization by Evaluation in Linear Logic. *Trends in Linear Logic and Applications 2017 (TLLA 17), Oxford, UK, September 3, 2017*, pages 1–2, 2017.

[CO-9] M. Diaz et T. Touili. Reachability Analysis of Dynamic Pushdown Networks with Priorities. *The 5th International Conference on Networked Systems (NETYS 2017), Marrakech, Morocco, May 17-19, 2017*, volume 10299 de *Lecture Notes in Computer Science*, pages 288–303, 2017.

[CO-8] C. Fouqueré et M. Quatrini. Refinement of universal quantification in Proof Theory. *Quantifiers and Determiners (QUAD), ESSLLI 2017 Workshop, Toulouse, France, July 17 - 21, 2017*, pages 1–15, 2017.

[CO-7] C. Fouqueré et M. Quatrini. Which Answers are Expected ? *(In)Coherence of discourse 4, Nancy, France, March 30-31, 2017*, pages 1–15, 2017.

[CO-6] M. Graiet et K. Klai. Track Report for Formal Verification of Service Based Systems : FVSBS 2017. *The 26th IEEE International Conference on Enabling Technologies : Infrastructure for Collaborative Enterprises (WETICE 2017), Poznan, Poland, June 21-23, 2017*, pages 238–238, 2017.

[CO-5] G. Guerrieri et L. Pellissier. Coherence, Taylor expansion, and box-connected proof-structures. *Trends in Linear Logic and Applications 2017 (TLLA 17), Oxford, UK, September 3, 2017*, pages 1–2, 2017.

[CO-4] S. Guerrini. Linear β-reduction. I. Cervesato et M. Fernandez, éditeurs, *Th 4th International Workshop on Linearity (LINEARITY 2016), Porto, Portugal, June 26, 2016*, volume 238 de *Electronic Proceedings in Theoretical Computer Science*, pages 44–53, 2017.

[CO-3] V. Mogbil et S. Del Vecchio. An Interpretation of CCS into Ludics. *The 12th Workshop on Logical and Semantic Frameworks, with Applications (LSFA 2017), Brasília, Brazil, September 23-24, 2017*, volume 338 de *Electronic Notes in Theoretical Computer Science*, pages 97–114. Elsevier, 2017.

[CO-2] L. T. D. Nguyen. Handsome proof nets for MLL+Mix with forbidden transitions. *Trends in Linear Logic and Applications 2017 (TLLA 17), Oxford, UK, September 3, 2017*, pages 1–2, 2017.

[CO-1] H. Ouni, K. Klai, C. Ameur Abid et B. Zouari. A Parallel Construction of the Symbolic Observation Graph : the Basis for Efficient Model Checking of Concurrent Systems. *The 8th International Symposium on Symbolic Computation in Software Science (SCSS 2017), Gammarth, Tunisia, April 6-9, 2017*, volume 45 de *EPiC Series in Computing*, pages 107–119, 2017.

Direction d'ouvrages (DO)

Livres

2022 [**LI-3**] H. Barendregt et G. Manzonetto. *A Lambda Calculus Satellite*, volume 94 de *Studies in Logic*. College Publications, 2022.

2018 [LI-2] C. Fouqueré, A. Lecomte, P. Livet, M. Quatrini et S. Tronçon. *Mathématique du dialogue : sens et interaction*. Hermann, 2018. 267 pages.

2017 [LI-1] E. Barbin, C. Goldstein, M. Moyon, S. Schwer et S. Vinatier. *Les travaux combinatoires en France (1870 - 1914) et leur actualité. Un hommage à Henri Delannoy*. Pulim, 2017.

Édition d'ouvrages collectifs

2022 [ED-5] L. Bernardinello et L. Petrucci. *Proceedings of the 43rd International Conference on Petri Nets and other Models of Concurrency (PetriNets'22), Bergen, Norway*, volume 13288 de *Lecture Notes in Computer Science*. Springer, 2022.

2019 [ED-4] E. André et M. Stoelinga. *The 17th International Conference Formal Modeling and Analysis of Timed Systems (FORMATS 2019), Amsterdam, The Netherlands, August 27-29, 2019*, volume 11750 de *Lecture Notes in Computer Science*. 2019.

[ED-3] T. Seiller et S. Jost. *The 3rd Joint Workshop on Developments in Implicit Computational complExity and Foundational & Practical Aspects of Resource Analysis (DICE & FOPARA 2019), Prague, Czech Republic, April 6-7,2019*, volume 298 de *Electronic Proceedings in Theoretical Computer Science*. 2019.

2017 [ED-2] S. Guerrini, H. Barendrengt et A. Piperno. Computing with lambda-terms : A special issue dedicated to Corrado Böhm for his 90th birthday. *Mathematical Structures in Computer Science*, 27(5), 2017.

[ED-1] L. Petrucci, C. Seceleanu et A. Cavalcanti. *The Joint 22nd International Workshop on Formal Methods for Industrial Critical Systems and 17th International Workshop on Automated Verification of Critical Systems, FMICS-AVoCS 2017, Turin, Italy, September 18-20, 2017*, volume 10471 de *Lecture Notes in Computer Science*. 2017.

Logiciels et Ressources (LO)

2023 [LO-20] J. Arias, K. Bae, L. Petrucci, C. Olarte, P. Ölveczky et F. Rømming. PITPN2Maude : Rewriting Logic Semantics for Parametric Time Petri Nets with Inhibitor Arcs. https://depot.lipn.univ-paris13.fr/arias/pitpn2maude, 2023.

[LO-19] J. Arias, M. T. Bennani, J. Desel, K. Klai et H. Ochi. sogMBT : Symbolic Observation Graph-Based Generation of Test Paths. https://depot.lipn.univ-paris13.fr/PMC-SOG/sogMBT, 2023.

2022 [LO-18] J. Arias, K. Bae, L. Petrucci, C. Olarte, P. Ölveczky et F. Rømming. PTA2Maude : Rewriting Logic Semantics for Parametric Timed Automata. https://depot.lipn.univ-paris13.fr/arias/pta2maude, 2022.

[LO-17] J. Arias, Ł. Maśko, C. Olarte, W. Penczek, L. Petrucci et T. Sidoruk. ADT2Maude : A rewriting logic specification for modeling ADTree and solve the optimal scheduling problem. https://depot.lipn.univ-paris13.fr/parties/tools/adt2maude, 2022.

[LO-16] J. Arias, L. Petrucci et V.-F. Le. CosyDraw : An Open Source Extensible Verification Environment. https://cosyverif.lipn.univ-paris13.fr/, 2022.

[LO-15] C. Aubert, T. Rubiano, N. Rusch et T. Seiller. pymwp : MWP analysis in Python. https://github.com/statycc/pymwp, 2022.

[LO-14] S. Boldo, F. Clément, M. Vincent, M. Mayero, F. Faissole, H. Mouhcine, L. Leclerc et S. Aubry. coq-num-analysis opam package (Numerical Analysis in Coq). https://depot.lipn.univ-paris13.fr/mayero/coq-num-analysis/, 2022.

[LO-13] C. Olarte, E. Pimentel et C. Rocha. L-Framework. https://carlosolarte.github.io/L-framework/, 2022.

2021 [LO-12] E. André. A Benchmarks Library for Extended Timed Automata. https://www.imitator.fr/library2.html, 2021.

[LO-11] E. André, J. Arias et L. Petrucci. IMITATOR : Tool for Parametric Verification and Robustness Analysis of Real-Time Systems with Parameter. https://www.imitator.fr/, 2021.

[LO-10] J. Arias, I. Garfatta et K. Klai. Solidity2CPN : Formal Verification of Solidity Smart Contracts using Coloured Petri Nets. https://github.com/garfatta/solidity2cpn, 2021.

[LO-9] J. Arias et L. Petrucci. Web user interface for IMITATOR. https://imitator.lipn.univ-paris13.fr/, 2021.

2020 [LO-8] C. Ameur Abid, J. Arias et K. Klai. PMC-SOG : Parallel Model Checking using the Symbolic Observation Graph. https://lipn.univ-paris13.fr/pmc-sog/, 2020.

[LO-7] J. Arias, W. Penczek, L. Petrucci et T. Sidoruk. ADT2AMAS : Managing Agents in Attack-Defence Scenarios. https://lipn.univ-paris13.fr/adt2amas/, 2020.

[LO-6] J.-V. Loddo. Mariotel, gestionnaire de salles de TP virtuelles sous GNU/Linux. https://tel.marionnet.org/info.php, 2020.

2019 [LO-5] T. Touili et X. Ye. SMODIC : A Model Checker for Self-modifying Code. https://lipn.univ-paris13.fr/~touili/smodic/, 2019.

2018 [LO-4] K. H. T. Dam et T. Touili. MADLIRA : a Malware Detector for Android, 2018.

[LO-3] K. H. T. Dam et T. Touili. STAMAD : a Malware Detector for PC Malware, 2018.

2017 [LO-2] E. André. IMITATOR Web page. https://www.imitator.fr/, 2017.

[LO-1] T. T. H. Nguyen, C. Rodríguez et C. Coti. Tool DPU. https://github.com/cesaro/dpu, 2017.

Thèses et Habilitations (TH)

Thèses

2022 [TH-14] I. Garfatta. Toward a Correct Blockchain-based Business Processes. Thèse, Université Sorbonne Paris Nord, 2022.

2021 [TH-13] D. Barbarossa. Une théorie de l'approximation des programmes basée sur la notion de ressources. Thèse, Université Sorbonne Paris Nord, 2021.

[TH-12] J. Jerray. Analyses formelles de systèmes temps-réel. Thèse, Université Sorbonne Paris Nord, 2021.

[TH-11] L. T. D. Nguyen. Automates implicites en logique linéaire et théorie catégorique des transducteurs. Thèse, Université Sorbonne Paris Nord, 2021.

[TH-10] D. A. Torres Gonzalez. Application-based fault tolerance for numerical linear algebra at large scale. Thèse, Université Sorbonne Paris Nord, 2021.

2019 [TH-9] A. Kaszczyc. Méthodes statiques pour la programmation fonctionnelle de simulateurs d'économies. Thèse, Université Sorbonne Paris Nord, 2019.

[TH-8] H. Ouni. Parallel verification of concurrent systems using the Symbolic Observation Graph. Thèse, Université Sorbonne Paris Nord, 2019.

[TH-7] M. Ramparison. On the theory and practice of updatable parametric timed automata. Thèse, Université Sorbonne Paris Nord, 2019.

2018 [TH-6] S. Del Vecchio. Process Algebras inside Ludics : an interpretation of the Calculus of Communicating Systems. Thèse, Université Sorbonne Paris Cité, 2018.

[TH-5] H. G. Nguyen. Efficient Parametric Verification of Parametric Timed Automata. Thèse, Université Sorbonne Paris Cité, 2018.

[TH-4] T. T. H. Nguyen. Quasi optimal model checking for concurrent systems. Thèse, Université Sorbonne Paris Cité, 2018.

2017 [TH-3] A. Pavaux. Inductive, Functional and Non-Linear Types in Ludics /Types inductifs, fonctionnels et non-linéaires en ludique. Thèse, Université Sorbonne Paris Cité, 2017.

[TH-2] L. Pellissier. Réductions et Approximations linéaires. Thèse, Université Sorbonne Paris Cité, 2017.

[TH-1] T. Rubiano. Implémentations de complexités implicites dans les compilateurs. Thèse, Université Sorbonne Paris Cité, 2017.

Habilitations

2018 [HR-3] E. André. Contributions to parametric timed model checking : Theory and algorithms. HdR, Université Sorbonne Paris Cité, 2018.

2017 [HR-2] G. Manzonetto. Lambda Calculus, Linear Logic and Symbolic Computation. HdR, Université Sorbonne Paris Cité, 2017.

[HR-1] D. Mazza. Polyadic Approximations in Logic and Computation. HdR, Université Paris 13, 2017.

Publications antérieures à l'arrivée au LIPN

Les références comportent « ↑ » et correspondent à des publications antérieures à l'arrivée au LIPN, non référencées par ailleurs.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2023 [↑RI-17] J. Arcile et E. André. Timed Automata as a Formalism for Expressing Security : A Survey on Theory and Practice. *ACM Computing Surveys*, 55(6) :127 :1–127 :36, 2023.

2022 [↑RI-16] E. André, D. Lime, D. Marinho et J. Sun. Guaranteeing Timed Opacity using Parametric Timed Model Checking. *ACM Transactions on Software Engineering and Methodology*, 31(4) :Article 64 : 1–36, 2022.

[↑RI-15] E. André, D. Lime et O. H. Roux. Reachability and liveness in parametric timed automata. *Logical Methods in Computer Science*, 18(1) :31 :1–31 :41, 2022.

[↑RI-14] M. Waga, E. André et I. Hasuo. Model-Bounded Monitoring of Hybrid Systems. *ACM Transactions on Cyber-Physical Systems*, 6(4) :30 :1–30 :26, 2022.

[↑RI-13] M. Waga, E. André et I. Hasuo. Parametric Timed Pattern Matching. *ACM Transactions on Software Engineering and Methodology*, 32(1) :10 :1–10 :35, 2022.

2020 [↑RI-12] E. André, D. Lime et N. Markey. Language Preservation Problems in Parametric Timed Automata. *Logical Methods in Computer Science*, 16(1) :Article 5 (31 pages), 2020.

[↑RI-11] E. André, T. H. Tan, M. Chen, S. Liu, J. Sun, Y. Liu et J. S. Dong. Automated Synthesis of Local Time Requirement for Service Composition. *International Journal on Software and Systems Modeling*, 19(4) :983–1013, 2020.

[↑RI-10] L. Brodo et C. Olarte. Verification Techniques for a Network Algebra. *Fundam. Informaticae*, 172(1) :1–38, 2020.

[↑RI-9] M. Falaschi, M. Gabbrielli, C. Olarte et C. Palamidessi. Dynamic Slicing for Concurrent Constraint Languages. *Fundam. Informaticae*, 177(3-4) :331–357, 2020.

2019 [↑RI-8] A. Ahmed Nacer, C. Godart, G. Rosinosky, A. Tari et S. Youcef. Business process outsourcing to the cloud : Balancing costs with security risks. *Computers in Industry*, 104 :59–74, 2019.

[↑RI-7] K. Chaudhuri, J. Despeyroux, C. Olarte et E. Pimentel. Hybrid linear logic, revisited. *Mathematical Structures in Computer Science*, 29(8) :1151–1176, 2019.

[↑RI-6] M. Kerjean et Y. Dabrowski. Models of Linear Logic based on the Schwartz epsilon-product. *Theory and Applications of Cateogories*, 34(45) :1440–1525, 2019.

2018 [↑RI-5] A. r. Ahmed Nace, E. Goettelmann, S. Youcef, A. Tari et C. Godart. A Design-Time Semi-Automatic Approach for Obfuscating a Business Process Model in a Trusted Multi-Cloud Deployment : A Design-Time Approach for BP Obfuscation. *International Journal of Web Services Research*, 15(4) :61–81, 2018.

[↑RI-4] M. Kerjean et C. Tasson. Mackey-complete spaces and power series - a topological model of differential linear logic. *Mathematical Structures in Computer Science*, 28(4) :472–507, 2018.

[↑RI-3] C. Olarte, E. Pimentel et C. Rueda. A concurrent constraint programming interpretation of access permissions. *Theory and Practice of Logic Programming*, 18(2) :252–295, 2018.

2017 [↑RI-2] C. Eberhart, T. Hirschowitz et T. Seiller. An intensionally fully-abstract sheaf model for pi (expanded version). *Logical Methods in Computer Science*, 13(4) :Article 9 (76 pages), 2017.

[↑RI-1] T. Seiller. Interaction graphs : graphings. *Annals of Pure and Applied Logic*, 168(2) :278–320, 2017.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2022 [↑CI-21] E. André, M. Waga, N. Urabe et I. Hasuo. Exemplifying parametric timed specifications over signals with bounded behavior. *The 14th NASA Formal Methods Symposium (NFM 2022), Pasaneda, CA, USA, May 24-27, 2022*, volume 13260 de *Lecture Notes in Computer Science*, pages 470–488, 2022.

[↑CI-20] J. Arcile et E. André. Zone extrapolations in parametric timed automata. *The 14th NASA Formal Methods Symposium (NFM 2022), Pasadena, CA, USA, May 24-27, 2022*, volume 13260 de *Lecture Notes in Computer Science*, pages 451–469. Springer, 2022.

[↑CI-19] B. Ghosh et E. André. Monitoring of scattered uncertain logs using uncertain linear dynamical systems. *The 42nd International Conference on Formal Techniques for Distributed Objects, Components, and Systems (FORTE 2022), Lucca, Italy, June 13-17, 2022*, volume 13273 de *Lecture Notes in Computer Science*, pages 67–87, 2022. arXiv document.

2021 [↑CI-18] E. André. IMITATOR 3 : Synthesis of timing parameters beyond decidability. *The 33rd International Conference on Computer Aided Verification (CAV 2021), Virtual Event, July 20-23, 2021*, volume 12759 de *Lecture Notes in Computer Science*, pages 552–565, 2021.

[↑CI-17] E. André, D. Marinho et J. van de Pol. A Benchmarks Library for Extended Timed Automata. *The 15th International Conference on Tests and Proofs (TAP 2021), Orléans, France, June 21–25, 2021*, volume 12740 de *Lecture Notes in Computer Science*, pages 39–50, 2021.

[↑CI-16] E. Pimentel, C. Olarte et V. Nigam. Process-As-Formula Interpretation : A Substructural Multimodal View (Invited Talk). N. Kobayashi, éditeur, *The 6th International Conference on Formal Structures for Computation and Deduction (FSCD 2021), Buenos Aires, Argentina, July 17-24, 2021,*, volume 195 de *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 3 :1–21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[↑CI-15] G. Rosinosky, S. Youcef, F. Charoy et E. Rivière. A Methodology for Tenant Migration in Legacy Shared-Table Multi-tenant Applications. *The 21st IFIP WG 6.1 International Conference on Distributed Applications and Interoperable Systems (DAIS 2021) held as Part of the 16th International Federated Conference on Distributed Computing Techniques, (DisCoTec 2021), Valletta, Malta, June 14-18, 2021*, volume 12718 de *Lecture Notes in Computer Science*, pages 3–20, 2021.

[↑CI-14] M. Waga, E. André et I. Hasuo. Model-bounded monitoring of hybrid systems. *The ACM/IEEE 12th International Conference on Cyber-Physical Systems (ICCPS '21), Nashville, Tennessee, USA, May 19-21, 2021*, pages 21–32, 2021.

2020 [↑CI-13] R. Affeldt, C. Cohen, M. Kerjean, A. Mahboubi, D. Rouhling et K. Sakaguchi. Competing Inheritance Paths in Dependent Type Theory : A Case Study in Functional Analysis. *The 10th International Joint Conference on Automated Reasoning (IJCAR 2020), Paris, France, July 1-4, 2020*, volume 12167 de *Lecture Notes in Computer Science*, pages 3–20, 2020.

[↑CI-12] E. André et A. Kryukov. Parametric non-interference in timed automata. *The 25th International Conference on Engineering of Complex Computer Systems (ICECCS 2020), Singapore, October 28-31, 2020*, pages 37–42, 2020.

[↑CI-11] S. Queiroz et C. Olarte. A semantic framework for PEGs. *The 13th ACM SIGPLAN International Conference on Software Language Engineering (SLE 2020), Virtual Event, USA, November 16-17, 2020*, pages 230–245, 2020.

2019 [↑CI-10] M. Kerjean et J.-S. Pacaud Lemay. Higher-Order Distributions for Differential Linear Logic. *The 22nd International Conference on Foundations of Software Science and Computation Structures (FOSSACS 2019), Prague, Czech Republic, April 6-11, 2019*, volume 11425 de *Lecture Notes in Computer Science*, pages 330–347, 2019.

[↑CI-9] T. Lang, C. Olarte, E. Pimentel et C. G. Fermüller. A Game Model for Proofs with Costs. S. Cerrito et A. Popescu, éditeurs, *The 28th International Conference on Automated Reasoning with Analytic Tableaux and Related Methods(TABLEAUX 2019), London, UK, September 3-5, 2019*, volume 11714 de *Lecture Notes in Computer Science*, pages 241–258, 2019.

2018 [↑CI-8] M. Falaschi et C. Olarte. An Assertion Language for Slicing Constraint Logic Languages. *The 28th International Symposium on Logic-Based Program Synthesis and Transformation (LOPSTR 2018), Frankfurt/Main, Germany, September 4-6, 2018, Revised Selected Papers*, volume 11408 de *Lecture Notes in Computer Science*, pages 148–165, 2018.

[↑CI-7] M. Kerjean. A Logical Account for Linear Partial Differential Equations. *The 33rd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2018), Oxford, UK, July 9-12, 2018*, pages 589–598, 2018.

[↑CI-6] G. Rosinosky, C. Labba, V. Ferme, S. Youcef, F. Charoy et C. Pautasso. Evaluating Multi-tenant Live Migrations Effects on Performance. *On the Move to Meaningful Internet Systems. OTM 2018 Conferences - Confederated International Conferences : CoopIS, C&TC,*

and ODBASE 2018, Valletta, Malta, October 22-26, 2018, volume 11229 de *Lecture Notes in Computer Science*, pages 61–77, 2018.

[↑CI-5] G. Rosinosky, S. Youcef et F. Charoy. A Genetic Algorithm for Cost-Aware Business Processes Execution in the Cloud. *The 16th International Conference on Service-Oriented Computing (ICSOC 2018), Hangzhou, China, November 12-15, 2018*, volume 11236 de *Lecture Notes in Computer Science*, pages 198–212, 2018.

2017

[↑CI-4] A. Ahmed Nacer, C. Godart, S. Youcef et A. Tari. A Metric for Evaluating the Privacy Level of a Business Process Logic in a Multi-cloud Deployment. *The 21st IEEE International Enterprise Distributed Object Computing Conference (EDOC 2017), Quebec City, QC, Canada, October 10-13, 2017*, pages 153–158, 2017.

[↑CI-3] J. Arias, P. Ciuciu, M. Dojat, F. Forbes, A. Frau-Pascual, T. Perret et J. M. Warnking. PyHRF : A Python Library for the Analysis of fMRI Data Based on Local Estimation of Hemodynamic Response Function. *The 16th Python in Science Conference (SciPy 2017), Austin, Texas, July 10 - July 16, 2017*, pages 34–40, 2017.

[↑CI-2] M. Cano, J. Arias et J. A. Pérez. Session-based Concurrency, Reactively. *The 37th IFIP WG 6.1 International Conference on Formal Techniques for Distributed Objects, Components, and Systems (FORTE 2017), Neuchâtel, Switzerland, June 19-22, 2017*, volume 10321 de *Lecture Notes in Computer Science*, pages 74–91, 2017.

[↑CI-1] G. Rosinosky, S. Youcef et F. Charoy. Efficient Migration-Aware Algorithms for Elastic BPMaaS. *The 15th International Conference on Business Process Management (BPM 2017), Barcelona, Spain, September 10-15, 2017, Proceedings*, volume 10445 de *Lecture Notes in Computer Science*, pages 147–163, 2017.

Autres articles dans des conférences ou workshops à comité de lecture

2020

[↑CO-6] L. Brodo et C. Olarte. A Constraint-based Language for Multiparty Interactions. C. Nalon et G. Reis, éditeurs, *The 15th International Workshop on Logical and Semantic Frameworks with Applications (LSFA 2020), Online, September 15, 2020*, volume 351 de *Electronic Notes in Theoretical Computer Science*, pages 25–50, 2020.

[↑CO-5] C. Olarte, E. Pimentel et B. Xavier. A Fresh View of Linear Logic as a Logical Framework. *The 15th International Workshop on Logical and Semantic Frameworks with Applications (LSFA 2020), Online, September 15, 2020*, volume 351 de *Electronic Notes in Theoretical Computer Science*, pages 143–165, 2020.

2018

[↑CO-4] M. Cano, J. Arias et J. A. Pérez. Refined Encodings of Sessions in ReactiveML. *The 11th International Workshop on Interaction and Concurrency Experience (ICE 2018), Madrid, Spain, June 20-21, 2018*, 2018.

[↑CO-3] J. Despeyroux, A. P., P. Liò et C. Olarte. A Logical Framework for Modelling Breast Cancer Progression. *The First International Symposium on Molecular Logic and Computational Synthetic Biology (MLCSB 2018), Santiago, Chile, December 17-18, 2018, Revised Selected Papers*, volume 11415 de *Lecture Notes in Computer Science*, pages 121–141, 2018.

[↑CO-2] C. Olarte, V. de, E. Pimentel et G. Reis. The ILLTP Library for Intuitionistic Linear Logic. T. Ehrhard, M. Fernández, V. de Paiva et L. T. de Falco, éditeurs, *The Joint International Workshop on Linearity & Trends in Linear Logic and Applications (Linearity-TLLA@FLoC 2018), Oxford, UK, 7-8 July 2018*, volume 292 de *Electronic Proceedings in Theoretical Computer Science*, pages 118–132, 2018.

[↑CO-1] C. Olarte, E. Pimentel et C. Rocha. Proving Structural Properties of Sequent Systems in Rewriting Logic. *The 12th International Workshop on Rewriting Logic and Its Applications (WRLA 2018), a Satellite Event of ETAPS, Thessaloniki, Greece, June 14-15, 2018*, volume 11152 de *Lecture Notes in Computer Science*, pages 115–135, 2018.

Thèses et Habilitations (TH)

Thèses

[↑TH-1] M. Kerjean. Reflexive spaces of smooth functions : a logical account of linear partial differential equations. Thèse, Université Sorbonne Paris Cité, 2018.

3.3 Domaine 4. Inscription des activités de recherche dans la société

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique.

L'équipe travaille à des sujets de nature fondamentale (pour l'axe logique) ou de nature potentiellement appliquée mais avec un regard fondamental (pour l'axe vérification). Dans les deux cas, les applications industrielles ne sont pas la préoccupation principale de l'équipe. L'axe vérification a néanmoins effectué des collaborations dans un contexte industriel, notamment avec Thales et dans l'industrie aérospatiale et de l'automobile :

- nous avons proposé [CI-75, CI-69] une traduction d'un formalisme (Time4sys) proposé par Thales, afin de permettre une vérification automatisée de systèmes temps-réel. Nous avons vérifié [**CI-70**] un algorithme distribué mis au point par Thales.
- Nous avons proposé [RI-33] une vérification d'un ordonnanceur d'Ariane avec le logiciel IMITATOR [↑CI-18]. Nous avons également proposé des algorithmes de monitoring efficaces avec application à des exemples issus de l'industrie automobile [**CI-51**, CI-53, **CI-52**, ↑RI-13] — travaux publiés dans plusieurs congrès de rang A ou A*, avec prix du meilleur article.

3.3.2 Référence 2. L'équipe développe des produits à destination du monde culturel, économique et social.

L'équipe a contribué aux actions suivantes :

- édition de la norme ISO-IEC 15909-3, publiée en novembre 2021 ;
- projet de maturation SATT ERGANEO pour création de startup dans le domaine de la détection de malware.

3.3.3 Référence 3. L'équipe partage ses connaissances avec le grand public et intervient dans des débats de société.

L'équipe est investie dans la diffusion et valorisation de la recherche vers le monde enseignant en interaction avec l'IREM Paris Nord, dont la direction est assurée par une membre de l'équipe. Parmi les actions entreprises dans ce cadre, nous mentionnons :

- le montage et pilotage du Diplôme Inter-Universitaire « EIL » (Enseigner l'Informatique au Lycée) pour 24 enseignants de l'académie de Créteil ;
- l'animation du groupe informatique avec participation de plusieurs chercheurs du LIPN ;
- l'organisation du stage « Coq en math ou l'art de raisonner », pour lycéennes et lycéens de 2nde su Lycée Joséphine Baker de Pierrefite (2022).

L'équipe a également conduit de nombreuses actions de partage de la connaissance avec le grand public et le jeune public :

- interview sur la détection de malware pour France Culture (2018) ;
- participation avec conférences et interviews à la « Fête de la science »/« Savante banlieue » (jusqu'à 2020) ;
- participation à la création de plusieurs laboratoires de mathématiques du plan Villani-Torossian ;
- présentations concernant le métier de chercheur/chercheuse et/ou les thématiques de l'équipe à des élèves de collège et lycée (toute la période) ;

— participation aux journées « Filles et Maths : une équation lumineuse », organisées par l'Institut Galilée et s'adressant à des lycéennes de Terminale, 1ère scientifique et de 2nde (2017–2020) ;
— animation d'ateliers Math.en.Jean (2017–2019).

L'équipe est aussi impliquée dans des actions en faveur de la promotion des logiciels libres et de l'ouverture des données de la recherche de l'échelle locale au niveau national et international :

— la participation (1 membre) au Conseil scientifique DoRANum disciplinaire, qui vise à accompagner les universitaires français(es) dans l'ouverture de leurs données de recherche (depuis 2021) ;
— la participation (1 membre) du groupe « EOSC (European Open Science Cloud) Future User Group », qui vise à préparer l'ouverture de la plateforme EOSC ayant, parmi d'autres, l'objectif d'ouvrir les données de la recherche des universitaires européen(ne)s (depuis 2021) ;
— la participation (1 ambassadeur) au projet Software Heritage, ayant pour ambition de collecter, préserver et partager tous les logiciels qui sont publiquement disponibles sous forme de code source.

Chapitre 4
Trajectoire de l'équipe LoVe

Sommaire

4.1 SWOT de l'équipe 275
4.2 Trajectoire 276
4.2.1 Logique, théorie de la programmation et complexité 277
4.2.2 Vérification 281

4.1 SWOT de l'équipe

- **Forces et Atouts**
  - **Production scientifique :** la production de l'équipe est de haute qualité, avec publication systématique (environ 2 papiers/an) dans des conférences de rang maximal (CORE A*, ERA A). Pour l'axe vérification, cette production est complémentée par le développement et la diffusion de logiciels.
  - **Dynamisme :** l'équipe développe des thèmes d'actualité, est porteuse d'initiatives scientifiques au niveau national et international, accueille régulièrement des chercheurs et chercheuses invité(e)s et fait partie de réseaux de recherche internationaux (Italie, Danemark).

  - **Personnel CNRS :** la présence de chercheurs CNRS, à la fois CR et DR, et d'un IR, est un atout important pour le développement de la recherche dans l'équipe.
  - **Attractivité :** l'équipe sait attirer des jeunes chercheurs (CR, MCF, post-doc, doctorant(e)s), et bénéficie aussi de la présence d'un bon vivier sur ses thématiques, en France comme en Europe.

- **Faiblesses**
  - **Manque de cohésion :** les deux axes de l'équipe ne collaborent presque pas entre eux.
  - **Manque de gros contrats internationaux** : l'équipe n'a pas eu de contrats de grande taille au niveau international, notamment dans les programmes cadres européens, ERC, etc.
  - **Reconnaissance internationale perfectible :** les membres de l'équipe sont reconnu(e)s dans leurs communautés internationales mais certains indicateurs de reconnaissance restent faibles : comités éditoriaux ; présidences de comités de conférences internationales de rang maximal ; pour l'axe logique, faible présence dans des comités de programme de conférences de rang maximal (mais nous avons un membre de PC de LICS 2023) ; invitations à des conférences internationales (mais nous avons des invitations en 2023).

- **Opportunités**
  - **Nouvelles forces :** l'équipe a récemment recruté des nouveaux membres, et un autre MCF sera recruté en 2023 (dans l'axe logique).
  - **Contexte local :** les deux axes de l'équipe participent désormais à la Fédération MathSTIC. La création de l'axe transversal « Complexités » est aussi une opportunité pour l'équipe, spécialement pour l'axe logique.
  - **Masse critique :** l'axe vérification a retrouvé la masse critique nécessaire pour constituer une équipe indépendante. Ceci nous donne l'opportunité de séparer l'équipe afin d'améliorer la visibilité des deux axes.

- **Menaces**

Trajectoire de l'équipe LoVe

  - **Dispersion dans l'axe logique :** le faible nombre de co-publications entre membres de l'axe logique est un indicateur d'une potentielle dispersion thématique, sur laquelle il faudra veiller.
  - **Vivier rang A réduit :** le vivier de recrutement des rangs A est beaucoup moins peuplé que pour les rangs B, spécialement pour l'axe logique. Ceci risque d'entraver la future croissance des axes/équipes.
  - **Responsabilités administratives lourdes :** les membres de l'équipe ont des responsabilités administrative importantes, au niveau local et national. La nature chronophage de ces tâches risque d'avoir une influence négative sur les activités scientifiques de l'équipe.
  - **Contexte régional :** la présence en Île-de-France d'équipes sur des thématiques similaires et, souvent, de taille plus importante et/ou jouissant d'un contexte économique plus favorable a toujours été et reste une menace pour l'équipe. Cependant, jusqu'ici nous avons réussi à garder notre spécificité et attractivité, et nous veillerons à cela dans le futur.

4.2 Trajectoire

L'équipe LoVe est issue de l'équipe LCR (Logique, Calcul, Raisonnement), qui a changé de nom en juin 2017 pour mieux afficher la présence de deux axes qui la composent.

L'équipe LCR était une équipe historique du LIPN et a toujours eu un caractère « multithématique ». En effet, avant la création de l'équipe CALIN en 2010, l'équipe LCR accueillait aussi un troisième axe de recherche, essentiellement indépendent des autres, portant sur la physique combinatoire.

L'évolution de l'équipe LCR sur le quinquennal 2012–2017 avait permis d'envisager la séparation en deux équipes, correspondantes aux axes « logique » et « vérification » affichés dans le nouveau nom de l'équipe. Pour se préparer à la séparation, une structure interne avec des « responsables d'axe » avait été mise en place (et est toujours existante). Cependant, déjà en 2016 ce projet a été affaibli par le départ d'un MCF de l'axe vérification (mise en disponibilité en entreprise). Le départ d'un deuxième MCF en 2018 (promu PU à Nancy), combiné à la perspective d'un autre départ (à la retraite) en 2020, nous ont conduit à renoncer au projet de séparation, car la nouvelle équipe de vérification n'aurait eu à ce moment que 4 membres permanents.

Aujourd'hui, après l'arrivée de deux MCF, un PU et un IR CNRS, l'axe vérification a retrouvé sa masse critique et nous allons reprendre la démarche de séparation. En termes de stabilité des effectifs, la situation est meilleure qu'en 2016/2017 car il n'y a aucun départ de membres de l'axe vérification envisagé à moyen terme, c'est-à-dire, sur le quinquennal 2023–2028. Cela dit, nous allons procéder avec un minimum de prudence car, avec 8 membres permanents, la nouvelle équipe de vérification sera la plus petite du laboratoire et il faudra s'assurer qu'elle ne naisse pas déjà faible. Concrètement, nous avons décidé d'attendre environ deux ans pour donner le temps à l'axe vérification de se stabiliser après les nouvelles arrivées (les dernières datant de fin 2022). La date envisagée pour la séparation de l'équipe est donc en 2025/2026.

La constitution des deux équipes séparées n'empêchera pas les collaborations, bien entendu. Nous savons déjà que les deux futures équipes interagiront, par exemple, dans le contexte de l'IRN CNRS avec le Danemark, dont la demande à été déposée en 2022, acceptée en 2023 et dont on attend la mise en place début 2024. Cet IRN réunit, du côté danois, des chercheurs des universités de Copenhague et d'Aarhus, et, du côté français, un certain nombre d'équipes dans des unités françaises (IRIF, IMJ, LAMA) et, notamment, plusieurs membres des deux axes de l'équipe LoVe. Scientifiquement, l'IRN est construit autour de la théorie des langages de programmation, des automates et de la spécification et vérification de leur propriétés. La plupart des collaborations attendues se situent entièrement dans un axe/équipe ou dans l'autre, mais le réseau a vocation à stimuler aussi les recherches inter-axe/équipe.

De manière générale, les interactions inter-axes envisageables aujourd'hui au sein de l'équipe LoVe deviendront des interactions inter-équipes, donc rien ne sera perdu dans ce sens. En revanche, la séparation aura des effets positifs par rapport à la structuration et à la visibilité des thématiques des deux axes, qui seront nettement améliorées.

Trajectoire de l'équipe LoVe

Le reste de cette présentation prend acte de la séparation en deux équipes et traitera de façon indépendante leurs trajectoires futures.

4.2.1 Logique, théorie de la programmation et complexité

Avant de discuter du projet scientifique de l'axe/équipe, nous devons faire un état des lieux de ses ressources humaines. En termes de départs et arrivées, les perspectives immédiates sont les suivantes :

départs : 2 MCF travailleront à 100% à l'UNIF à partir de 2023, 1 PU prendra sa retraite en 2025 ;

arrivées : 1 recrutement de MCF en 2023, 1 MCF en échange de poste en 2024.

Si on compte qu'il y aura toujours 1 MCF à 40% (voire moins) qui ne contribuera pas à la recherche, 1 MCF qui sera probablement recruté PU ailleurs dans le futur proche, et 1 PU qui continuera dans ses fonctions de président de l'université, les effectifs réels de l'axe/équipe seront d'environ 10 membres, dont environ 80% rang B. La nouvelle

équipe de logique sera donc essentiellement de la même taille que la nouvelle équipe de vérification.

Avec des effectifs réduits, l'intégration des nouveaux MCF sera cruciale pour le bon fonctionnement de l'axe/équipe. Pour le MCF qui arrivera en échange de poste, il effectuera d'abord un échange de service, à partir de l'automne 2023, et nous aurons donc déjà entamé son intégration au moment de son arrivée officielle. Pour le/la MCF recruté(e) en 2023, le déroulement de l'intégration scientifique, les personnes avec qui il/elle pourra collaborer plus facilement, etc. dépendront, bien évidemment, de la personne qui sera recrutée mais nous porterons une attention particulière à cet aspect.

Il faudra veiller aux capacités d'encadrement de l'axe/équipe : après le départ probable d'un MCF habilité, il ne restera que 3 membres habilités et seulement 2 seront en mesure d'encadrer des thèses sur des sujets fédérateurs pour l'équipe. Cette situation devra être palliée par l'obtention de l'HDR par des membres de l'équipe, ou le recrutement d'un rang A, qui, toutefois, n'est pas évident (le vivier PU en logique est assez réduit).

Pour ce qui concerne l'horizon scientifique, l'axe/équipe continuera à se positionner sur ses trois thématiques principales : sémantique de la programmation, complexité et preuves formelles.

- **Sémantique de la programmation**

La recherche contemporaine en sémantique de la programmation est guidée, entre autres, par :

1. le développement de langages appliqués à des domaines spécifiques, comprenant en particulier la possibilité de prouver la correction des programmes écrits dans de tels langages ;
2. l'introduction de nouveaux outils mathématiques pour la modélisation des programmes.

La direction (2) est en principe indépendente du point (1), mais elle s'y intersecte souvent car ces nouvelles approches sémantiques sont susceptibles d'être appliquées aux domaines dont il est question au point (1).

Trajectoire de l'équipe LoVe

L'axe/équipe s'insérera dans le contexte (1) avec des contributions à la programmation probabiliste et différentiable, deux domaines particulièrement actifs à l'heure actuelle et qui resteront vraisemblablement plutôt dynamiques dans les années à venir, à cause de leurs applications à l'inférence bayésienne, à l'optimisation et à l'apprentissage profond, parmi d'autres. Dans ce cadre, le développement de systèmes de types assurant le bon comportement des programmes est une expertise clé sur laquelle l'équipe pourra s'appuyer pour contribuer à l'avancement de cette direction de recherche. Nous pensons en particulier aux systèmes inspirés de la logique linéaire et ses variantes différentielle ou gradée.

Pour ce qui concerne les aspects sémantiques du point (2), parmi les nouveaux outils auxquels on fait allusion, l'équipe vise particulièrement à développer les suivants :

— outils pour modèles quantitatifs, surtout à base d'espaces vectoriels ou d'analyse fonctionnelle, ou bien de fibrations de catégories pour capturer les types gradés ou les types intersection, dans le but de modéliser, entre autres, des propriétés de programmes probabilistes ou différentiables.

— Catégories de dimension supérieure, en particulier 2-catégories, doubles catégories, et opérades supérieures. Il s'agit ici de la frontière contemporaine de la sémantique dénotationnelle [3, 6, **CI-117**, 1], où l'évaluation des programmes n'est plus interprétée par une égalité (comme dans la sémantique traditionnelle) mais par un isomorphisme ou, plus généralement, un morphisme de dimension supérieure, afin de modéliser plus finement la dynamique des programmes. Nous avons déjà recruté un post-doc ayant une expertise dans cette thématique [7, 2], qui travaillera dans l'axe/équipe jusqu'à décembre 2025.

— Outils pour modèles interactifs, comme jeux ou réalisabilité, surtout à base d'actions de groupe et systèmes dynamiques [**RI-13**], capables de prendre en compte la dynamique des programmes d'une façon différente et complémentaire à la sémantique catégorique supérieure.

Ces recherches pourront aussi certainement donner lieu à des interactions fructueuses avec le LAGA (laboratoire de mathématiques), dans le cadre du nouvel axe « Catégories : entre Calcul et Topologie » de la Fédération MathSTIC.

- **Complexité**

La complexité sera l'autre grande thématique de recherche de l'axe/équipe. Nous avons ici un projet ambitieux, dont les fondations on été jetées pendant la période d'évaluation, qui vise à faire du LIPN (et, en particulier, de la nouvelle équipe de logique) un centre reconnu au niveau national dans le sujet de la complexité algorithmique. En effet, bien que de nombreux chercheurs et enseignants-chercheurs travaillent sur le sujet en France, il n'existe pas d'équipe pouvant s'afficher officiellement de la « théorie de la complexité », malgré la présence de nombreux experts. Ceux-ci sont souvent membres d'équipes dont le thème principal est autre, mais lié aux facettes de la complexité sur lesquels ils travaillent (algorithmique, cryptographie, calcul quantique, théorie des graphes. . .).

Au LIPN, la complexité apparait dans les thématiques de l'axe logique de l'équipe LoVe mais également dans les équipes CALIN et AOC, avec en particulier un nouvel intérêt dans l'équipe CALIN pour les approches algébriques et géométriques. Certains aspects de la complexité pourraient, par ailleurs, toucher aussi les autres équipes/axes du laboratoire. Après l'organisation de plusieurs écoles de recherche sur le sujet (Caleidoscope 2019, 2020 et 2021 ; GCT 2022 ; EPIT 2023), notre équipe a proposé la création d'un axe transversal « Complexités » au LIPN, dont elle assure actuellement la coordination.

La création de l'axe « Complexités » a été un premier pas pour donner de la visibilité à cette thématique. L'axe est centré sur un séminaire quasi-hebdomadaire, auquel participent régulièrement entre 10 et 15 membres du LIPN, sans compter la présence occasionnelle de membres d'autres laboratoires, qui peuvent participer en visioconférence (les séances sont systématiquement diffusées en ligne et enregistrées).

L'axe/équipe continuera à travailler à la réalisation de ce projet de fédération des recherches en complexité au niveau national, dans plusieurs directions :

animation scientifique : nous comptons de pérenniser l'organisation de l'école de recherche « Caleidoscope » (*The Complexity Kaleidoscope*), et de la soutenir avec des invitations, pour des périodes relativement longues (minimum un mois), de personnalités scientifiques de renommée, autour desquelles on pourra organiser, par exemple, des « mois thématiques » sur une des multiples facettes de la complexité. Ces événements scientifiques pourront impliquer, au niveau de leur organisation, des chercheurs français externes au LIPN, provenant par exemple de l'IRIF, de l'IMJ, du LaBRI ou du LAMA (Chambéry), où il existe des personnes ayant exprimé leur intérêt pour ce genre d'activités.

recrutement : nous nous donnerons les moyens pour d'attirer des recrutements de chercheurs/enseignants chercheurs travaillant en complexité. Au niveau de notre axe/équipe, nous souhaiterions commencer par le recrutement d'un PU. Le vivier est très petit, mais des candidats potentiels existent et ont été contactés. L'axe/équipe compte aussi sur le soutien de la direction du laboratoire, dont la politique scientifique encouragera le recrutement de chercheurs et chercheuses avec une coloration « complexité » aussi dans les autres équipes, lorsque ceci sera possible/pertinent du point de vue scientifique. Cela contribuera au renforcement de l'axe transversal « Complexités », dont notre axe/équipe ne pourra que bénéficier, même dans le cas où le recrutement concerne directement une autre équipe.

Le recrutement de post-docs avec une expertise en complexité sera aussi activement poursuivi par l'axe/équipe. Ce sera une façon d'intégrer des compétences

actuellement faibles ou absentes dans l'équipe et dans le laboratoire, et de tisser en même temps des nouveaux liens avec d'autres établissements internationaux déjà bien positionnés sur la thématique (ceux d'où proviendront les post-doctorant(e)s).

avancées scientifiques : l'axe/équipe va naturellement continuer ses recherches en complexité, qui contribueront à renforcer son expertise et son rayonnement au sein de la communauté. Nous envisageons d'avancer sur plusieurs fronts :

— systèmes dynamiques et bornes inférieures : l'axe/équipe porte un projet novateur basé sur le développement des fondements du calcul à partir des systèmes dynamiques (ANR JCJC DySCo, 2022–2027). Parmi les objectifs du projet, nous mentionnons le transfert d'outils et méthodes de la théorie mathématique des systèmes dynamiques pour obtenir des résultats de bornes inférieures en complexité algorithmique, spécialement dans le cadre de la complexité algébrique [8].

— Complexité descriptive et catégories : la complexité descriptive est une théorie bien connue mettant en relation les classes de complexité avec l'expressivité de certains fragments de la logique [4]. Des membres de l'axe/équipe ont récemment entamé un programme de recherche visant à refonder cette théorie sur la logique catégorique [5], remplaçant les classes de complexité par des classes de morphismes dans une certaine catégorie d'objets se comportant comme des « espaces de données ». Une thèse sur ce sujet est actuellement en cours, dont la soutenance est prévue en 2026, et d'autres pourront commencer au cours du prochain quinquennal.

— Complexité implicite : il s'agit d'une direction de recherche qui a progressivement perdu de l'élan ces dernières années, mais il reste des voies prometteuses dont l'exploration vient d'être entamée, notamment pour les classes parallèles [CI-108] ou pour ce qui concerne les applications à l'analyse statique et la vérification de programmes [CI-113].

- **Preuves formelles**

Moins centrale par rapport à la sémantique des programmes et à la complexité, cette thématique restera néanmoins importante pour l'axe/équipe, pour plusieurs raisons. Tout d'abord, les preuves formelles sont un domaine très développé en France, dont l'importance ne diminuera pas dans le futur à moyen/long terme. Deuxièmement, il s'agit d'un thème potentiellement fédérateur, à plusieurs niveaux.

Trajectoire de l'équipe LoVe

Outre les travaux sur la formalisation de l'analyse réelle pour l'analyse numérique et, en particulier, la méthode des éléments finis, en collaboration avec des mathématiciens et logiciens de l'Inria et l'**UTC!** (bibliothèque coq-num-analysis), l'axe/équipe continuera à développer les recherches avec l'axe/équipe de vérification et, plus généralement, avec le reste du laboratoire. Avec l'axe/équipe de vérification il s'agit de prouver formellement des programmes concurrents critiques utilisés dans le domaine de la vérification. Avec le reste du laboratoire, mentionnons par exemple l'interaction des assistants à la preuve avec l'apprentissage artificiel (équipe A3), une direction qui semble particulièrement intéressante pour l'avenir. Il s'agit par exemple d'appliquer les avancées dans le domaine de l'apprentissage pour développer l'automatisation et la recherche de preuves dans les outils d'aide à la preuve. D'autres interactions sont également en cours et à venir, en lien avec les travaux sur les preuves de programmes concurrents, par exemple (équipe AOC).

Additionnellement aux recherches citées précédemment nous envisageons un investissement croissant en terme d'utilisation des assistants d'aide à la preuve dans l'enseignement. Il s'agit d'un sujet émergeant où de véritables questions de recherche se posent actuellement. L'axe/équipe a déjà une experience importante en la matière, et elle contribuera notamment au défi Inria LiberAbaci, consacré précisément à ce type de recherche. Des délégations Inria ont d'ores et déjà été budgétées pour les années à

venir.

Tout comme un groupe de travail transversal sur la complexité a été créé au sein du laboratoire (qui a successivement évolué en axe transversal), nous pensons qu'il serait pertinent de créer un groupe de travail transversal sur les preuves formelles dans divers domaines liées aux recherches du laboratoire, au sein duquel l'axe/équipe jouera naturellement un rôle de coordination.

4.2.2 Vérification

Le projet de l'axe Vérification de l'équipe LoVe s'inscrit dans une dynamique allant des aspects fondamentaux aux applications de la vérification en passant par la proposition d'algorithmes efficaces.

- **Vérification fondamentale**

Nos recherches posent un cadre théorique comprenant à la fois les modèles étudiés, les logiques pour exprimer leurs propriétés, et les nouvelles approches de model checking et de synthèse pour vérifier ces propriétés.

Synthèse de paramètres temporels et de stratégies. Les systèmes cyber-physiques comprennent divers composants fonctionnant en parallèle qui interagissent avec leur environnement, en temps contraint. Les modèles système/environnement peuvent être représentés par des modèles de jeux. Nous souhaitons développer des algorithmes symbolique efficaces pour les modèles de jeux temporisés paramétrés grâce à notre expertise dans le cadre des automates temporisés paramétrés. Il s'agit alors non plus de synthétiser seulement des paramètres temporels, mais aussi des stratégies garantissant le bon fonctionnement du système en présence d'un environnement hostile (adversaire).

Par ailleurs, la définition récente de la logique STCTL [**CI-120**] ouvre de nouvelles perspectives pour la synthèse de paramètres et de stratégies dans les systèmes multi-agents. Le model-checking de telles propriétés dépend en particulier de la sémantique considérée, ouvrant un grand champ de possibilités.

Propriétés de sécurité. Le versant fondamental des aspects sécurité développés dans l'axe Vérification constitue une étape préliminaire en vue de travaux visant à identifier des classes de systèmes quantitatifs permettant de décider si un système est opaque, c'est-à-dire imperméable à ce qu'une attaque déduise des faits devant être gardés secrets. Notre objectif est non seulement de décider si un système est opaque, mais aussi de pouvoir rendre un système opaque, de deux façons : 1) en modifiant certaines valeurs temporelles (*synthèse de paramètres*) et 2) en contrôlant les comportements pertinents à l'exécution (*synthèse de contrôleur*). Ces travaux sont par ailleurs susceptibles d'entrer dans les thématiques de l'ANR BisoUS (2023-2027). Par ailleurs, nous comptons étendre nos travaux sur la supervision de l'opacité à des systèmes probabilistes. L'objectif est tout d'abord de quantifier la propriété d'opacité (actuellement évaluée à vrai ou faux) et ensuite synthétiser un superviseur garantissant une opacité au dessus d'un seuil prédéfini par les experts du système.

Logique de réécriture et vérification. L'intégration de la réécriture-modulo-SMT dans le contexte des théories de réécriture en temps réel offrira la possibilité d'effectuer des analyses correctes et complètes pour les spécifications Real-Time Maude. De plus, en s'appuyant sur nos publications [CO-36] et [**?**], nous envisageons d'intégrer des techniques paramétrées pour la vérification de ces systèmes. Nous prévoyons également l'intégration d'outils (p.ex., des model-checkers parallèles) et des techniques de réduction d'états développés par l'équipe, et offrir un ensemble plus robuste d'outils d'analyse disponibles pour la logique de réécriture.

- **Algorithmes efficaces**

Les algorithmes développés se heurtent à des problèmes d'explosion d'espaces d'états ou d'indécidabilité. Nous développons des stratégies d'exploration d'espaces d'états.

Monitoring et vérification efficace. Sur certains systèmes trop complexes pour être vérifiés intégralement par model checking, la vérification légère (*monitoring*) permet d'apporter des garanties partielles mais formelles sur le système. En ce qui concerne le monitoring, les garanties apportées par méthodes formelles vont concerner non pas toutes les exécutions, mais une exécution ou un ensemble limité d'exécutions. Cela permet une détection de bugs relativement en amont du processus de conception, ou peut au contraire certifier la validité d'exécutions avec échantillonnage partiel.

À partir de nos travaux récents [↑RI-14, ↑RI-13, ↑CI-19], nous souhaitons proposer de nouvelles techniques efficaces pour le monitoring de systèmes cyber-physiques. Les particularités de l'axe Vérification concernent notamment nos capacités à mettre au point des techniques quantitatives (temps, coût [**CI-52**]) mais aussi en présence d'incertitude [↑RI-13, ↑CI-19].

Cette thématique de notre projet est importante par son caractère pivot entre la théorie des méthodes formelles et ses applications pratiques.

Vérification parallèle et distribuée. Nos perspectives dans ce domaine sont d'une part de combiner les méthodes de réduction par ordre partiel avec les structures de données symboliques développés dans l'équipe (le SOG) ; et d'autre part de poursuivre les expérimentations des algorithmes existants de réduction d'ordre partiel sur une large base de données de modèles afin d'identifier des pistes d'amélioration (papier accepté à Petri Nets 2023). Aussi, dans la continuité de nos travaux sur des algorithmes de vérification distribuée œuvrant sur des architectures RDMA [**CI-46**, CI-45, CI-110], nous voudrions adapter d'autres techniques de réduction (p.ex., la réduction par ordre partiel) ou algorithme de vérification (p.ex., la vérification de propriétés temporelles) pour ce contexte.

- **Applications à des domaines concrets**

Les recherches sont également appliquées dans des domaines spécifiques, bénéficiant alors de caractéristiques propres.

Détection de malware. Nos travaux dans la détection de malware ont permis d'analyser des malwares séquentiels. Nous comptons étendre nos approches pour pouvoir traiter les malwares présentant de la *concurrence*, en plus de la récursion et du code auto-modifiant. Nous comptons aussi considérer des modèles plus précis des comportements néfastes, afin de pouvoir détecter plus de malware. Ces travaux feront partie des projets ANR DefMal (2022-2028) et ERGANEO MALWARE (2022-2024). Notons que ce dernier projet est un projet de maturation dont le but est d'améliorer notre détecteur de malware afin de faire une déclaration d'œuvre numérique et de le déposer à l'APP.

Vérification de processus dans la Blockchain Nos traveaux sur la vérification des processus dans la Blockchain [**CI-92**, LO-10, CO-27, CI-109] ouvrent plusieurs perspectives intéressantes dont la prise en compte d'autres propriétés de vulnérabilité des contrats intelligents, l'extension à d'autres langages que Solidity ainsi que l'extension aux applications IoT interagissant avec la Blockchain.

Références complémentaires

[1] P. Clairambault and S. Forest. The cartesian closed bicategory of thin spans of groupoids. In *To appear in Proceedings of LICS*, 2023.

[2] P. Clairambault, F. Olimpieri, and H. Paquet. From thin concurrent games to generalized species of structures. In *To appear in Proceedings of LICS*, 2023.

[3] M. Fiore, N. Gambino, M. Hyland, and G. Winskel. The cartesian closed bicategory of generalized species of structure. *Journal of the London Mathematical Society*, 77(1) :203–220, 2008.

[4] N. Immerman. *Descriptive complexity*. Graduate texts in computer science. Springer, 1999.

[5] P. T. Johnstone. *Sketches of en Elephant. A Topos Theory Compendium. Volume 2*. Oxford University Press, 2002.

[6] P.-A. Melliès. Template games and differential linear logic. In *Proceedings of LICS*, pages 1–13. IEEE, 2019.

[7] H. Paquet and P. Saville. Strong pseudomonads and premonoidal bicategories. *CoRR*, abs/2304.11014, 2023.

[8] L. Pellissier and T. Seiller. Prams over integers do not compute maxflow efficiently. *CoRR*, abs/1811.06787, 2018.

Partie VI

RCLN

Représentation des Connaissances et Langage Naturel

Sommaire

Informations générales sur l'équipe RCLN **291**
1.1 Membres de l'équipe RCLN . . . 292
1.2 Les thématiques scientifiques et leurs enjeux . . . 294
1.3 Profil d'activités liées à la recherche . . . 300
1.4 Prise en compte des recommandations du précédent rapport . . . 300

Introduction du portfolio **303**

Auto-évaluation de l'équipe RCLN **305**
3.1 Domaine 2. Attractivité . . . 306
3.2 Domaine 3. Production scientifique . . . 311
3.3 Domaine 4. Inscription des activités de recherche dans la société . . . 327

Trajectoire de l'équipe RCLN **329**
4.1 SWOT . . . 329
4.2 Trajectoire . . . 331

Overview

For almost three decades, research carried out by the RCLN team has focused on Natural Language Processing (NLP) and knowledge representation. Knowledge is generally transmitted through natural language, and thus Knowledge Representation, a major issue in Artificial Intelligence, Cognitive Science and Semantic Web, is concerned with processing the knowledge expressed in natural language. In a complementary way, text understanding may involve knowledge and reasoning.

The RCLN team brings together skills in automatic language processing, corpus linguistics, semantic web, as well as in data mining and machine learning. These complementary skills give it an original positioning and allow it to carry out innovative work for the analysis and exploration of text corpora as well as for the acquisition of knowledge from texts or knowledge graphs. The team is particularly active in problems that are little studied in language, or complex problems such as the treatment of under-endowed or specialised languages, the detection of neologisms, the analysis of microblogs or the generation and evaluation of abstractive summaries. The team is also working on approaches to deal with the complexity of problems in texts or knowledge graphs, such as syntactic analysis, simultaneous extraction of semantic relations and named entities, or graph completion using methods from deep learning, combinatorial optimisation, data mining or inductive logic programming. The team is structured around three closely related research areas: Algorithms for syntactic and semantic analysis of texts, Exploration and generation of texts and Knowledge Acquisition. These three axes are complementary: the syntactic-semantic analysis of corpora can serve as a basis for text exploration and knowledge acquisition, and conversely the analysis algorithms can take advantage of the knowledge acquired.

These researches have led to results published in the major NLP conferences (EMNLP, NAACL, COLING) and in the fields of semantic web and knowledge acquisition (Semantic Web, KBS, ISWC, ESWC, K-CAP), as well as Data Mining (IDA). They have also led to an important software development: about thirteen software tools are publicly available.

The RCLN research team is strongly involved in the LabEx scientific project "Empirical Foundations of Linguistics" (EFL), in which we play different key coordination roles.

For the next five years, the team will focus on three poles of innovative research: syntactic and semantic analysis taking advantage of deep learning methods and combinatorial optimisation, ressources and analysis methods for under-endowed languages, new usages of large language models (LLM) that are based on knowledge graphs, and knowledge graphs enrichment methods. Taking advantage of the results obtained so far, the team will continue its research on its federative project: scientific literature mining.

Chapitre 1

Informations générales sur l'équipe RCLN

Sommaire

1.1 Membres de l'équipe RCLN 292
1.1.1 Membres permanents 292
1.1.2 Membres non permanents au 01/01/2023 293
1.1.3 Doctorants associés au 01/01/2023 293
1.1.4 Évolution de l'équipe (permanents) 294
1.2 Les thématiques scientifiques et leurs enjeux 294
1.2.1 Algorithmes pour l'analyse syntaxique et sémantique des textes . . 295
1.2.2 Exploration et génération de textes 296
1.2.3 Acquisition de connaissances à partir de textes et de graphes de connaissances 297
1.2.4 Axe transverse concernant la fouille de littérature scientifique . . . 299
1.3 Profil d'activités liées à la recherche 300
1.4 Prise en compte des recommandations du précédent rapport . . . 300

Nom de l'équipe : RCLN Représentation des Connaissances et Langage Naturel (RCLN)

Responsable de l'équipe : Nathalie Pernelle

1.1 Membres de l'équipe RCLN

Corps	2017	2018	2019	2020	2021	2022	(2023)
DR	0	0	0	0	0	0	(0)
CR	0	0	0	0	0	0	(0)
PU	3	2	2	3	3	3	(3)
MCF	10	10	10	10	10	10	(10)
Prag	0	0	0	0	0	0	(0)
IR	2	2	2	2	2	2	(2)
Post-docs et ATER	2	1	1	1	1	1	(0)
Ingénieurs (CDD)	2	0	0	0	1	1	(0)
Doctorants	11	10	5	5	7	8	(7)
Total	*30*	*25*	*20*	*21*	*24*	*25*	*(22)*

1.1.1 Membres permanents

(membres ou, *en italique, anciens membres*, et, **en gras, responsable de l'équipe** au 01/01/2023)

	Nom	*Prénom*	*Situation*	*Institution*
	Audibert	Laurent	MCF	IUTV
	Buscaldi	Davide	MCF	IUTV
	Cartier	Emmanuel	MCF	UFR LLSHS
	Charnois	Thierry	PU	IUTV
	Garcia Flores	Jorge	IR	CNRS
	Gayral	*Françoise*	MCF	IUTV
	Gangemi	Aldo	PU	IG
	Grezka	*Aude*	*IR*	*CNRS*
	Le Roux	Joseph	MCF	IG
	Lévy	François	PU émérite	IUTV
	Nazarenko	Adeline	PU	IG
	Pernelle	**Nathalie**	PU	IG
	Recanati	Catherine	MCF	IG

Informations générales sur l'équipe RCLN

	Rozenknop	Antoine	MCF	IG
	Salotti	Sylvie	MCF	IUTV
	Tomeh	Nadi	MCF	IG
	Zargayouna	Haïfa	MCF	IUTV
	Zarrouk	Manel	MCF	IG

1.1.2 Membres non permanents au 01/01/2023

	Nom	*Prénom*	*Situation*	*Institution*
	Beugnet	Vincent	CIFRE	Thales DMS
	El Khbir	Niama	Contrat Doctoral	LabEx
	Felhi	Ghazi	Allocataire	USPN
	Floquet	Nicolas	Allocataire	USPN
	Liyanage	Vijini	Allocataire	USPN
	Schulz	Alexandre	Contrat Doctoral	Huawei
	Zaratiana	Urchade	CIFRE	Initiativas

1.1.3 Doctorants associés au 01/01/2023

(*Les doctorants associés sont encadrés par des membres du LIPN et dépendent d'une école doctorale extérieure à l'Université Paris 13.*)

Nom	*Prénom*	*Situation*	*Institution*
Abdine	Hadi	Contrat doctoral	Ecole Polytechnique
Aleman Manzanarez	Gerardo	Contrat doctoral	Tec de Monterrey
Annanda	Sousa	Contrat doctoral (SFI)	Univ. Galway
Azevedo	Lucas	Contrat doctoral (SFI)	Univ. Galway
Boukhatem	Nabil Moncef	CIFRE	Ecole Polytechnique
Khajeh Nassiri	Armita	Contrat doctoral	Paris Saclay
Kulic	Ammar	Allocation Doctorale	USPN (ED Erasme)
Lovera	Fernando	Salarié	Univ. Simon Bolivar
Simonne	Lucas	Contrat doctoral	Paris Saclay
Wozniak	Weronica	Cotutelle	Université de Lodz

1.1.4 Évolution de l'équipe (permanents)

Les effectifs de l'équipe RCLN sont restés stables si l'on considère le début et la fin de la période. Le recrutement d'une professeure des Universités en décembre 2019 et d'une maître de conférence en janvier 2020 a permis de compenser un départ en retraite et un départ en disponibilité de longue durée. Ce renouvellement a permis de renforcer les compétences de l'équipe en Acquisition de connaissances et Graphes de connaissances. L'arrivée d'une IR en linguistique en septembre 2018 a permis d'accroître nos compétences en linguistique de corpus. Cependant, il faut noter l'engagement d'Adeline Nazarenko (PU) au CNRS en tant que DAS sur toute la période et la mise en disponibilité d'un an d'un MCF pour raisons personnelles. Ainsi, l'équipe a été particulièrement affectée par l'ensemble de ces départs et engagements en 2018 et 2019.

Départs

- **Sept 2017** A. Gangemi, PU IG (disponibilité)
- **Sept 2018** N. Tomeh, MCF IG (disponibilité)
- **Mars 2019** F. Gayral, MCF IUTV (retraite)

Arrivées

- **Sept 2018** A. Grezka, IR, CNRS (mutation du LDI)
- **Sept 2019** N. Tomeh, MCF IG (retour de disponibilité)
- **Dec 2019** N. Pernelle, PU IG (MCF Paris Saclay)
- **Jan 2020** M. Zarrouk, MCF IG (Senior Postdoc Univ. Galway)

1.2 Les thématiques scientifiques et leurs enjeux

L'équipe RCLN (Représentation des Connaissances et Langage Naturel) réunit des compétences en traitement automatique des langues, linguistique de corpus, web sémantique, ainsi qu'en fouille de données et en apprentissage automatique. Ces compétences complémentaires lui confère un positionnement original et permet de mener des travaux novateurs pour l'analyse et l'exploration des corpus de textes ainsi que pour l'acquisition de connaissances à partir de textes ou de graphes de connaissances. L'équipe se positionne en particulier sur des problèmes peu étudiés en langue ou difficiles comme le traitement de langues sous-dotées, les langues de spécialité, la détection de néologismes sémantiques, l'analyse de microblogs ou encore la génération et l'évaluation de résumés abstractifs. L'équipe s'attache également à définir des approches permettant de gérer la complexité des problèmes dans les textes ou les graphes de connaissance, tels que l'analyse syntaxique , l'extraction simultanée de relations sémantiques et d'entités nommées, ou encore la découverte de règles expressives dans les graphes en faisant appel à des méthodes issues de l'apprentissage profond, de l'optimisation combinatoire, de la fouille de données ou encore de la programmation logique inductive.

Informations générales sur l'équipe RCLN

L'équipe est fortement impliquée dans l'animation et les travaux du LabEx EFL *Empirical Foundations of Linguistics*, et en particulier dans l'axe 5 *Analyse sémantique computationnelle*. Adeline Nazarenko a été co-directrice du labex depuis sa création jusqu'en 2022, co-direction aujourd'hui assurée par Thierry Charnois qui assure également la coordination de l'axe 5 et Emmanuel Cartier a co-piloté les activités de formation jusque février 2023. L'équipe participent activement à la préparation d'un nouveau Labex qui sera soumis début 2024.

L'équipe est structurée autour de trois axes de recherche fortement liés : *Algorithmes pour l'analyse syntaxique et sémantique des textes*, *Exploration et génération de textes* et enfin *Acquisition de connaissances à partir de de textes ou de graphes de connaissances*. Ces trois axes sont complémentaires : l'analyse syntaxico-sémantique de corpus peut servir de base à l'exploration de texte et à l'acquisition de connaissances,

et réciproquement les algorithmes d'analyse peuvent tirer parti des connaissances acquises.

1.2.1 Algorithmes pour l'analyse syntaxique et sémantique des textes

Dans cet axe, l'équipe développe des approches d'analyse syntaxique et sémantique en s'appuyant sur des techniques issues de l'optimisation combinatoire et de l'apprentissage automatique, en particulier de l'apprentissage profond.

- **Analyse Syntaxique**

L'objectif de l'analyse syntaxique de dépendances basée sur les graphes est de découvrir la structure la plus adéquate pour une phrase d'entrée en calculant un score pour toutes les analyses candidates possibles et en renvoyant celle qui a obtenu le score le plus élevé. Cependant, le nombre d'analyses candidates est exponentiel. L'équipe a développé des techniques d'analyse syntaxique pour le TAL fondées sur l'apprentissage profond et les méthodes issues de l'optimisation dans les graphes telles que la relaxation lagrangienne, décomposition, algorithme de Frank-Wolfe, ou encore la montée par bloc de coordonnées pour obtenir des résultats à l'état de l'art en analyse syntaxique [CI-27, **CI-11**, **CI-5**]. Ces méthodes, développées pour certaines d'entre elles en collaboration avec l'équipe AOC, permettent de combiner des contraintes permettant de garantir la bonne formation des structures produites, avec des réseaux de neurones permettant de paramétrer les analyseurs à partir de corpus d'exemples. Les représentations en dépendances peuvent également encoder certaines informations sémantiques de la phrase qui peuvent varier selon les jeux de données. Nous avons tiré partie de deux jeux de données disponibles pour l'arabe dont l'un représente des informations syntaxiques et morpho-syntaxiques et l'autre relativement plus focalisé sur les relations sémantiques. Nous avons proposé un algorithme multitâche d'analyse syntaxique des dépendances, basé sur des LSTM hiérarchiques de type easy-first, capable de décoder une phrase dans plusieurs formalismes. Nous avons montré qu'un tel système conjoint est plus performant que le système de référence mono-tâche sur les banques de dépendances du CATiB et de l'UD Arabic et que les informations linguistiques disponibles dans chaque formalisme aident à faire de meilleures prédictions pour l'autre [**CI-15**]. Enfin, l'équipe a également défini des méthodes de génération non supervisées, en particulier pour le désenchevêtrement (*i.e. disentanglement*) des rôles syntaxiques et de la sémantique lexicale, particularité essentielle pour les travaux cherchant à obtenir des méthodes interprétables. Ces méthodes sont fondées sur les auto-encodeurs variationnels et le mécanisme d'attention croisée [**CI-34**].

- **Analyse sémantique**

En ce qui concerne l'analyse sémantique des textes, nous nous intéressons à la fois à la tâche d'extraction d'entités nommées, et à l'extraction de relations ou d'évènements.

Pour la tâche d'extraction d'entités nommées, nous avons notamment proposé une méthode de filtrage des espaces de recherche de modèles CRF semi-Markovien qui permet leur passage à l'échelle et l'amélioration de leur performance [CI-28]. La classification de relations consiste à trouver les instances des relations sémantiques, c'est-à-dire les ensembles d'entités, et à classer la relation qui les lie selon une typologie existante. Dans un cadre non supervisé, les types de relations doivent être déduits directement des données. Dans ce cadre, nous avons montré comment des plongements de mots issus de modèles pré-entrainés pouvaient être utilisés pour identifier des types de relations sémantiques en utilisant une mesure de similarité relationnelle originale basée sur la combinaison des vecteurs d'entrée et de sortie CBOW de word2vec [CI-4].

L'équipe s'est également focalisée sur la thématique d'extraction d'information spécifique à la langue arabe. Nous avons ainsi développé une collaboration étroite avec le Camel Lab (Univ. de NY, Abu Dabi) pour concevoir des systèmes adaptés à la complexité de cette langue. Les travaux existants se concentrent principalement sur la reconnaissance d'entités nommées et peu de travaux ont été réalisés sur l'extraction de relations et la reconnaissance d'événements en arabe, ce qui constitue une tâche complexe en raison de la richesse morphologique et des idiosyncrasies de cette langue. L'équipe a développé des modèles basés sur de l'apprentissage profond supervisé permettant d'extraire, à partir d'un texte, le graphe d'entités avec leurs relations et les événements avec leurs arguments [CO-60]. Pour atteindre les résultats de l'état de l'art sur ces tâches, nous nous appuyons sur l'analyse morphologique pour répondre aux difficultés spécifiques à l'arabe et utilisons un algorithme de recherche en faisceau différentiable pour permettre l'apprentissage *end-to-end* des paramètres du modèle.

Enfin, nous avons proposé la première tâche sur l'extraction de relations sémantiques à la compétition internationale SemEval 2018. Le défi se concentre sur les relations sémantiques spécifiques au domaine scientifique et comprend trois sous-tâches différentes qui permettent de comparer et quantifier l'effet de différentes étapes de prétraitement sur les résultats de la classification des relations. Cette tâche a attiré un total de 32 participants avec 158 soumissions aux différents scénarios. Les jeux de données ont depuis été utilisés par de nombreuses approches : le track est cité par 173 publications [**CO-23**].

1.2.2 Exploration et génération de textes

Cet axe de recherche réunit des travaux qui cherchent à combiner analyse des textes et connaissances pour classifier, annoter, résumer ou générer des documents.

- **Exploration de textes**

L'exploration des microblogs fait l'objet de nombreuses recherches car ceux-ci peuvent être utilisés pour identifier des tendances, détecter de nouveaux évènements, ou pour prédire des phénomènes physiques et sociaux. L'équipe a abordé le problème de la géolocalisation de microblogs en combinant des architectures neuronales avec des bases de connaissances pour extraire des informations contextuelles, même en l'absence de références géographiques directes [CI-13]. L'algorithme, appelé Sherloc, exploite les toponymes contenus dans le message, extrait leur sémantique d'une ressource géographique et les intègre dans un espace métrique qui capture la distance sémantique entre eux. Les toponymes les plus proches sémantiquement d'un message sont identifiés et regroupés en fonction de leur emplacement spatial. Sherloc ne nécessite aucun entrainement préalable et peut déduire la localisation dans une ville avec une grande précision. D'autres travaux sur les microblogs ont porté sur l'analyse de sentiments [CO-51], ou encore la détection de trolls [CO-45].

Informations générales sur l'équipe RCLN

- **Génération de textes**

Nous nous sommes intéressés à doter les robots de capacités en langage naturel leur permettant de raconter leur expérience dans l'exécution d'une certaine tâche. Ces capacités narratives peuvent être utilisées pour surveiller l'activité d'un robot dans un environnement autonome ou améliorer la qualité de la communication avec des interlocuteurs. Contrairement aux travaux antérieurs sur la verbalisation des robots, qui se concentrent exclusivement sur sa navigation spatiale, ce travail considère les mouvements et actions réalisées, les lieux où la tâche a été exécutée et ses résultats. Nous avons défini un modèle de dialogue narratif, un graphe de situations, qui peut être

instancié en utilisant l'analyse de l'historique des tâches du robot et ses éléments significatifs et qui est utilisé pour générer une phrase et ses connecteurs linguistiques [RI-8].

L'équipe s'est également intéressée à la génération automatique de résumés abstractifs et à leur évaluation. Nous avons proposé un modèle génératif auto-régressif basé sur une architecture neuronale Transformer de type BART spécifique à la langue arabe [**CO-58**] . Ce modèle a atteint l'état de l'art sur plusieurs jeux de données de caractéristiques différentes. L'équipe a également contribué aux méthodes d'évaluation de résumés automatiques. ROUGE est sans doute la métrique la plus largement utilisée, mais celle-ci se base sur les chevauchements lexicaux entre les résumés candidats et les résumés de référence et n'est donc pas adaptée aux résumés abstractifs. Aussi, nous avons développé une méthode permettant d'évaluer la qualité des résumés abstractifs appelée GeSERA, qui généralise la méthode SERA définie pour les données médicales, et qui s'affranchit de la présence de chevauchements lexicaux en utilisant des vecteurs de similarité sémantiques basés sur Wikipédia [CI-25].

Les modèles de génération de texte excellent aujourd'hui dans la production de textes qui paraissent être générés par un humain. Ces outils peuvent être utilisés à mauvais escient et générer par exemple des textes comportant de fausses informations. Proposer des outils capables de distinguer un texte généré automatiquement d'un texte écrit par un humain devient essentiel. Nous avons proposé un jeu de données constitué d'articles originaux et d'articles générés automatiquement mais partiellement modifiés. Ces derniers permettent de simuler une utilisation plus réaliste des outils de génération automatique par rapport aux jeux de données existants dans lesquels les textes sont générés complètement automatiquement [CI-31]. Les résultats préliminaires montrent que les systèmes existants ont moins de facilité pour repérer des textes de ce type.

1.2.3 Acquisition de connaissances à partir de textes et de graphes de connaissances

Dans cet axe, nous nous intéressons aux approches permettant d'exploiter des corpus de textes ou des Graphes de connaissances volumineux pour construire ou enrichir des ressources langagières structurées ou des graphes de connaissances.

- **Construction de ressources langagières**

La ressource lexicale publique Morfetik dont nous sommes responsables permet d'obtenir, pour tout mot français, toutes ses formes (féminin et pluriel, formes conjuguées des verbes, etc.), ou, à l'inverse, d'identifier le mot (i.e. "lemme") correspondant à toute forme fléchie [RI-15]. Cette ressource, construite de manière semi-automatique en collaboration avec des linguistes, décrit aujourd'hui plus de 1000 000 formes grâce à l'introduction d'expressions et de mots composés en 2022. L'interface offre également la possibilité de consulter la ressource interne Néoveille [RE-3], notre plateforme de repérage et de suivi des néologismes en corpus dynamique et d'obtenir ainsi pour le mot recherché une visualisation des contextes dans lequel le mot apparait ainsi que des indicateurs métalinguistiques (pays, journal, domaine). L'équipe a en effet poursuivi ses travaux sur l'identification et la caractérisation des innovations lexicales dans un cadre multilingue, en développant des outils de gestion de corpus dynamiques et en diachronie, de fouille et de suivis lexicaux mobilisant différentes approches, qu'il s'agisse de méthodes d'identification supervisées à base de ressources lexicales, de méthodes de détection des événements rares sur la base de patrons syntaxiques prototypiques ou encore de modèles de langue contextuels [CL-9, RI-9]. L'intérêt s'est porté essentiellement sur les innovations lexicales dites formelles, par identification des procédés de dérivation et de composition.

- **Construction et enrichissement de Graphes de Connaissances**

En collaboration avec quatre universités européennes (Univ. Cagliari, Univ. Seville, Open Univ. of Milton Keynes, Univ. of Milano Bicocca), l'équipe RCLN a défini l'approche SCICERO [RI-20] qui permet de générer des graphes de connaissances à grande échelle, sémantiquement riches et précis à partir de la littérature scientifique, tels que CS-KG pour l'informatique (41 millions d'énoncés extraits de 6,7 millions de résumés) [CI-35], ou AI-KG [**CI-16**]. Cette approche peut être combinée avec l'approche de complétion de graphes SciCheck qui incorpore un ensemble de caractéristiques et d'heuristiques conçues pour capturer les connaissances scientifiques [RE-6] et qui est plus précise que l'état de l'art des approches basées sur les plongements ou sur des méthodes classiques de TAL. Le grand nombre d' articles scientifiques publiés chaque année, implique une difficulté croissante à suivre les nouveaux thèmes des différents domaines scientifiques et leur importance. Nous avons également proposé une stratégie de construction automatique de taxonomies de thèmes permettant d'exploiter des millions de résumés d'articles scientifiques en nous appuyant sur des modèles simples qui sont plus faciles à évaluer automatiquement, mais en améliorant nos modèles grâce à l'application d'un grand nombre d'itérations [CI-36].

Les données scientifiques constituent également l'une des sources de connaissances pertinente pour la recherche d'experts et de leur expertise. Les méthodes existantes de recherche d'experts dans les publications ne tiennent pas compte de la validation implicite par les pairs exprimés par les liens de collaborations. En collaboration avec l'équipe A3, nous avons proposé une méthode originale qui extrait l'expertise et les liens de collaborations d'un corpus de publications scientifiques annoté sémantiquement par le biais d'une ontologie, représente les informations extraites sous la forme de graphes bipartites attribués, et enfin énumère des motifs fermés apparaissant dans des noyaux (i.e Core closed pattern mining) en introduisant des critères de sélection originaux adaptés à la recherche d'experts et à leur expertise [CI-21, CO-20].

L'équipe RCLN s'est également intéressée aux approches de fouille de règles dans les graphes de connaissances. Ces règles peuvent venir enrichir le graphe de connaissance et être utilisées pour rendre explicite les régularités dans les données, compléter ou intégrer les données, détecter des données erronées, ou concevoir des systèmes hybrides. Dans les graphes de connaissances, les approches de découverte de règles doivent faire face au volume de données et à l'hypothèse du monde ouvert (OWA) qui ne permet pas de disposer facilement de contrexemples. Nous avons étudié les différences théoriques et expérimentales entre différentes sémantiques de règles de liage [RI-12], qui varient en fonction de la façon de prendre en compte l'hypothèse du monde ouvert. Dans le cadre du projet PSPC AIDA, nous avons collaboré avec le LISN pour concevoir le système RE-Miner qui découvre efficacement des règles de liage basées sur des expressions référentielles expressives représenter sous forme de motifs de graphes [**CI-14**]. Ces motifs caractérisent un individu de manière unique et nous avons montré sur différents jeux de données réels que cette approche surpassait les approches de liage à base de règles existantes (approche arrivée 1ère en terme de F-Mesure pour un track OAEI 2020) [CO-42]. Nous avons également montré que les liens erronés générés par ces approches automatiques pouvaient être en partie détectés grâce à une approche uniquement basée sur les communautés pouvant être détectées dans le graphe formé par les liens d'identité [RI-13] (collaboration avec la Vrije Univ. d'Amsterdam). De même, nous avons collaboré avec l'équipe A3 pour définir une approche nommée CRA-Miner qui permet de découvrir des règles de prédiction de la présence d'amiante dans les éléments de construction pour le CSTB (Centre Scientifique et Technique du Bâtiment) [CI-24], en s'inspirant d'approches top-down issues de la Programmation Logique Inductive mais en offrant la possibilité de prendre en compte des informations temporelles et les liens de subsomption de manière efficace. Enfin, peu d'approches s'intéressent à la découverte de règles causales dans les graphes de connaissances. Nous avons collaboré avec l'INRAE de Montpellier pour définir les différentes variantes du système DICARE

Informations générales sur l'équipe RCLN

(Dicare-E, Dicare-C et Dicare E) qui découvrent des règles différentielles causales représentées en logique du premier ordre, en se basant sur le cadre du potenital outcome framework où la valeur du contrefactuel est estimée [**CI-20**, CI-29]. Le système a été utilisé pour expliquer pourquoi certaines personnes veulent réduire leur consommation de viande plus que d'autres en utilisant un graphe de connaissance représentant des données collectées à l'INRAE .

1.2.4 Axe transverse concernant la fouille de littérature scientifique

Les publications de recherche ou les ensembles de données partagées devraient être mieux exploités par des systèmes intelligents pour soutenir et accélérer les efforts scientifiques, en facilitant la recherche d'experts, la génération d'état de l'art, la génération de nouvelles hypothèses scientifiques, ou pour soutenir ou contredire certaines affirmations. Nous avons défini la fouille de littérature scientifique comme un axe transverse de l'équipe RCLN depuis 2017. Ces textes emploient une (des) langue(s) de spécialité et introduisent différentes difficultés : (1) le nombre de publications scientifiques croit de manière exponentielle et cela rend les possibilités d'analyse manuelle de plus en plus difficile ou parcellaire (e.g. le nombre d'études publiées sur le changement climatique a été par exemple multiplié par 100 depuis 1990), (2) le vocabulaire y est particulièrement évolutif que ce soit pour exprimer des entités désignant des méthodes, des outils, des organisations ou des objets d'étude, (3) les relations et entités sont souvent ambigües, difficiles à segmenter et enfin la recherche d'expert pour un domaine de spécialité se complique également avec le volume des données et la qualité variable des publications. Comme cela apparaît dans les descriptions des axes présentées ci-dessus, nous avons défini différentes approches qui portent sur la fouille de littérature scientifique et qui peuvent être employées de manière complémentaires. Ces approches vont de l'extraction de relations à la création de taxonomies ou de graphes de connaissances à partir de résumés d'articles scientifiques, en passant par une énumération et une caractérisation des experts de domaine prenant en compte une forme de validation par les pairs. De plus, nous avons constitué des ressources permettant à d'autres approches de tester leur capacité à traiter des textes de littérature scientifique (i.e. track SemEval) et nous avons construit et mis à disposition des graphes de connaissances générés à partir de millions de résumés de publications du domaine de l'informatique (CS-KG) et plus spécifiquement pour le domaine de l'intelligence artificielle (AI-KG).

1.3 Profil d’activités liées à la recherche

Activités (répartir 100 points sur ces 7 items)	
Administration et animation de la recherche : pilotage de la recherche (VP, direction d’institut, DAS, par exemple), participation à des instances d’évaluation (CNU, CoNRS, CSS, Hcéres, par exemple), responsabilité de dispositifs Idex ou Isite, direction de projets (ANR, Horizon Europe, ERC, CPER, PIA, France 2030, par exemple), responsabilités éditoriales dans des revues ou collections nationales et internationales.	40
Aide aux politiques publiques et expertise technique : pouvoirs publics aux niveaux européen, national et régional, entreprises, instances internationales comme FAO, OMS, etc.	0
Contribution à l’adossement d’enseignements innovants à la recherche : EUR, SFRI, etc.	0
Dissémination de la recherche : partage de connaissances avec le grand public, médiation scientifique, interface sciences et société	5
Recherche et encadrement de la recherche.	45
Valorisation, transfert, innovation.	20
Autres activités. (à préciser en une ligne maximum).	0

1.4 Prise en compte des recommandations du précédent rapport

Dans cette section, nous présentons les actions entreprises pour tenir compte des recommandations HCERES données à la précédente évaluation.

— Manque de collaborations intra-équipe : pour favoriser les collaborations intra-équipe, nous avons défini un projet transverse concernant la fouille de littérature scientifique, collaboré pour constituer ou participer à des tracks tels que la campagne d’évaluation SemEval 2017, SemEval 2018, le Défi national DEFT 2019, qui implique à chaque fois deux ou trois membres permanents, incité les maîtres de conférences à participer au co-encadrement des thèses dirigées par des habilités et enfin 5 membres collaborent pour définir et implémenter la plateforme ChêneTal qui doit permettre aux linguistes d’accéder et de combiner certains des outils développés au sein de l’équipe. Ainsi, sur cette période, plus d’un tiers des articles de conférences internationales ont parmi leurs auteurs 2 ou 3 membres permanents RCLN.

— Manque de collaborations internationales : sur cette période 6 thèses ont été co-encadrées avec des laboratoires étrangers. Sur les 21 revues internationales publiées, 17 ont des co-signataires membres de laboratoires étrangers, de même sur 36 conférences internationales, 13 sont effectuées avec des collaborateurs étrangers.

— Manque de projets internationaux : nous avons pu augmenter le nombre de projets internationaux portés par l’équipe en incitant les enseignants chercheurs de l’équipe à inviter des chercheurs étrangers (8 enseignants chercheurs invités sur la période), en formaliser ces collaborations internationales via des projets nationaux tels que des projets émergence permettant de financer des missions, puis via des projets bi-latéraux, pour ensuite participer au montage de projets plus conséquents. Sur la période précédente, seulement 2 projets bilatéraux avaient été menés (Ecos-Nord et AUF-PIRAT). Sur cette période, un projet émergence implique le Camel Lab (Trauma), et 5 projets bilatéraux ont vu le jour (4 projets PHC et 1 projet Ecos-Nord). Nous sommes également partenaires d’un projet Horizon CL-23 qui vient d’être déposé en 2023.

— Risque concernant la diminution des ressources suite à une fin possible du Labex EFL : la fin du Labex EFL aurait eu en effet pour conséquence une importante diminution des ressources de l'équipe RCLN. Le Labex a été reconduit jusque 2024 et l'équipe s'implique fortement cette fois encore pour que le Labex soit renouvelé en 2024. En parallèle, nous avons augmenter nos soumissions à des appels à projet et nos collaborations industrielles ce qui a conduit à une augmentation des thèses Cifres et des projets nationaux auxquels nous participons.
— Manque d'HDR : nous sommes conscients que le nombre d'HDR est trop faible compte tenu de la taille de l'équipe : un maître de conférence a soutenu son habilitation en 2018 en linguistique, mais seulement les trois membres PU sont habilités en informatique au 1er janvier 2023 dont une qui est extrêmement prise par ses responsabilités au sein du CNRS. Nous avons incité les enseignants chercheurs à co-encadrer les thèses menées au sein de l'équipe ainsi que des thèses co-encadrées avec d'autres laboratoire et deux maîtres de conférences sont actuellement en train d'écrire leur manuscrit en vue de soutenir leur habilitation avant la fin de l'année.
— Durée des thèses trop élevée : nous avons été plus attentifs à la durée des thèses qui dépassait les 4 ans en moyenne sur la période précédente. Nous avons veillé à une bonne intégration des doctorants au sein de l'équipe, et en particulier à leur participation active au groupe de lecture et aux séminaires de l'équipe. La durée a pu être ramenée à 43,9 mois sur cette période.
— Vie de l'équipe non explicitée : L'équipe se réunit mensuellement, permanents et non-permanents, pour discuter des appels à projet, des évènements liés au Labex, des besoins en poste de permanents ou d'ATER, des recrutements de docteurs et post-doctorants, et des besoins en équipement. Un groupe de lecture bi-mensuel a été mis en place depuis la dernière période, qui est animé par Joseph Leroux et Nadi Tomeh, et les séminaires sont organisés par Davide Buscaldi. L'équipe est représentée au conseil de laboratoire par deux membres de l'équipe, et par trois membres au comité d'expert pour la section 27.

Chapitre 2
Introduction du portfolio

L'équipe RCLN propose 5 portfolios : (1) Higher-Order Dependency Parsing for Arc-Polynomial Score Functions via Gradient-Based Methods and Genetic Algorithm, (2) Traitement automatique de la langue arabe (3) Generating Knowledge Graphs from Scientific Literature (4) Campagne d'évaluation en extraction et classification de relations sémantiques dans les papiers scientifiques, et enfin (5) Morfetik, ressource morphologique.

Les portfolios 1 et 3 ont été choisis afin de présenter deux contributions scientifiques de l'équipe en analyse syntaxique et en acquisition de connaissances à partir de texte, contributions qui sont au coeur des travaux de l'équipe dans deux axes différents, et qui ont été valorisées dans des conférences et journaux de haut niveau.

Le portfolio concernant le traitement automatique de la langue arabe est un exemple permettant d'illustrer nos travaux actuels sur les langues sous-dotées, recherches que nous menons en collaboration avec le Camel Lab.

Les portfolios 4 et 5 sont des exemples de ressources que nous partageons avec la communauté scientifique. Il s'agit tout d'abord de la première campagne internationale d'évaluation sur l'extraction de relations dans les textes scientifiques, campagne qui a impliqué de nombreux participants et dont les ressources sont réutilisées jusqu'à aujourd'hui par les approches d'extraction pour évaluer leurs résultats. Enfin, le portfolio 5 présente une ressource morphologique qui est surtout à destination des linguistes. Morfetik permet d'accéder à l'ensemble existant le plus complet des formes simples et

composées du français, avec leurs variations.

Chapitre 3

Auto-évaluation de l'équipe RCLN

Sommaire

3.1 Domaine 2. Attractivité 306

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche. 306

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels. 308

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs. 309

3.1.4 Référence 4. L'équipe est attractive par la qualité de ses équipements et de ses compétences techniques. 310

3.2 Domaine 3. Production scientifique 311

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité. 311

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels. 311

3.2.3 Publications 312

3.3 Domaine 4. Inscription des activités de recherche dans la société . 327

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique. 327
3.3.2 Référence 2. L'équipe développe des produits à destination du monde culturel, économique et social. 328

3.1 Domaine 2. Attractivité

3.1.1 Référence 1. L'équipe est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche.

Le rayonnement scientifique est visible à travers les invitations dans les colloques internationaux ou nationaux. Les membres de l'équipe sont aussi très régulièrement sollicités pour participer à des comités de programme des conférences majeures au niveau international ou pour l'organisation de manifestations scientifiques dans les domaines du TAL, du web sémantique, ou encore de la linguistique, ce qui témoigne de leur reconnaissance dans ces domaines. Il se traduit aussi par une forte implication dans le pilotage de la recherche notamment au niveau national (CNRS) ou local (LabEx EFL).

Ces éléments sont détaillés ci-dessous.

Invitations dans des colloques ou conférences Les membres de l'équipe sont régulièrement invités à présenter leurs travaux dans des colloques, conférences ou ateliers de niveau international ou national :

- Conférence CNRIA, Thies (Sénégal), 15 au 17 juin 2022 ;
- Spanish Conference on Information Retrieval *CERI 2018* à Saragosse (Espagne) du 25 au 27 juin 2018 ;
- 7th International Symposium on Language & Knowledge Engineering *LKE 2019*, Dublin (Irlande), octobre 2019 ;
- Workshop international *DeepOntoNLP 2020*, Bolzano (Italie) le 16 septembre 2020 ;
- Webinaire *DOING network*, LIFO, Université d'Orléans, le 17 juin 2021 ;
- Journée d'étude à l'Université de Niigata (Japon), 6 mars 2017 ;
- Journée d'étude à l'Université de Tokyo (Japon), 8 mars 2017 ;
- Journée d'étude à l'Université de Nagoya (Japon), 16 décembre 2017 ;
- Séminaire de l'Université de technologie de Chiba, Chiba-kodai (Japon), 26 octobre 2018 ;
- Colloque International La Perception : langue, discours, cognition, 6 et 7 décembre 2019, Sorbonne Université ;
- Webinaire de l'École Doctorale des Pays de Vysegràd, Université Jagellonne de Cracovie et Institut de Néophilologie de l'Université Pédagogique de Cracovie, le 23 et 24 septembre (2021).

Auto-évaluation de l'équipe RCLN

Responsabilités éditoriales

- Comité éditorial et de rédaction du périodique Buscila-infos (Bulletin d'information en Sciences du langage), jusqu'en 2019 ;
- Comité scientifique de la collection Dixit Grammatica, L'Harmattan (Paris). Revue internationale électronique de grammaire, linguistique et langue françaises, depuis 2015 ;
- Comité de rédaction de la revue Traitement automatique du langage (3 membres de l'équipe sur la période).

Participation à des comités de programme de conférences internationales importantes. L'équipe est très investie dans les PC (comités de programme) des conférences internationales de premier plan, ainsi que dans les principales conférences nationales du domaine.

- 4 Area Chair (COLING 2020, ACL 2020, COLING 2022, CoNLL 2022)
- Action Editor pour ACL Rolling Review (depuis 2021)
- plus de 40 PC ou comités scientifiques de conférences internationales (en tant que membres) : IJCAI 2019, IJCAI-ECAI 2018,EMNLP-IJCNLP 2019,NAACL 2019,ACL 2020,AACL-IJCNLP 2020,EMNLP 2020,ACL 2021,NAACL 2021,EACL 2021,COLING 2022,CONLL 2022, EMNLP 2022, ECML-PKDD 2016 et 2017, LREC 2016, 2018, 2020, 2022, ESWC 2017, 2019, 2020, ISWC 2017, 2022, WebConf 2017, 2018, 2019, 2020, 2021, 2022, BioNLP-OST 2019, WSDL 2022, EKAW 2020,21,22, ESWC 2020,21,22, ICCS 2020, 21, 22 ;
- PC de conférences nationales : TALN, EGC, IC, CORIA, CNRIA

Organisation de conférences ou ateliers internationaux

- 3 Co-chairs de la tâche 7 (Semantic Relation Extraction and Classification in Scientific Papers) du challenge Semeval 2018 ;
- 2 Co-chair du track Knowledge Management & Ontology engineering de la conférence IEEE AICCSA (2017, 2018) ;
- 2 comités de pilotage SemEval 2020 (en tant que membre) ;
- 4 co-organisations du workshop "Deep Learning for Knowledge Graphs" (DL4KG2020) associé à ESWC ou ISWC (2019, 2020, 2021, 2022) ;
- 4 co-organisations du workshop "Sentic Computing, Sentiment Analysis, Opinion mining and Emotion Detection", satellite workshop at ESWC, (2018) ;
- 1 co-organisation du workshop "Interactions between Data Mining and Natural Language Processing" associé à ECML-PKDD, (2017) ;
- 5 co-organisations (Publication Chair) du workshop "Arabic Natural Language Processing", (2017, 2019, 2020, 2021, 2022) ;
- 1 co-organisation du workshop "Processing Language Variation : Digital Armenian (DigitAm)" associé à LREC (2022) ;
- 2 co-organisation du workshop "Language Technology for Equality, Diversity, Inclusion" associé à EACL 2021 et ACL 2022", (2021, 2022) ;
- 1 co-organisation du congrès international FRASEOLOGÍA Y TRADUCCIÓN EN HISPANOAMÉRICA, Université Ricardo Palma, (2018).

Responsabilités dans les sociétés savantes

- 1 membre du comité de pilotage (Advisory Committee) de l'ACL Special Interest Group on Arabic Natural Language Processing (depuis 2022) ;
- 1 vice-présidence de l'ATALA - Association pour le Traitement Automatique des Langues - (2017 à 2021) ;
- 2 membres du collège SIC de l'AFIA (depuis 2015 et depuis 2018) ;
- 1 membre du comité d'orientation de l'ARIA (depuis 2016) ;
- 1 membre du comité d'administration de l'ATALA (2017- 2021) ;
- 1 membre du comité de pilotage de la conférence nationale IC (depuis 2018).
- 1 membre du comité de pilotage collège HLT AFIA (depuis 2019).

Auto-évaluation de l'équipe RCLN

Participation à des instances de pilotage ou évaluation de la recherche L'équipe est fortement investie dans les instances de pilotage de la recherche avec des responsabilités scientifiques ou administratives importantes notamment au niveau du CNRS et du LabEx EFL, ainsi que dans de multiples instances d'évaluation telles que l'HCERES, l'ANR.

Participation à des instances de pilotage

- CNRS : Direction adjointe scientifique INS2I CNRS (2016-2022) ;
- Responsabilité de la mission pour l'expertise scientifique du CNRS (2022) – puis Direction de l'institut INS2I depuis février 2023 – ;
- Co-direction du Laboratoire d'excellence "Fondements Empiriques de la Linguistique", LabEx EFL (depuis 2011) ;
- Responsabilité de l'axe 5 "Analyse sémantique computationnelle" du LabEx EFL (depuis 2011) ;
- Responsabilité de la formation du LabEx EFL (2018-2022) ;
- 2 Membre élus de la Commission Recherche et du Conseil Académique de l'université Sorbonne Paris Nord (2016 – 2020) et (2021-2022) ;
- Membre élu du Conseil d'Administration de l'université Sorbonne Paris Nord (2018-2020) ;
- Co-responsable axe "Accès à l'information et fouille de textes" du GDR TAL (depuis 2018).

Participation à des instances d'évaluation de la recherche

- Membre extérieur Conseil Scientifique LabEx « Les passés dans le présent » (depuis 2019) ;
- Membre extérieur Conseil Scientifique du LIRMM (depuis 2022) ;
- Comité HCERES (1 présidence et 1 membre expert), évaluation Laboratoire Informatique d'Avignon, (2022)
- Membre du comité d'évaluation scientifique CE23 Intelligence Artificielle de l'ANR, Appel à projets générique (2018, 2020 et 2022) ;
- Experte pour les campagnes d'avancement des IT CNRS, commission régionale d ?interclassement de la BAP D (circonscription Ile de France Villejuif) ;
- Membre du comité d'évaluation des chercheurs du CIRAD Montpellier commission ACACONGUA (2022).

En outre, les membres ont été impliqués dans les responsabilités liées à l'enseignement :

- Vice présidente du département d'informatique de l'Institut Galilée (depuis 2021) ;
- Cheffe du département GEA de l'IUT de Villetaneuse (2016–2021) ;
- Responsable du BUT GEA (depuis 2021).

3.1.2 Référence 2. L'équipe est attractive par la qualité de sa politique d'accompagnement des personnels.

Sur la période considérée, l'équipe a vu l'arrivée d'une ingénieure de recherche CNRS en sciences du langage (début 2018) par mutation interne, d'une maîtresse de conférences (décembre 2019) et d'une professeure (début 2020), toutes deux en section 27 du CNU. En revanche, l'équipe peine à attirer des candidats sur les postes CR CNRS, difficultés qui peuvent s'expliquer par un vivier réduit et surtout du fait des perspectives offertes dans l'industrie, sur les domaines du traitement des langues et de l'apprentissage, c'est-à-dire de l'intelligence artificielle, un domaine, actuellement et depuis quelques années, particulièrement actif et d'actualité.

Auto-évaluation de l'équipe RCLN

Postdoctorants, doctorants et ingénieurs. Sur la période on compte 19 doctorant(e)s dont 7 en cours, 11 soutenances (au 31 décembre 2022) et un abandon. L'équipe a su remarquablement diversifier l'origine des financements de ces thèses. Les allocations ministérielles représentent 42% (8 thèses), le reste des financements se composent de 2 allocations LabEx EFL, 1 double culture (ex COMUE SPC), 1 ANR, 1 FUI, 3 financements extérieurs (Huawei, CEA et IRSTEA) et 3 CIFRE. La durée moyenne des thèses sur la période est de 43,9 mois dont 2 thèses qui ont été menées difficilement à terme et dont la durée a dépassé 4 ans. La situation sanitaire dûe à la pandémie en 2020 et 2021

a clairement eu un impact sur cette durée. En effet, des doctorants se sont retrouvés confinés dans des petits studios durant plusieurs mois, et les contacts par visio que nous avons maintenus n'ont pas remplacé une présence en laboratoire. Cependant comparée à la dernière période, la durée moyenne des thèses s'est nettement améliorée : elle était de plus de 4 ans dans la période précédente. À l'issue de leur soutenance un doctorant a effectué un postdoc à l'étranger et a été recruté comme Maître de conférence (il a été aussi classé à l'issue du concours CR CNRS). Tous les autres autres doctorant(e)s ont été recrutés dans l'industrie. On peut aussi noter que la quasi-totalité des doctorants (18 sur 19) sont co-encadrés par un MCF, ce qui permet aux non-HDR de participer à cette activité.

Les membres de l'équipe participent aussi aux co-encadrements de doctorant(e)s d'une école doctorale hors Galilée dans le cadre de co-directions ou de co-encadrements de thèses dont 6 à l'étranger (Vénézuela, Mexique, 3 Irlande et Pologne), celle de Pologne étant en co-tutelle, auxquels il faut ajouter 2 thèse à LIX, 3 à Orsay et 1 à l'école doctorale Erasme à l'USPN. Sur la période sont ainsi recensés 12 doctorant(e)s hors école doctorale Galilée dont 10 en cours au 31 décembre 2023.

L'équipe a par ailleurs recruté 4 postdoctant(e)s dont 3 femmes, sur des financements LabEx, FUI, et projet IDEX Neoveille. Deux postocs ont ensuite obtenu un poste de Maître de Conférences.

L'équipe est particulièrement attentive au suivi et au devenir des personnels contractuels recrutés (doctorant(e), postocs et ingénieurs sur CDD). Elle les accompagne dans leurs démarches, les conseille, relis leur dossier et les prépare à des auditions dans le cas d'un concours. Ces personnels sont aussi intégrés dans l'équipe sous forme d'invitations aux réunions d'équipe, d'invitations à présenter leurs travaux en séminaire ou encore à participer aux groupes de lecture organisés régulièrement (à un rythme bimensuel ou mensuel). Les doctorant(e) et postdocs sont aussi financés par l'équipe sur projets ou par le laboratoire pour présenter leurs travaux publiés dans les conférences, et sont encouragés à participer à une école d'été. L'équipe bénéficie de deux ingénieurs de recherche permanents qui font partie de la DevTeam au niveau du laboratoire et apporte leur support technique à tous les personnels de l'équipe. Les deux IR participent aux travaux scientifique de l'équipe. Arrivée au début de la période (janvier 2018) par mutation interne, une IR a récemment été promue en classe exceptionnelle, ce qui est un fait notable et témoigne de la qualité de son intégration.

Invitations de chercheurs. L'équipe invite régulièrement des chercheurs provenant de laboratoires internationaux pour développer des collaborations scientifiques ou dans le cadre de co-encadrements de doctorant(e). Ces invitations sont financées sur projets, sur fonds propres du LIPN ou par l'Université Sorbonne Paris Nord. La liste ci-dessous donne les invitations d'au moins 2 semaines et est limitée aux invités internationaux :

- Yudith Cardinale, Univeristé Simon Bolivar, Vénézuela, 2019 ;
- Raoul Monroy, TEC de Monterrey, Mexique, 2022 ;
- Adam Wyner, Swansea University, UK, 2017 et 2022 ;
- Franseco Osborne, Open University, Milton Keynes, UK, 2018 et 2019 ;
- Mustafa Jarrar, Bir Zeit University, Palestine, 2018 ;
- Nizar Habach,New York University Abu Dhabi, UAE, CamelLab, Abou Dhabi, 2018 ;
- Fernando Perez-Tellez, Tecnicol University Dublin, Ireland, 2022 ;
- Emiliano Sanchis, Universitat Politècnica de València, Espagne, 2021.

Auto-évaluation de l'équipe RCLN

3.1.3 Référence 3. L'équipe est attractive par la reconnaissance de ses succès à des appels à projets compétitifs.

Les résultats de l'équipe en matière de réponse à des appels à projets sont concluants :

- Au niveau international, 4 projets bilatéraux PHC (2 Polonium, 1 Barrande, 1 Galilée) 25k€ au total, 1 ECOS-Nord ;

- Au niveau PIA :
  - 1 projet IDEX ex COMUE Sorbonne Paris Cité, Neoveille (2015-2017), 211k€ ;
  - 1 LabEx EFL (2011-2014), plus de 470 k€ depuis 2017 pour l'équipe ;
- Au niveau national :
  - 1 ANR PARSITI (2016–2021), 134k€ ;
  - 1 ANR DALIH (2021–2025), 130k€ ;
  - 2 DGLFLF (Délégation Générale à la Langue française et aux Langues de France) Ministère de la Culture, (2017–2018 et 2017–2020), 64k€ au total ;
  - 1 émergence TRAUMA CNRS (2020), 10k€.

  On peut aussi considérer le projet FUI PCUv2 avec des industriels et les financements CIFRE (voir en section 3.3.1).

Sur la période, l'ensemble de ces projets a ainsi permis de financer 5 thèses, 4 post-doctorant(e)s et 2 ingénieurs d'étude, et plus larement de dynamiser les activités de l'équipe, comme le montre le financement LabEx.

En effet, l'équipe RCLN est fortement impliquée dans le pilotage et les travaux du LabEx EFL dont l'ensemble des activités de recherche s'inscrivent dans le projet scientifique du LabEx. Dans ce cadre, plusieurs opérations financées par le LabEx et portées par des membres de RCLN ont clairement contribué à renforcer les thématiques de l'équipe par le biais de recrutements de postdoctorants (trois sur la période financés à 100%) et doctorants (trois dont deux financés entièrement), de missions, d'équipement (serveur de calcul GPU), de bourses de mobilité entrante (5 invitations de doctorants à l'international) et de soutiens de colloques (2 ateliers internationaux, DL4KG 2019 2020) - l'ensemble de ces financements s'élevant à plus de 470 k€ pour la période considérée (366k€ engagés et 104k€ obtenus et à engager)-. Ce contexte a également favorisé le développement de collaborations au sein du LabEx mais aussi à l'international. La visibilité du LabEx a aussi permis à l'équipe d'attirer de très bons candidats pour les recrutements doctoraux et postdoctoraux.

3.1.4 Référence 4. L'équipe est attractive par la qualité de ses équipements et de ses compétences techniques.

Auto-évaluation de l'équipe RCLN

À l'initiative de l'équipe RCLN, le laboratoire a développé depuis 2014 une infrastructure de calcul dédiée à l'intelligence artificielle et au traitement automatique des langues. Aujourd'hui, cette infrastructure met à disposition des chercheur·euse·s 5 serveurs de calcul avec des accélérateurs GPU, pour un total de 140 coeurs, 1,2 To de RAM et 26 GPUs (8 A40, 11 Quadro RTX 5000, 3 RTX2080 et 4 RTX2080Ti). Cette infrastructure rend service à une cinquantaine d'utilisateur·rices appartenant au LIPN, le L2TI, le LabEx EFL ainsi qu'à des collaborateur·rice·s internationaux·ales. L'infrastructure permet de mettre au point les paramètres pour entraîner des grands modèles de langues (LLM) avant d'éventuellement continuer l'entraînement sur des infrastructures nationales, comme Jean-Zay. Les serveurs sont administrés par un IR de l'équipe RCLN et une IE de l'USPN. Quant aux démos et systèmes mis en ligne par l'équipe, le serveur (80 coeurs et 640 Go de RAM) sert de vitrine à la mise en ligne des logiciels Neoveille (détection de néologismes), Morfetik (ressource morphologique) et SDMC (extraction de motifs). Il hébergera aussi la plateforme de mise en ligne de chaînes de traitements ChêneTAL, en cours de développement. Ce serveur est ouvert aussi aux collaborateur·rice·s internationaux·ales de l'équipe RCLN, et il compte plus de 120 utilisateur·rices inscrit·e·s.

3.2 Domaine 3. Production scientifique

	2017	2018	2019	2020	2021	2022	2023	*Total*
Livres	0	1	0	0	0	0	0	*1*
Chapitres de livre	3	3	1	3	1	0	0	*11*
Revues internationales	7	2	2	5	2	3	0	*21*
Autres revues	0	1	2	0	2	2	0	*7*
Conférences internationales	7	3	3	4	9	10	0	*36*
Autres communications	18	14	6	10	4	10	2	*64*
Total	*35*	*24*	*14*	*22*	*18*	*25*	*2*	*140*

3.2.1 Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité.

Comme le montrent ses publications, l'équipe est reconnue pour ces travaux en analyse syntaxique et sémantique qui s'appuient sur des méthodologies basées sur de l'apprentissage profond mais aussi sur de l'optimisation combinatoire. Elle est également reconnue pour son expertise concernant la construction ou l'enrichissement de ressources langagières et de graphes de connaissances à partir de corpus ou de données structurées. Dans ces domaines, les conférences internationales sont le moyen de publication privilégié par la communauté particulièrement en traitement automatique des langues, ce qui explique le nombre relativement plus important de publications dans ces supports par rapport aux revues internationales.

Les membres de l'équipe ont ainsi publié dans les conférences les plus réputées de ces domaines : EMNLP, NAACL, AACL, Coling, EACL, CoNLL et ISWC pour les publications réputées de rang A. Les publications dans les revues sont aussi de très bon niveau notamment les revues classées Q1 dans scimagojr : Semantic Web, Knowledge Based Systems, Journal of Geographical Information Science, Future Generation Computer Systems, et celles classées Q1 ou Q2 selon les années : Open Linguistics, Data intelligence, Big Data Research.

Signalons l'article SemEval 2018 qui porte sur l'organisation par l'équipe en collaboration avec des chercheurs allemands du premier challenge sur l'extraction et la classification de relations sémantiques dans le domaine de la littérature scientifique. Cet article n'est pas classé dans les conférences internationales mais a une visibilité importante notamment en raison des jeux de données textuelles que nous avons mis à disposition : il a fait l'objet à ce jour de 173 citations (source Google Scholar). Ce travail est décrit dans le portfolio.

3.2.2 Référence 2. La production scientifique de l'équipe est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels.

Le nombre d'équivalents temps plein sur la période pour les membres permanents et tenant compte de la présence partielle de certains membres est $35,75$. On obtient :

— $21/35,75 = 0,58$ article de revue internationale par personne et par an ;

— $32/35,75 = 0.90$ article de conférence internationale par personne et par an.

Si un article est compté le nombre de fois égal au nombre de membres de l'équipe ou du laboratoire l'ayant co-signé, on obtient les nombres suivants :

— $0,62$ article de revue internationale par personne et par an, soit sensiblement le même nombre que dans le cas précédent (1 seul article avec 2 signatures de membres de l'équipe) ;

— $1,31$ article de conférence internationale par personne et par an. 11 articles de conférences ont 2 membres parmi les signataires, et 2 articles en ont 3.

À la différences des articles de conférence pour lesquels il est donc fréquent que plusieurs membres publient ensemble, les revues sont plus fréquemment co-signées par un seul membre de l'équipe mais avec des collaborateurs étrangers. S'agissant des conférences internationales sélectives, il est aussi intéressant de noter la bonne dynamique de publication qui a été enclenchée durant la période : le nombre de publications dans cette catégorie est passé de 3 en 2018 et 2019, à 9 en 2021, puis à 10 en 2022.

Les doctorants participent également à la production scientifique de l'équipe et ont tous publié durant la préparation de leur thèse au moins 2 articles dont au moins un au niveau international (6 doctorants ont soutenu en 2017 ou 2018 et ont des publications antérieures qui n'apparaissent pas dans ce document). 1 article de revue internationale 11 articles de conférence internationale ont été signés par les doctorants, le plus souvent avec leur(s) encadrant(s), dont certains dans de très bonnes conférences (NAACL, AACL, IWPT, EMNLP) sans compter les publications des 6 doctorants ayant publié antérieurement à 2017). Les postdoctorant(e)s accueillis plus de 6 mois ont également publié durant leur séjour dans l'équipe (EMNLP, CiCling).

Un petit nombre de membres permanents (environ 16% sur la période) ne participent que peu ou pas à la production scientifique de l'équipe. Il s'agit pour l'essentiel d'un investissement dans des tâches lourdes d'enseignement et de responsabilités liées à l'enseignement à l'IUT. S'ils le souhaitent, l'équipe est toujours disponible pour les aider à reprendre des activités de recherche comme par exemple les intégrer dans des projets de recherche.

3.2.3 Publications

Les références
- comportant « * » sont acceptées et à paraître.
- indiquées sur fond gris sont les productions phare de chaque catégorie.
- en gras constituent les 20% les meilleures de la production scientifique de chaque catégorie.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2022

[RI-21] E. Cartier, A. Onysko, E. Winter-Froemel, E. Zenner, G. Andersen, B. Hilberink-Schulpen, U. Nederstigt, E. Peterson et F. van Meurs. Linguistic repercussions of COVID-19 : A corpus study on four languages. *Open Linguistics*, 8(1) :751–766, 2022.

[RI-20] D. Dessi, F. Osborne, D. Reforgiato Recupero, D. Buscaldi et E. Motta. SCICERO : A deep learning and NLP approach for generating scientific knowledge graphs in the computer science domain. *Knowledge-Based Systems*, 258 :109945 (14 pages), 2022.

[RI-19] T. Gaillat, A. Simpkin, N. Ballier, B. Stearns, A. Sousa, M. Bouyé et M. Zarrouk. Predicting CEFR levels in learners of English : The use of microsystem criterial features in a machine learning approach. *ReCALL*, 34(2) :130–146, 2022.

2021

[**RI-18**] D. Dessi, F. Osborne, D. Reforgiato Recupero, D. Buscaldi et E. Motta. Generating knowledge graphs by employing Natural Language Processing and Machine Learning techniques within the scholarly domain. *Future Generation Computer Systems*, 116 :253–264, 2021.

[**RI-17**] L. Di Rocco, F. Dassereto, M. Bertolotto, D. Buscaldi, B. Catania et G. Guerrini. Sherloc : a knowledge-driven algorithm for geolocating microblog messages at sub-city level. *International Journal of Geographical Information Science*, 35(1) :84–115, 2021.

2020 [RI-16] N. Firoozeh, A. Nazarenko, F. Alizon et B. Daille. Keyword extraction : Issues and methods. *Natural Language Engineering*, 26(3) :259–291, 2020.

[RI-15] A. Grezka. Variabilité et traitement automatique des langues. *Lingvisticae Investigationes*, 43(2) :280–299, 2020.

[RI-14] P. Mogorrón Huerta, A. Grezka et L. Navarro-Brotons. Les variations diatopiques dans les expressions figées. *Lingvisticæ Investigationes*, 43(2) :169–171, 2020.

[RI-13] J. Raad, W. Beek, F. Van Harmelen, J. Wielemaker, N. Pernelle et F. Saïs. Constructing and cleaning identity graphs in the LOD cloud. *Data Intelligence*, 2(3) :323–352, 2020.

[RI-12] D. Symeonidou, V. Armant et N. Pernelle. BECKEY : Understanding, comparing and discovering keys of different semantics in knowledge bases. *Knowledge-Based Systems*, 195 :105708 (26 pages), 2020.

2019 [RI-11] A. Grezka et A. Kijima. L'expression de la perception visuelle : regard franco-japonais. *Journal in English Lexicology (LEXIS)*, 13(13) :19 pages, 2019.

[RI-10] D. Reforgiato Recupero, M. Alam, D. Buscaldi, A. Grezka et F. Tavazoee. Frame-based detection of figurative language in tweets [application notes]. *IEEE Computational Intelligence Magazine*, 14(4) :77–88, 2019.

2018 [RI-9] E. Cartier. Noms propres et innovations lexicales : étude linguistique et statistique à partir de Néoveille. *Cahiers de Lexicologie*, 113(113) :203–224, 2018.

[RI-8] J. Garcia Flores, I. V. Meza Ruiz, E. Colin, C. Gardent, A. Gangemi et L. A. Pineda. Robot experience stories : First person generation of robotic task narratives in $SitLog$[1]. *Journal of Intelligent and Fuzzy Systems*, 34(5) :3291–3300, 2018.

2017 [RI-7] M. Alam, D. Reforgiato Recupero, M. Mongiovì, A. Gangemi et P. Ristoski. Event-based knowledge reconciliation using frame embeddings and frame similarity. *Knowledge Based Systems*, 135 :192–203, 2017.

[RI-6] S. Consoli, V. Presutti, D. Reforgiato Recupero, A. G. Nuzzolese, S. Peroni, M. Mongiovì et A. Gangemi. Producing Linked Data for Smart Cities : The Case of Catania. *Big Data Research*, 7 :1–15, 2017.

[RI-5] A. Gangemi, S. Peroni, D. M. Shotton et F. Vitali. The Publishing Workflow Ontology (PWO). *Semantic Web*, 8(5) :703–718, 2017.

[RI-4] A. Gangemi, V. Presutti, D. Reforgiato Recupero, A. G. Nuzzolese, F. Draicchio et M. Mongiovì. Semantic Web Machine Reading with FRED. *Semantic Web*, 8(6) :873–893, 2017.

[RI-3] A. Hernandez-Castañeda, H. Calvo, A. Gelbukh et J. Garcia Flores. Cross-domain deception detection using support vector networks. *Soft Computing*, 21(3) :585–595, 2017.

[RI-2] A. Kemmar, Y. Lebbah, S. Loudni, P. Boizumault et T. Charnois. Prefix-projection global constraint and top-k approach for sequential pattern mining. *Constraints*, 22(2) :265–306, 2017.

[RI-1] A. G. Nuzzolese, V. Presutti, A. Gangemi, S. Peroni et P. Ciancarini. Aemoo : Linked Data exploration based on Knowledge Patterns. *Semantic Web*, 8(1) :87–112, 2017.

Articles dans des revues nationales ou internationales

2022 [RE-7] J. Altmanova, E. Cartier, J. Luzzy, S. Pinto et S. Piscopo. Innovations lexicales dans le domaine de l'environnement et de la biodiversité : le cas de bio en français et en italien. *Neologica, numéro spécial Néologie et environnement*, 16 :85–110, 2022.

[RE-6] A. Borrego, D. Dessi, I. Hernandez, F. Osborne, D. Reforgiato Recupero, D. Ruiz, D. Buscaldi et E. Motta. Completing Scientific Facts in Knowledge Graphs of Research Concepts. *IEEE Access*, 10 :125867–125880, 2022.

2021 [RE-5] E. Cartier et R. Huyghe. La concurrence affixale en diachronie : le cas des préfixes de haut degré en français. *Linx, numéro spécial : Entre vieillissement et innovation : le changement linguistique*, 82 :1–26, 2021.

[RE-4] C. Recanati. La démarche originale de Jean Nicod et son intérêt en sémantique et en sciences cognitives. *Intellectica*, 75(2) :257–284, 2021.

2019 [RE-3] E. Cartier. Neoveille, plateforme de repérage et de suivi des néologismes en corpus dynamique. *Néologica*, 13 :23–54, 2019.

[RE-2] A. Grezka, M. Niziolek et D. Buscaldi. Description de quelques procédés linguistiques de l'ironie, par le biais des tweets sur les transports en commun en français et en polonais. *Studia Romanica Posnaniensia*, 46(1) :43–64, 2019.

2018 [RE-1] A. Grezka, H. Takada et A. Kijima. A record of the lecture given by Dr. Grezka. *Universality and Individuality in Language*, 9 :71–93, 2018. (en japonais).

Chapitres de livre

2021 [CL-11] E. Cartier. *Discourse traditions and computational linguistics*, E. Winter-Froemel et A. S. Octavio de Toledo y Huerta, éditeurs, *Manual of Discourse Traditions in Romance*, volume 30 de *Manuals of Romance Linguistics*, pages 669–690. De Gruyter, 2021.

2020 [CL-10] E. Cartier. *Etude de la variation et du changement lexical à l'ère du numérique*, G. Tallarico, J. Humbley et C. Jacquet-Pfau, éditeurs, *Nouveaux horizons pour la néologie en français*, pages 199–244. Lambert-Lucas, 2020.

[CL-9] E. Cartier. *L'innovation lexicale en français contemporain : étude linguistique, sociopragmatique et statistique de la presse féminine*, F. Diémoz, G. Dostie, P. Hadermann et F. Lefeuvre, éditeurs, *Le français innovant*, volume 130 de *Sciences pour la communication*, pages 21–56. Peter Lang, 2020.

[CL-8] E. Cartier et J. Lazar. *Anglicismes en français et en tchèque contemporains : le cas des formes en -ing*, R. Mudrochová, E. Cartier et A. Kacprzak, éditeurs, *Les emprunts néologiques et leurs équivalents autochtones, études outillées sur corpus*, volume 2020 de *Philologica*, pages 117–132. Charles University Karolinum Press, 2020.

2019 [CL-7] E. Cartier. *Emprunts en français contemporain : étude linguistique et statistique à partir de la plateforme Néoveille*, A. Kacprzak, R. Mudrochová et J.-F. Sablayrolles, éditeurs, *L'emprunt en question(s)*, pages 145–186. Lambert Lucas, 2019.

2018 [CL-6] D. Legallois, T. Charnois et M. Larjavaara. *Grammar of genres and styles : an overview*, D. Legallois, T. Charnois et M. Larjavaara, éditeurs, *The Grammar of Genres and Styles*, pages 1–14. Numéro 320 Trends in Linguistics. Studies and Monographs [TiLSM]. De Gruyter, 2018.

[CL-5] D. Legallois, T. Charnois et M. Larjavaara. *The balance between quantitative and qualitative literary stylistics : how the method of "motifs" can help*, D. Legallois, T. Charnois et M. Larjavaara, éditeurs, *The Grammar of Genres and Styles*, pages 164–193. Numéro 320 Trends in Linguistics. Studies and Monographs [TiLSM]. De Gruyter, 2018.

[CL-4] C. Roussey, H. Hernandez et H. Zargayouna. *Domain ontologies*, P. A. Fuertes-Olivera, éditeur, *The Routledge Handbook of Lexicography*, Chapitre 14, pages 217–234. Routledge Taylor & Francis group, 2018.

2017 [CL-3] E. Cartier et V. Julie. *Etude de la pénétration des anglicismes de type N ou ADJ(-)Ving à partir d'un corpus contemporain journalistique : les exemples de bashing et shaming en français contemporain*, C. Jacquet-Pfau, A. Napieralski et J.-F. Sablayrolles, éditeurs, *Emprunts néologiques et équivalents autochtones : études interlangues*, pages 11–34. Wydawnictwo Uniwersytetu Łódzkiego, 2017.

[CL-2] B. Daille et A. Nazarenko. *Le tournant des données en traitement automatique des langues*, M. Bouzeghoub et R. Mosseri, éditeurs, *Les big data à découvert*, pages 118–119. CNRS Editions, 2017.

[CL-1] G. Lodi, L. Asprino, A. G. Nuzzolese, V. Presutti, A. Gangemi, D. Reforgiato Recupero, C. Veninata et A. Orsini. *Semantic Web for Cultural Heritage Valorisation*, S. Hai-Jew, éditeur, *Data Analytics in Digital Humanities*, pages 3–37. Springer, 2017.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2022 [CI-36] Y. Dauxais, U. Zaratiana, M. Laneuville, S. D. Hernandez-Perez, P. Holat et C. Grosman. Towards Automation of Topic Taxonomy Construction. *The 20th International Symposium on Intelligent Data Analysis (IDA 2022), Rennes, France, April 20-22, 2022*, volume 13205 de *Lecture Notes in Computer Science*, pages 26–38, 2022.

[CI-35] D. Dessì, F. Osborne, D. Reforgiato Recupero, D. Buscaldi et E. Motta. CS-KG : A Large-Scale Knowledge Graph of Research Entities and Claims in Computer Science. *The 21st International Semantic Web Conference (ISWC 2022), Hanghzou, China, October 23-27, 2022*, volume 13489 de *Lecture Notes in Computer Science*, pages 678–696, 2022.

[**CI-34**] G. Felhi, J. Le Roux et D. Seddah. Exploiting Inductive Bias in Transformers for Unsupervised Disentanglement of Syntax and Semantics with VAEs. *The 2022 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies (NAACL 2022), Seattle, United States, July 10-15, 2022*, pages 5763–5776, 2022.

[CI-33] J. L. Garcia Mendoza, D. Buscaldi et F. Orihuela-Espina. Evaluation of a New Representation for Noise Reduction in Distant Supervision. *The 21st Mexican International Conference on Artificial Intelligence (MICAI 2022), Monterrey, Mexico, October 24-29, 2022*, volume 13613 de *Lecture Notes in Artificial Intelligence*, pages 101–113, 2022.

[CI-32] J.-A. González, D. Buscaldi, E. Sanchis et L.-F. Hurtado. Transformer-Based Models for the Automatic Indexing of Scientific Documents in French. *The 27th International Conference on Applications of Natural Language to Information Systems (NLDB 2022), Valencia, Spain, June 15-17, 2022*, volume 13286 de *Lecture Notes in Computer Science*, pages 60–72. Springer, 2022.

[CI-31] V. Liyanage, D. Buscaldi et A. Nazarenko. A Benchmark Corpus for the Detection of Automatically Generated Text in Academic Publications. *The 13th Language Resources and Evaluation Conference,(LREC 2022), Marseille, France, June 20-25, 2022*, pages 4692–4700, 2022.

[CI-30] L. Simonne, N. Pernelle et F. Saïs. Counter Effect Rules Mining in Knowledge Graphs. *The 23rd International Conference on Knowledge Engineering and Knowledge Management (EKAW 2022), Bolzano, Italy, September 26-29, 2022*, volume 13514 de *Lecture Notes in Computer Science*, pages 167–173, 2022.

[CI-29] L. Simonne, N. Pernelle, F. Sais et R. Thamopoulos. Discovering Causal Rules in Knowledge Graphs using Graph Embeddings. *IEEE-ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT),*, page 6, Niagara Falls, Canada, 2022.

[CI-28] U. Zaratiana, N. Tomeh, P. Holat et T. Charnois. GNNer : Reducing Overlapping in Span-based NER Using Graph Neural Networks. *The 60th Annual Meeting of the Association for Computational Linguistics : Student Research Workshop (ACL 2022), Dublin, Ireland, May 22-27, 2022*, pages 97–103, 2022.

[**CI-27**] X. Zhang, J. Le Roux et T. Charnois. Higher-Order Dependency Parsing for Arc-Polynomial Score Functions via Gradient-Based Methods and Genetic Algorithm. *The 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (AACL/IJCNLP 2022), Online Only, November 20-23, 2022*, pages 1158–1171, 2022.

2021

[CI-26] L. Azevedo, M. d'Aquin, B. Davis et M. Zarrouk. LUX (Linguistic aspects Under eXamination) : Discourse Analysis for Automatic Fake News Classification. *Findings of the Association for Computational Linguistics (ACL/IJCNLP 2021), Bangkok, Thailand, August 1-6, 2021*, pages 41–56, 2021.

[CI-25] J. Lopez Espejel, G. De Chalendar, J. Garcia Flores, T. Charnois et I. V. Meza Ruiz. GeSERA : General-domain Summary Evaluation by Relevance Analysis. *Recent Advances in Natural Language Processing 2021 (RANLP 2021), Varna, Bulgaria, Sep 1–3, 2021*, pages 856–867, 2021.

[CI-24] T. Mecharnia, N. Pernelle, C. Rouveirol, F. Hamdi et L. Chibout. Mining Contextual Rules to Predict Asbestos in Buildings. *The 26th International Conference on Conceptual Structures (ICCS 2021), Berlin, Germany, September 20-22, 2021*, volume 12879 de *Lecture Notes in Artificial Intelligence*, pages 170–184, 2021.

[CI-23] A. Nazarenko, F. Lévy et A. Wyner. A Pragmatic Approach to Semantic Annotation for Search of Legal Texts - An Experiment on GDPR. *The 34th Annual Conference on Legal Knowledge and Information Systems (JURIX 2021), Vilnius, Lithuania, 8-10 December 2021*, volume 346 de *Frontiers in Artificial Intelligence and Applications*, pages 23–32, 2021.

[CI-22] G. Santini, H. Soldano et S. Zevio. Experiments on F-Restricted Bi-pattern Mining. *The 10th International Conference on Complex Networks and Their Applications (COMPLEX NETWORKS X), Madrid, Spain, November 30-December 2, 2021*, volume 1015 de *Studies in Computational Intelligence*, pages 539–550, 2021.

[CI-21] G. Santini, H. Soldano et S. Zevio. Restricted Bi-pattern Mining of Attributed Networks. *The 26th International Conference on Conceptual Structures (ICCS 2021), Virtual Event, September 20-22, 2021*, volume 12879 de *Lecture Notes in Artificial Intelligence*, pages 200–207, 2021.

[**CI-20**] L. Simonne, N. Pernelle, F. Saïs et R. Thomopoulos. Differential Causal Rules Mining in Knowledge Graphs. *The 11th on Knowledge Capture Conference (K-CAP '21), Virtual Event, USA, December 2-3, 2021*, pages 105–112, 2021.

[CI-19] H. Vornhagen, M. Zarrouk, B. Davis et K. Young. Do city dashboards make sense ? Conceptualising user experiences and challenges in using city dashboards. A case study. *The 22nd Annual International Conference on Digital Government Research (DG.O'21), Omaha, NE, USA, June 9-11, 2021*, pages 219–226, 2021.

[CI-18] X. Zhang, J. Le Roux et T. Charnois. Strength in numbers : averaging and clustering effects in mixture of experts for graph-based dependency parsing. *The 17th International Conference on Parsing Technologies (IWPT 2021), Bangkok, Thailand, August 6, 2021*, pages 106–118, 2021.

2020

[CI-17] E. Cartier. Borrowings Characteristics and Life-Cycles : a Corpus-Driven Study in Contemporary French . *The Fifth International Conference on Language Contact in Times of Globalization (LCTG 5), Klagenfurt, Austria, September 9–10, 2021*, pages 31–32, 2020.

[**CI-16**] D. Dessì, F. Osborne, D. Reforgiato Recupero, D. Buscaldi, E. Motta et H. Sack. AI-KG : An Automatically Generated Knowledge Graph of Artificial Intelligence. *The 19th International Semantic Web Conference (ISWC 2020), Athens, Greece, November 2-6, 2020*, volume 12507 de *Lecture Notes in Computer Science*, pages 127–143, 2020.

[**CI-15**] Y. Kankanampati, J. Le Roux, N. Tomeh, D. Taji et N. Habash. Multitask Easy-First Dependency Parsing : Exploiting Complementarities of Different Dependency Representations. *The 28th International Conference on Computational Linguistics (COLING 2020), Barcelona, Spain, December 8-13, 2020*, pages 2497–2508, 2020.

[**CI-14**] A. Khajeh, N. Pernelle, F. Saïs et G. Quercini. Generating Referring Expressions from RDF Knowledge Graphs for Data Linking. *The 19th International Semantic Web Conference (ISWC 2020), Athens, Greece, November 2-6, 2020*, volume 12506 de *Lecture Notes in Computer Science*, pages 311–329, 2020.

2019

[CI-13] L. Di Rocco, D. Buscaldi, M. Bertolotto, B. Catania et G. Guerrini. The role of geographic knowledge in sub-city level geolocation. *The 34th ACM/SIGAPP Symposium on Applied Computing (SAC 2019), Limassol, Cyprus, April 8-12, 2019*, pages 687–689, 2019.

[CI-12] I. Garrido-Marquez, A. Nazarenko et F. Lévy. Combining Free Collaborative and Controlled Annotation in Social Media. *The 6th International Conference on Social Networks Analysis, Management and Security (SNAMS 2019), Granada, Spain, October 22-25, 2019*, pages 572–577, 2019.

[**CI-11**] J. Le Roux, A. Rozenknop et M. Lacroix. Representation Learning and Dynamic Programming for Arc-Hybrid Parsing. *The 23rd Conference on Computational Natural Language Learning (CoNLL 2019), Hong Kong, China, November 3-4, 2019*, pages 238–248, 2019.

2018

[CI-10] P. Alizadeh, P. Cellier, T. Charnois, B. Crémilleux et A. Zimmermann. An Experimental Approach For Information Extraction in Multi-Party Dialogue Discourse. *The19th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing 2018), Hanoi, Vietnam, March 18-24, 2018, Revised Selected Papers*, volume 13396 de *Lecture Notes in Computer Science*, pages 180–192, 2018. http ://www.cicling.org/2018/intranet/pre-print/papers/paper_173.pdf.

[CI-9] I. Garrido Marquez, J. Garcia Flores, F. Lévy et A. Nazarenko. Controlling the drift of semantic indexing systems. *The 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KEOD 2018), Seville, Spain, September 18-20, 2018*, pages 199–206, 2018.

[CI-8] A. Nazarenko, F. Lévy et A. Z. Wyner. An Annotation Language for Semantic Search of Legal Sources. *The 11th International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan, May 7-12, 2018*, pages 1096–1100, 2018.

2017 [CI-7] H. Calvo, A. Hernandez-Castaneda et J. Garcia Flores. Author Identification Using Latent Dirichlet Allocation. *The 18th Conference on Computational Linguistics and Intelligent Text Processing (CICLing 2017), Budapest, Hungary, April 17-23, 2017, Revised Selected Papers*, volume 10762 de *Lecture Notes in Computer Science*, page 12, 2017.

[CI-6] E. Cartier. Néoveille, a Web Platform for Neologism Tracking. *The 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2017), Valencia, Spain, April 3-7, 2017, Software DemonstrationsT*, pages 95–98, 2017.

[**CI-5**] C. Corro, J. Le Roux et M. Lacroix. Efficient Discontinuous Phrase-Structure Parsing via the Generalized Maximum Spanning Arborescence. *The 2017 Conference on Empirical Methods in Natural Language Processing (EMLNP 2017), Copenhagen, Denmark, September 7–11, 2017*, pages 1644 – 1654, 2017.

[**CI-4**] K. Gábor, H. Zargayouna, I. Tellier, D. Buscaldi et T. Charnois. Exploring Vector Spaces for Semantic Relations. *The 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP 2017), Copenhagen, Denmark, September 9-11, 2017*, pages 1814–1823, 2017.

[CI-3] A. Meloni, D. Reforgiato Recupero et A. Gangemi. AMR2FRED, A Tool for Translating Abstract Meaning Representation to Motif-Based Linguistic Knowledge Graphs. *The 14th Extended Semantic Web Conference (ESWC 2017) Satellite Events, Portoro\vz, Slovenia, May 28 - June 1, 2017, Revised Selected Papers*, volume 10577 de *Lecture Notes in Computer Science*, pages 43–47, 2017.

[CI-2] J. Reynaud, M. Alam, Y. Toussaint et A. Napoli. A Proposal for Classifying the Content of the Web of Data Based on FCA and Pattern Structures. *23rd International Symposium on Methodologies for Intelligent Systems (ISMIS 2017), Warsaw, Poland, June 26-29, 2017*, volume 10352 de *Lecture Notes in Artificial Intelligence*, pages 684–694, 2017.

[CI-1] A. Wyner, F. Gough, F. Lévy, M. Lynch et A. Nazarenko. On annotation of the textual contents of scottish legal instruments. *The 13th Annual Conference Legal Knowledge and Information Systems (Jurix 2017), Luxembourg, December 13-15, 2017*, volume 302 de *Frontiers in Artificial Intelligence and Applications*, pages 101–106, 2017.

Autres articles dans des conférences ou workshops à comité de lecture

2023 [CO-64] E. Cartier. Semantic Change Tracking Through the Prism of Distributionnalism and Construction Grammars : an Experiment in Contemporary French. *The 14th International Cognitive Linguistics Conference : Linguistic Diversity and Cognitive Linguistics (ICLC-14), 10—14 July 2017 Tartu, Estonia*, pages 214–215, 2023.

[CO-63] A. Khajeh, N. Pernelle et F. Saïs. Enrichissement de règles de Horn par des predicats numériques. *Extraction et Gestion des Connaissances (EGC 2023), Lyon, France, 16 - 20 janvier 2023*, volume E-39 de *Revue des Nouvelles Technologies de l'Information*, pages 425–432, 2023.

2022 [CO-62] N.-M. Boukhatem, D. Buscaldi et L. Liberti. Empirical Comparison of Semantic Similarity Measures for Technical Question Answering. *The 26th European Conference on Advances in Databases and Information Systems : New Trends in Database and Information Systems (ADBIS 2022) - Shorts Papers, Turin, Italy, September 5-8, 2022*, volume 1652 de *Communications in Computer and Information Science*, pages 167–177, 2022. (Short paper).

[CO-61] E. Cartier. Diachronic Semantic Evolution Automatic Tracking : a pilot study in Modern and Contemporary French combining Dependency Analysis and Contextual Embeddings. *The 4th Globalex Workshop on Lexicography and Neology (GWLN4 @EURALEX22), Mannheim, Germany, July 15, 2022*, 2022. https ://videolectures.net/euralex2022_cartier_diachronic_semantic/.

[CO-60] N. Elkhbir, N. Tomeh et T. Charnois. ArabIE : Joint Entity, Relation and Event Extraction for Arabic. *The 7th Arabic Natural Language Processing Workshop (WANLP@EMNLP 2022), Abu Dhabi, United Arab Emirates, December 8, 2022*, pages 331–345, 2022.

[CO-59] G. Felhi, J. Le Roux et D. Seddah. Towards Unsupervised Content Disentanglement in Sentence Representations via Syntactic Roles. *The 1st Controllable Generative Modeling in Language and Vision Workshop (CtrlGen Workshop) at NeurIPS 2021, Virtual Conference, December 13, 2021*, virtual, France, 2022.

[**CO-58**] M. Kamal Eddine, N. Tomeh, N. Habash, J. Le Roux et M. Vazirgiannis. AraBART : a Pretrained Arabic Sequence-to-Sequence Model for Abstractive Summarization. *The Seventh Arabic Natural Language Processing Workshop (WANLP@EMNLP 2022), Abu Dhabi, United Arab Emirates, December 8, 2022*, pages 31–42, 2022.

[CO-57] A. Khajeh, N. Pernelle, F. Saïs et G. Quercini. Knowledge Graph Refinement based on Triplet BERT-Networks. *Deep Learning meets Ontologies and Natural Language Processing (DeepOntoNLP), 3rd International Workshop, in conjunction with ESWC 2022, Hersonissos, Greece, May 29-June 2, 2022*, pages 1–9, The 3rd International Workshop on Deep Learning meets Ontologies, 2022.

[CO-56] L. Simonne, N. Pernelle et F. Saïs. Découverte de règles causales dans les graphes de connaissances à l'aide de plongements dans les graphes. *33es Journées francophones d'Ingénierie des Connaissances (IC 2022), Saint-Étienne, France, 29 juin–1 juillet 2022*, pages 1–10, 2022.

[CO-55] U. Zaratiana, N. Elkhbir, P. Holat, N. Tomeh et T. Charnois. Global Span Selection for Named Entity Recognition. *The 2022 Workshop on Unimodal and Multimodal Induction of Linguistic Structures (UM-IoS@EMNLP 2022), Abu Dhabi, United Arab Emirates, December 7, 2022*, pages 11–17, 2022.

[**CO-54**] U. Zaratiana, N. Elkhbir, D. Núñez, P. Holat, N. Tomeh et T. Charnois. DyREx : Dynamic Query Representation for Extractive Question Answering. *The 2nd Workshop on Efficient Natural Language and Speech Processing (ENSLP@NeurIPS 2022), New Orleans, LO, United States, December 2, 2022*, pages 1–6. 2nd Workshop on Efficient Natural Language and Speech Processing @ Neurips 2022, 2022.

[CO-53] U. Zaratiana, N. Tomeh, P. Holat et T. Charnois. Named Entity Recognition as Structured Span Prediction. *The First Workshop on Unimodal and Multimodal Induction of Linguistic Structures (UM-IoS@EMNLP 2022), Abu Dhabi, United Arab Emirates, December 7-8, 2022*, pages 1–10, 2022.

2021

[CO-52] G. Felhi, J. Le Roux et D. Seddah. Challenging the Semi-Supervised VAE Framework for Text Classification. *The Second Workshop on Insights from Negative Results in NLP, Punta Cana, Dominican Republic, November 10, 2021*, pages 136–143, 2021.

[CO-51] F. Lovera, Y. Cardinale, D. Buscaldi, T. Charnois et N. H. Masun. Deep Learning Enhanced with Graph Knowledge for Sentiment Analysis. *6th International Workshop on Explainable Sentiment Mining and Emotion Detection (X-SENTIMENT 2021) co-located with 18th Extended Semantic Web Conference 2021 (ESWC 2021), Hersonissos, Greece, June 6 –7, 2021*, volume 2918 de *CEUR Workshop Proceedings*, pages 74–86, 2021.

[CO-50] T. Mecharnia, L. Chibout Khelifa, F. Hamdi, N. Pernelle et C. Rouveirol. Découverte de règles contextuelles pour prédire la présence d'amiante dans les bâtiments. *32èmes Journées francophones d'Ingénierie des Connaissances (IC 2021), Bordeaux, France, June 30 - July 2, 2021*, pages 81–88, 2021.

[CO-49] L. Simonne, N. Pernelle et F. Saïs. Fouille de règles différentielles causales dans les graphes de connaissances. *Extraction et Gestion des Connaissances (EGC 2021), 25-29 Janvier 2021, Montpellier, France*, volume E-37 de *Revue des Nouvelles Technologies de l'Information*, pages 293–300, 2021.

2020

[CO-48] P. Alizadeh, I. Vladimir et J. Garcia Flores. Apprentissage par renforcement pour la recherche d'experts sur le web. *Extraction et Gestion des Connaissances (EGC 2020), Brussels, Belgium, January 27-31, 2020*, volume E-36 de *Revue des Nouvelles Technologies de l'Information*, pages 261–268. Éditions RNTI, 2020.

[CO-47] D. Buscaldi, G. Felhi, D. Ghoul, J. Le Roux, G. Lejeune et X. Zhang. Calcul de similarité entre phrases : quelles mesures et quels descripteurs ? *6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition), Atelier DÉfi Fouille de Textes, Nancy, France, June 8-19, 2020*, pages 14–25, 2020.

[CO-46] T. Gaillat, N. Ballier, A. Sousa, M. Bouyé, A. Simpkin, B. Stearns et M. Zarrouk. Un prototype en ligne pour la prédiction du niveau de compétence en anglais des productions écrites. *6ème conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition), Nancy, France, June 8-19, 2020*, volume 4, pages 30–33, 2020.

[CO-45] B. Ghanem, D. Buscaldi et P. Rosso. TexTrolls : Identifying Trolls on Twitter with Textual and Affective Features. *The 2020 Workshop on Online Misinformation and Harm-Aware Recommender Systems (OHARS'20) co-located with 14th ACM Conference on Recommender Systems (RecSys 2020), Rio de Janeiro, Brazil, September 25, 2020*, volume 2758 de *CEUR Workshop Proceedings*, pages 4–22, 2020.

[CO-44] A. Grezka. Verbes de perception et traitement de la polysémie : pourquoi et comment ? *Actes du colloque international "La perception : langue, discours, cognition", CELIS0, Paris, France, 6-7 décembre 2019*, pages 29–44. Irina Thomières-Shakovskaya, 2020.

[CO-43] A. Khajeh, N. Pernelle et F. Saïs. Découverte d'expressions référentielles dans les graphes de connaissances. *Extraction et Gestion des Connaissances (EGC 2020), Brussels, Belgium, January 27-31, 2020*, volume E-36 de *Revue des Nouvelles Technologies de l'Information*, pages 357–364, 2020.

[CO-42] A. Khajeh, N. Pernelle, F. Saïs et G. Quercini. RE-miner for data linking results for OAEI 2020. *The 15th International Workshop on Ontology Matching co-located with the 19th International Semantic Web Conference (OM@ISWC 2020), Athens, Greece, November 2, 2020*, volume 2788 de *CEUR Workshop Proceedings*, pages 211–215, 2020.

[CO-41] A. Sousa, N. Ballier, T. Gaillat, B. Stearns, M. Zarrouk, A. Simpkin et M. Bouyé. From Linguistic Research Projects to Language Technology Platforms : A Case Study in Learner Data. *Proceedings of the 1st International Workshop on Language Technology Platforms (IWLTP 2020), Marseille, France, May 11-12, 2020*, Proceedings of the 1st International Workshop on Language Technology Platforms (IWLTP 2020), pages 112–120, 2020.

[CO-40] A. Sousa, M. d'Aquin, M. Zarrouk et J. Holloway. Person-Independent Multimodal Emotion Detection for Children with High-Functioning Autism. *The First Workshop on Artificial Intelligence for Function, Disability, and Health co-located with the 2020 International Joint Conference on Artificial Intelligence - Pacific Rim Conference on Artificial Intelligence (IJCAI-PRICAI 2020), Online, January 7-8, 2021*, volume 2760 de *CEUR Workshop Proceedings*, pages 14–20, 2020.

[CO-39] S. Zevio, G. Santini, H. Soldano, H. Zargayouna et T. Charnois. A Combination of Semantic Annotation and Graph Mining for Expert Finding in Scholarly Data. *The 2020 Workshop on Graph Embedding and Mining Worshop (GEM 2020) co-located with European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2020), Virtual Event, September 18, 2020*, page 16, 2020.

2019 [CO-38] D. Buscaldi, D. Dessì, E. Motta, F. Osborne et D. Reforgiato Recupero. Mining Scholarly Data for Fine-Grained Knowledge Graph Construction. *The Workshop on Deep Learning for Knowledge Graphs (DL4KG2019) Co-located with the 16th Extended Semantic Web Conference 2019 (ESWC 2019), Portoroz, Slovenia, June 2, 2019*, volume 2377 de *CEUR Workshop Proceedings*, pages 21–30, 2019.

[CO-37] D. Buscaldi, D. Dessì, E. Motta, F. Osborne et D. Reforgiato Recupero. Mining Scholarly Publications for Scientific Knowledge Graph Construction. *The 16th Extended Semantic Web Conference - Satellite Events (ESWC 2019 Satellite Events), Portorož, Slovenia, June 2-6, 2019, Revised Selected Papers*, volume 11762 de *Lecture Notes in Computer Science*, pages 8–12, 2019.

[CO-36] D. Buscaldi, D. Ghoul, J. Le Roux et G. Lejeune. Indexation et appariements de documents cliniques pour le Deft 2019. *Edition 2019 du Défi Fouille de Textes (DEFT 2019), Toulouse, France, 9-15 Mai 2019*, pages 49–56, 2019.

[CO-35] A. Grezka, M. Niziolek et D. Buscaldi. Description linguistique et détection automatique de l'ironie dans les tweets en français et en polonais. *5ème Forum franco-polonais de la recherche et de l'innovation, Varsovie, Pologne, 19 Novembre 2019*. 5ème forum franco-polonais de la recherche et de l'innovation, 2019. (poster).

[CO-34] H. Murrieta, I. Meza, P. Alizadeh et J. Garcia Flores. Towards Identifying for Evidence of Drain Brain from Web Search Results using Reinforcement Learning. *LatinX in AI*

Research Workshop (LXAI) at the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, December 8-14, 2019, pages 1–3, 2019. (poster).

[CO-33] S. Zevio, G. Santini, H. Zargayouna et T. Charnois. Fouille de texte et fouille de graphe appliquées à la recherche d'experts. *30èmes journées francophones d'Ingénierie des Connaissances (IC 2019), Toulouse, France, 2-4 juillet 2019*, pages 222–223, 2019.

2018

[CO-32] P. Alizadeh, P. Cellier, T. Charnois, B. Crémilleux et A. Zimmermann. Etude expérimentale d'extraction d'information dans des retranscriptions de réunions. *25e conférence sur le Traitement Automatique des Langues Naturelles (TALN 2018), Rennes, France, May 14-18, 2018*, pages 465–472, rennes, france, 2018.

[CO-31] D. Buscaldi. Improving access to scientific literature : a semantic IR perspective. *The 5th Spanish Conference on Information Retrieval (CERI 2018), Zaragoza, Spain, June 26-27, 2018*, page 1 :1, 2018.

[CO-30] D. Buscaldi. Tweetaneuse @ AMI EVALITA2018 : Character-based Models for the Automatic Misogyny Identification Task (Short Paper). *The Sixth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2018) co-located with the Fifth Italian Conference on Computational Linguistics (CLiC-it 2018), Turin, Italy, December 12-13, 2018*, volume 2263 de *CEUR Workshop Proceedings*, 2018.

[CO-29] D. Buscaldi, J. Le Roux et G. Lejeune. Modèles en Caractères pour la Détection de Polarité dans les Tweets. *13ème Edition de l'Atelier Défi Fouille de Textes (DEFT@CORIA/TALN 2018), Rennes, France, 15 mai 2018*, pages 249–258, 2018.

[CO-28] E. Cartier. Néologie sémantique : modélisation, expérimentations automatiques multilingues dans le cadre de Néoveille. *4ème Congrès international de néologie des langues romanes (CINEO 2018), Lyon, France, July 4-6, 2018*, pages 1–2, 2018.

[CO-27] E. Cartier. Néonaute : mining web archives for linguistic analysis. *IIPC Web archiving Conference 2018, Wellington (New-Zealand), November 12-15, 2018.*, 2018. sans proceedings.

[CO-26] E. Cartier. Néoveille, an Automatic System for Lexical Units Life-Cycle Tracking. *The XVIII EURALEX International Congress, Ljubljana, Slovenia, July 17-21, 2018*, pages 34–35, XVIII EURALEX International Congress, Ljubljana, Slovenia, 2018. sans proceedings.

[CO-25] E. Cartier, L. Galand, P. Stirling et S. Aubry. Néonaute, Enrichissement sémantique pour la recherche d'information. *25ème conférence sur le Traitement Automatique des Langues Naturelles (TALN 2018), Rennes, France, May 14-18, 2018*, pages 201–204, 2018.

[CO-24] E. Cartier, J.-F. Sablayrolles, N. Boutmgharine, J. Humbley, M. Bertocci, C. Jacquet-Pfau, N. Kubler et G. Tallarico. Détection automatique, description linguistique et suivi des néologismes en corpus : point d'étape sur les tendances du français contemporain. *6ème Congrès Mondial de Linguistique Française, Mons, Belgique, 9-13 juillet 2018*, volume 46 de *SHS Web of Conferences*, pages 08002 : 1–20, 2018.

[**CO-23**] K. Gábor, D. Buscaldi, A.-K. Schumann, B. Q. Zadeh, H. Zargayouna et T. Charnois. SemEval-2018 Task 7 : Semantic Relation Extraction and Classification in Scientific Papers. *The 12th International Workshop on Semantic Evaluation (SemEval@NAACL-HLT 2018), New Orleans, Louisiana, USA, June 5-6, 2018*, pages 679–688, 2018.

[CO-22] K. Gábor, N. Lechevrel, I. Tellier, D. Buscaldi, H. Zargayouna et T. Charnois. Apport des dépendances syntaxiques et des patrons séquentiels à l'extraction de relations. *25e conférence sur le Traitement Automatique des Langues Naturelles (TALN 2018) , Rennes, France, May 14-18, 2018*, pages 557–566, 2018.

[CO-21] S. Zevio. Identifying Transmedia Works from User-Generated Knowledge Bases : Japanese Pop Culture Study Case. *The 8th Joint International Semantic Technology Conference co-located with the 8th Joint International Semantic Technology Conference (JIST 2018), Awaji City, Hyogo, Japan, November 26-28, 2018 (Workshops and Posters)*, volume 2293 de *CEUR Workshop Proceedings*, pages 33–42, 2018.

[CO-20] S. Zevio. Knowledge Discovery and Enrichment from Scholarly Data for Expert Finding. *The EKAW Doctoral Consortium 2018 co-located with the 21st International Conference on Knowledge Engineering and Knowledge Management Doctoral Consortium (EKAW 2018), , Nancy, France, November 13, 2018*, volume 2306 de *CEUR Workshop Proceedings*, pages 1–8, 2018.

[CO-19] S. Zevio, H. Zargayouna, G. Santini et T. Charnois. Vers une cartographie automatique des thématiques et profils d'experts associés à une conférence scientifique : 9 ans d'ateliers recherche d'information semantique. *10ème Atelier Recherche d'Information Sémantique (RISE@CORIA/TALN 2018), Rennes, France, 15 Mai 2018*, rennes, france, 2018.

2017 [CO-18] M. Alam, D. Reforgiato Recupero, M. Mongiovì, A. Gangemi et P. Ristoski. Frame Embeddings for Event-Based Knowledge Reconciliation. *The ISWC 2017 Posters & Demonstrations and Industry Tracks co-located with 16th International Semantic Web Conference (ISWC 2017), Vienna, Austria, October 23-25, 2017*, volume 1963 de *CEUR Workshop Proceedings*, 2017.

[CO-17] M. Alam, D. Reforgiato Recupero, M. Mongiovì, A. Gangemi et P. Ristoski. Reconciling Event-Based Knowledge Through RDF2VEC. *The International Workshops on Hybrid Statistical Semantic Understanding and Emerging Semantics, and Semantic Statistics co-located with 16th International Semantic Web Conference (HybridSemStats@ISWC 2017), Vienna, Austria, October 22, 2017*, volume 1923 de *CEUR Workshop Proceedings*, pages Article 3 : 1–10, 2017.

[CO-16] I. Bannour, H. Zargayouna et A. Nazarenko. Modèle de recherche d'information sémantique en graphe : interrogation par propagation d'activation. *28èmes Journées francophones d'Ingénierie des Connaissances (IC 2017), Caen, France, July 3-7, 2017*, pages 181–192, 2017.

[CO-15] D. Buscaldi, A. Grezka et G. Lejeune. Tweetaneuse : Fouille de motifs en caractères et plongement lexical à l'assaut du DEFT 2017. *13ème Défi Fouille de Texte (DEFT 2017), TALN 2017, Orléans, France, 26-30 Juin, 2017*, pages 65–76, 2017.

[CO-14] D. Buscaldi, S. D. Hernandez-Perez et T. Charnois. Classification of keyphrases from scientific publications using WordNet and word embeddings. *1er Atelier Valorisation et Analyse des Données de la Recherche (VADOR 2017) durant la 35e édition du congrès Informatique des Organisations et Systèmes d'Information et de Décision (INFORSID 2017), Toulouse, France, May 31, 2017*, volume 1860 de *CEUR Workshop Proceedings*, pages 50–57, 2017.

[CO-13] D. Buscaldi et B. Priego. LIPN-UAM at EmoInt-2017 : Combination of Lexicon-based features and Sentence-level Vector Representations for Emotion Intensity Determination. *The 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA@EMNLP 2017), Copenhagen, Denmark, September 8, 2017*, pages 255–258, 2017.

[CO-12] E. Cartier. Detection and Life-cycles Tracking of Neologisms : Experiments in the Neoveille Project. *The Dynamics of Lexical Innovation Workshop 2017, Munich, Germany, June 28-30, 2017*, pages 8–8, 2017.

[CO-11] E. Cartier. Sémantique lexicale et distributionnalisme : éléments pour le repérage automatique du sens en corpus. *7ème Colloque international sur la Représentations du Sens Linguistique (RSL VII) Sherbrooke, Québec, Canada, 25-27 octobre 2017*, 2017.

[CO-10] E. Cartier, G. Lejeune, K. Gábor et T. Charnois. Neoveille, a System for Multilingual Online Neologism Tracking. *The 18th International Conference on Computational Linguistics and Intelligent Text Processing (CICLING 2017), Budapest, Hungary, April 17–23, 2017*, pages 13–25, 2017. (Supplementary Proceedings).

[CO-9] C. Corro et J. Le Roux. Transforming Dependency Structures to LTAG Derivation Trees. *The 13th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+13), Umea, Sweden, September 4–6, 2017*, pages 112–121, 2017.

[CO-8] I. Garrido Marquez, A. Nazarenko, F. Lévy et J. Garcia Flores. Mesurer la qualité des systèmes de catégories de blogs. *28èmes Journées francophones d'Ingénierie des Connaissances (IC 2017), Caen, France, July 3-7, 2017*, pages 217–222, 2017.

[CO-7] S. D. Hernandez-Perez, D. Buscaldi et T. Charnois. LIPN at SemEval-2017 Task 10 : Filtering Candidate Keyphrases from Scientific Publications with Part-of-Speech Tag Sequences to Train a Sequence Labeling Model. *The 11th International Workshop on Semantic Evaluation (SemEval@ACL 2017), Vancouver, Canada, August 3-4, 2017*, pages 995–999, 2017.

[CO-6] N. Lechevrel, K. Gábor, I. Tellier, T. Charnois, H. Zargayouna et D. Buscaldi. Combining syntactic and sequential patterns for unsupervised semantic relation extraction. *The*

2017 Workshop on Interactions between Data Mining and Natural Language Processing 2017 (DMNLP 2017) co-located with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2017) Skopje, Macedonia, September 22, 2017, volume 1880 de *CEUR Workshop Proceedings*, pages 81–84, 2017.

[CO-5] G. Lejeune et E. Cartier. Character Based Pattern Mining for Neology Detection. *The First Workshop on Subword and Character Level Models in NLP, Copenhagen, Denmark, September 7, 2017*, pages 25–30, 2017.

[**CO-4**] A. Meroño Peñuela, R. Hoekstra, A. Gangemi, P. Bloem, R. de Valk, B. Stringer, B. Janssen, V. de Boer, A. Allik, S. Schlobach et K. R. Page. The MIDI Linked Data Cloud. *The 16th International Semantic Web Conference (ISWC 2017), Vienna, Austria, October 21-25, 2017*, volume 10588 de *Lecture Notes in Computer Science*, pages 156–164, 2017.

[CO-3] N. Mimouni, A. Nazarenko et S. Salotti. Answering Complex Queries on Legal Networks : A Direct and a Structured IR Approaches. *AI Approaches to the Complexity of Legal Systems - AICOL International Workshops 2015-2017 : AICOL-VI@JURIX 2015, AICOL-VII@EKAW 2016, AICOL-VIII@JURIX 2016, AICOL-IX@ICAIL 2017, and AICOL-X@JURIX 2017, Revised Selected Papers*, volume 10791 de *Lecture Notes in Computer Science*, pages 451–464, 2017.

[CO-2] C. Santos, A. Gangemi et M. Alam. Detecting and Editing Privacy Policy Pitfalls on the Web. *The 1st Workshop on Technologies for Regulatory Compliance co-located with the 30th International Conference on Legal Knowledge and Information Systems (JURIX 2017), Luxembourg, December 13, 2017*, volume 2049 de *CEUR Workshop Proceedings*, pages 87–99, 2017.

[CO-1] H. Zargayouna, C. Roussey et S. Ouardani. Annotation sémantique à partir de textes : Cas des observations dans les Bulletins de Santé du végétal. *9ème atelier Recherche d'Information SEmantique (RISE 2017), Caen, France, 4 Juillet 2017*, pages 1–8, 2017.

Direction d'ouvrages (DO)

Livres

2018 [LI-1] D. Legallois, T. Charnois et M. Larjavaara. *The Grammar of genres and styles. From Discrete to Non-Discrete Units*, volume 320 de *Trends in Lingustics book series*. De Gruyter Mouton, 2018. to appear.

Édition d'ouvrages collectifs

2023 [ED-17] A. Kacprzak, E. Cartier et R. Mudrochová. *La néologie par emprunt en français, en polonais et en tchèque : tendances actuelles*, volume 58 de *Études de linguistique, Littérature et Art*. Peter Lang, 2023.

2022 [ED-16] M. Alam, D. Buscaldi, M. Cochez, F. Osborne, D. Reforgiato Recupero et H. Sack. *Special Issue on Deep Learning and Knowledge Graphs*, volume 13 de *Semantic Web*. AKA, 2022.

[**ED-15**] H. Bouamor, H. S. Al-Khalifa, K. Darwish, O. Rambow, F. Bougares, A. Abdelali, N. Tomeh, S. Khalifa et W. Zaghouani. *The 7th Arabic Natural Language Processing Workshop (WANLP@EMNLP 2022), Abu Dhabi, United Arab Emirates, December 8, 2022*. Association for Computational Linguistics, 2022.

[ED-14] D. Buscaldi, M. Dragoni, F. Frasincar et D. Reforgiato Recupero. *Special Issue on Sentimental Health : Future Generation Sentiment Analysis Systems*, volume 129 de *Future Generation Computer Systems*. Elsevier, 2022.

[ED-13] B. R. Chakravarthi, B. Bharathi, J. P. McCrae, M. Zarrouk, K. Bali et P. Buitelaar. *The 2nd Workshop on Language Technology for Equality, Diversity and Inclusion (LT-EDI 2022), Dublin, Ireland, May 27, 2022*. Association for Computational Linguistics, 2022.

[ED-12] V. Khurshudian, N. Tomeh, D. Nouvel, A. Donabédian et C. Vidal-Gorène. *The LREC 2022 Workshop on Processing Language Variation : Digital Armenian (DigitAm), Marseille, France, June 20, 2022*. European Language Resources Association (ELRA), 2022.

[ED-11] B. Piwowarski, F. Boudin, G. Dias, J.-P. Chevallet, J. Moreno, M. Roche et T. Charnois. *Actes de la 3ème journée : Technologies du langage humain et accès interactif à l'information, Paris, Avril 2022*. 2022.

2021 [ED-10] S. Ben Abbès, R. Hantach, P. Calvez, D. Buscaldi, D. Dessì, M. Dragoni, D. Reforgiato Recupero et H. Sack. *The 2nd International Workshop on Deep Learning meets Ontologies and Natural Language Processing (DeepOntoNLP 2021) & 6th International Workshop on Explainable Sentiment Mining and Emotion Detection (X-SENTIMENT 2021) co-located with 18th Extended Semantic Web Conference (ESWC 2021), Hersonissos, Greece, June 6-7, 2021*, volume 2918 de *CEUR Workshop Proceedings*. 2021.

[ED-9] B. R. Chakravarthi, B. Bharathi, J. P. McCrae, M. Zarrouk, K. Bali et P. Buitelaar. *The First Workshop on Language Technology for Equality, Diversity and Inclusion (LT-EDI@EACL 2021), Online, April 19, 2021*, *The First Workshop on Language Technology for Equality, Diversity and Inclusion (LT-EDI@EACL 2021), Online, April 19, 2021*. Association for Computational Linguistics, 2021.

[ED-8] N. Habash, H. Bouamor, H. M. Hajj, W. Magdy, W. Zaghouani, F. Bougares, N. Tomeh, I. Abu Farha et S. Touileb. *The 6th Arabic Natural Language Processing Workshop, (WANLP 2021), Kyiv, Ukraine, April 9, 2021*. Association for Computational Linguistics, 2021.

2020 [ED-7] M. Alam, D. Buscaldi, M. Cochez, F. Osborne, D. Reforgiato Recupero et H. Sack. *The 2020 Workshop on Deep Learning for Knowledge Graphs (DL4KG2020) co-located with the 17th Extended Semantic Web Conference 2020 (ESWC 2020), Heraklion, Greece, June 2, 2020*, volume 2635 de *CEUR Workshop Proceedings*. 2020.

[ED-6] R. Mudrochová, E. Cartier et A. Kaprczak. *Les emprunts néologiques et leurs équivalents autochtones, études outillées sur corpus*, volume 2020 de *Philologica*. Charles University Karolinum Press, 2020.

[ED-5] I. Zitouni, M. Abdul-Mageed, H. Bouamor, F. Bougares, M. El-Haj, N. Tomeh et W. Zaghouani. *The 5th Arabic Natural Language Processing Workshop (WANLP@COLING 2020), Barcelona, Spain, December 12, 2020*. Association for Computational Linguistics, 2020.

2018 [ED-4] D. Buscaldi, A. Gangemi et D. Reforgiato Recupero. *The 5th Semantic Web Challenges (SemWebEval 2018) at ESWC 2018, Heraklion, Greece, June 3-7, 2018, Revised Selected Papers*, volume 927 de *Communications in Computer and Information Science*. Springer, 2018.

[ED-3] D. Reforgiato Recupero, M. Dragoni, D. Buscaldi, M. Alam et E. Cambria. *The 4th Workshop on Sentic Computing, Sentiment Analysis, Opinion Mining, and Emotion Detection (EMSASW 2018) Co-located with the 15th Extended Semantic Web Conference 2018 (ESWC 2018), Heraklion, Greece, June 4, 2018*, volume 2111 de *CEUR Workshop Proceedings*. 2018.

2017 [ED-2] P. Cellier, T. Charnois, A. Hotho, S. Matwin, M.-F. Moens et Y. Toussaint. *Proceedings of the Workshop on Interactions between Data Mining and Natural Language Processing 2017 (DMNLP 2017) co-located with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2017) Skopje, Macedonia, September 22, 2017*, volume 880 de *CEUR Workshop Proceedings*. 2017.

[ED-1] N. Habash, M. T. Diab, K. Darwish, W. El-Hajj, H. S. Al-Khalifa, H. Bouamor, N. Tomeh et M. El-Haj. *The 3rd Arabic Natural Language Processing Workshop, (WANLP 2017@EACL), Valencia, Spain, April 3, 2017*. Association for Computational Linguistics, 2017.

Logiciels et Ressources (LO)

2022 [LO-27] D. Antonov et D. Buscaldi. Relation-prediction-from-embeddings. https://github.com/chejuro/Relation-prediction-from-embeddings, 2022.

[LO-26] D. Buscaldi, D. Dessì, D. Reforgiato Recupero, F. Osborne et E. Motta. CS-KG : A Large-Scale Knowledge Graph of Research Entities and Claims in Computer Science. http://scholkg.kmi.open.ac.uk/, 2022.

[LO-25] M. Daychman et D. Buscaldi. Prédiction de relations AIKG / Wordnet (sémantique). https://github.com/Pylyr/Internship_SBERT, 2022.

[LO-24] G. Felhi, J. Le Roux et D. Seddah. QKVAE : A model for Unsupervised Disentanglement of Syntax and Semantics (code). https://github.com/ghazi-f/QKVAE, 2022.

[LO-23] J. A. Gonzalez, D. Buscaldi, L. Hurtado et E. Sanchis. MBartCopyGenerator. https://gist.github.com/jogonba2/, 2022.

[LO-22] M. Kamal Eddine, N. Tomeh, J. Le Roux, N. Habash et M. Vazirgiannis. AraBART : a Pretrained Arabic Sequence-to-Sequence Model for Abstractive Summarization (code). https://huggingface.co/moussaKam/AraBART, 2022.

[LO-21] U. Zaratiana, N. Elkhbir, P. Holat, N. Tomeh et T. Charnois. Global Span Selection for Named Entity Recognition (code). https://github.com/urchade/global-span-selection, 2022.

[LO-20] U. Zaratiana, N. Elkhbir, D. H. Nunez Fernandez, P. Holat, N. Tomeh et T. Charnois. DyREx : Dynamic Query Representation for Extractive Question Answering (code). https://github.com/urchade/DyReX, 2022.

[LO-19] U. Zaratiana, P. Holat, N. Tomeh et T. Charnois. Hierarchical Transformer Model for Scientific Named Entity Recognition (code). https://github.com/urchade/HNER, 2022.

[LO-18] U. Zaratiana, N. Tomeh, P. Holat et T. Charnois. GNNer : Reducing Overlapping in Span-based NER Using Graph Neural Networks (code). https://github.com/urchade/gnner, 2022.

[LO-17] U. Zaratiana, N. Tomeh, P. Holat et T. Charnois. Named Entity Recognition as Structured Span Prediction (code). https://github.com/urchade/span-structured-prediction, 2022.

[LO-16] X. Zhang, J. Le Roux et T. Charnois. Basic Parser, linear and second order. https://github.com/kidlestar/BParser, 2022.

[LO-15] X. Zhang, J. Le Roux et T. Charnois. Higher-Order Dependency Parsing for Arc-Polynomial Score Functions via Gradient-Based Methods and Genetic Algorithm (code). https://github.com/kidlestar/PolyParser, 2022.

2021 [LO-14] D. Buscaldi. Sentence Representations based on Dependency Graphs and Graph Convolutional Networks (DG-GCN). https://github.com/dbuscaldi/PreTens22, 2021.

[LO-13] A. Grezka, J. Garcia Flores et T. Charnois. ChêneTAL : plate-forme d'intégration d'outils de traitement automatique des langues. *XXIVe École Doctorale des Pays de Vysegràd, Autour du texte dans les études françaises„ Cracovie, Pologne 23-25 septembre 2021*, 2021.

[LO-12] J. Lopez Espejel, P. Alizadeh et J. Garcia Flores. RENFO / DiasporasRL (code). https://github.com/jorgegarciaflo/DiasporasRL, 2021.

[LO-11] J. Lopez Espejel, G. De Chalendar, J. Garcia Flores, T. Charnois et I. V. Meza Ruiz. GeSERA : General-domain Summary Evaluation by Relevance Analysis (code). https://github.com/JessicaLopezEspejel/GeSERA, 2021.

[LO-10] X. Zhang, J. Le Roux et T. Charnois. Strength in numbers : averaging and clustering effects in mixture of experts for graph-based dependency parsing (code). https://github.com/kidlestar/MOE, 2021.

2020 [LO-9] Y. Kankanampati, J. Le Roux, N. Tomeh, D. Taji et N. Habash. Multitask Easy-First Dependency Parsing : Exploiting Complementarities of Different Dependency Representations (code). https://github.com/yash-reddy/MEF_parser, 2020.

2019 [LO-8] E. Cartier. Neoveille, plateforme de repérage et de suivi des néologismes en corpus dynamique (outil en ligne). https://tal.lipn.univ-paris13.fr/neoveille/, 2019.

[LO-7] J. Garcia Flores, H. Ferreira, A. Okotan, F. Mantilla et F. Abdoulvahide. Pamparios : Traducteur espagnol - Wixárika (langue amérindienne) (démo). https://tal.lipn.univ-paris13.fr/wixarika/, 2019.

[LO-6] A. Grezka. Morfetik, ressource morphologique. https://tal.lipn.univ-paris13.fr/morfetik, 2019.

[LO-5] H. Kaddouh, M. Labiodh, N. Ghourabi, J. Garcia Flores, A. Hafsa et C. Ourtirane. YuCorrectMaya : validation app of comparable corpus for the Maya Yucatec language (code). https://github.com/rcln/maya, 2019.

2018 [LO-4] D. Buscaldi. Tweetaneuse @ AMI EVALITA2018 : Character-based Models for the Automatic Misogyny Identification Task (code). https://github.com/dbuscaldi/tweetaneuse2018, 2018.

[LO-3] I. Garrido Marquez, J. Garcia Flores, F. Lévy et A. Nazarenko. Kilroy : tag prediction and semantic indexing system (code). https://github.com/IvanGarridoMarquez/predictOverTime, 2018.

2017 [LO-2] J. Garcia Flores, I. V. Meza Ruiz et L. A. Pineda Cortes. Golfred : Robot Experience Stories Generator (code). https://github.com/jorgegarciaflo/Golfred, 2017.

[LO-1] I. V. Meza Ruiz, G. Fuentes-Pineda, J. Garcia Flores, M. Chabouni et Z. Khezane. minhashing : Topic mining data processing and visualization (mostly on wikipedia), 2017.

Thèses et Habilitations (TH)

Thèses

2023 [TH-12] G. Felhi. Interpretable Sentence Representation with Variational Autoencoders and Attention. Thèse, Université Sorbonne Paris Nord, 2023.

2022 [TH-11] X. Zhang. Multitask Deep Learning for Joint Syntactic and Semantic Structure Prediction. Thèse, Université Sorbonne Paris Nord, 2022.

2021 [TH-10] J. Lopez Espejel. Résumé automatique abstractif des textes médicaux longs avec un Transformer à multi-encodeurs et évaluation de résumé du domaine général avec wikiSERA. Thèse, Université Sorbonne Paris Nord, 2021.

[TH-9] S. Zevio. Découverte et enrichissement de connaissances à partir de textes pour la recherche d'experts. Thèse, Université Sorbonne Paris Nord, 2021.

2019 [TH-8] I. Garrido Marquez. Dynamics in semantic annotation, a perspective of information access system. Thèse, Université Sorbonne Paris Nord, 2019.

[TH-7] S. D. Hernandez-Perez. Détection de topiques et leur similarité dans les textes scientifiques. Thèse, Université Sorbonne Paris Nord, 2019.

2018 [TH-6] C. Corro. Lagrangian Based Approaches for Lexicalized Tree Adjoining Grammar Parsing. Thèse, Université Sorbonne Paris Cité, 2018.

[TH-5] N. Firoozeh. Semantic-oriented Recommendation for Content Enrichment / Recommandation sémantique pour l'enrichissement de contenus textuels. Thèse, Université Sorbonne Paris Cité, 2018.

[TH-4] P. Holat. Fouille de motifs et modélisation statistique pour l'extraction de connaissances textuelles. Thèse, Université Sorbonne Paris Cité, 2018.

[TH-3] O. Seminck. Modèles cognitifs et computationnels de la résolution des pronoms. Thèse, Université Paris Cité, 2018. Thèse encadrée par Adeline Nazarenko.

2017 [TH-2] I. Bannour. Recherche d'information sémantique : graphe sémantico-documentaire et propagation d'activation. Thèse, Université Sorbonne Paris Cité, 2017.

[TH-1] E. Hassan. Reconnaissance d'expériences vécues dans les avis d'utilisateurs : une méthode basée sur les événements. Thèse, Université Sorbonne Paris Cité, 2017.

Habilitations

2018 [HR-1] E. Cartier. Dynamique lexicale des langues : éléments théoriques, méthodes automatiques, expérimentations en français contemporain. HdR, Université Sorbonne Paris Cité, 2018.

Rapports de contrats

2019 [RC-1] C. Nédellec, A. Nazarenko, F. Andre, C. Balivo, B. Daille, A. Drouot, J. Garcia Flores, L. Freyt-Caffin, N. Gandon, J. Lafait, N. Le Ba, C. Martin, N. Morcrette, S. Respingue-Perrin, S. Rennes et P. Zweigenbaum. Recommandations sur l'analyse automatique de documents : acquisition, gestion, exploration. Rapport Technique, Comité pour la science ouverte, 2019.

Publications antérieures à l'arrivée au LIPN

Les références comportent « ↑ » et correspondent à des publications antérieures à l'arrivée au LIPN, non référencées par ailleurs.

Signification des soulignements :
- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2017 **[↑RI-1]** A. Grezka et L. Zhu. Du figement au défigement : la reconnaissance de néologismes polylexicaux. *Etudes de Linguistique Appliquée*, 2/2017(186) :185–196, 2017.

Articles dans des revues nationales ou internationales

2018 **[↑RE-3]** J. Raad, W. Beek, N. Pernelle, F. Saïs et F. van Harmelen. Détection de liens d'identité erronés en utilisant la détection de communautés dans les graphes d'identité. *Ingénierie des Systèmes d Infrmation*, 23(3-4) :61–88, 2018.

[↑RE-2] J. Raad, N. Pernelle, F. Saïs, J. Dibie, L. Ibanescu et S. Dervaux. Comment représenter et découvrir des liens d'identité contextuelle dans une base de connaissances : Application à des données expérimentales en sciences du vivant. *Revue d'Intelligence Artificielle*, 32(3) :345–372, 2018.

2017 **[↑RE-1]** A. Grezka. Morfetik, un dictionnaire morphologique : illustration avec le lexique de la perception. *Écho des études romanes*, 13(1) :89–99, 2017.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2019 **[↑CI-5]** N. Ballier, T. Gaillat, A. Simpkin, B. Stearns, M. Bouyé et M. Zarrouk. A Supervised Learning Model for the Automatic Assessment of Language Levels Based on Learner Errors. *The 14th European Conference on Technology Enhanced Learning (EC-TEL 2019), Delft, The Netherlands, September 16-19, 2019*, volume 11722 de *Lecture Notes in Computer Science*, pages 308–320, 2019.

[↑CI-4] T. Mecharnia, L. Chibout Khelifa, N. Pernelle et F. Hamdi. An Approach Toward a Prediction of the Presence of Asbestos in Buildings Based on Incomplete Temporal Descriptions of Marketed Products. *The 10th International Conference on Knowledge Capture (K-CAP 2019), Marina Del Rey, CA, USA, November 19-21, 2019*, pages 239–242, 2019.

2018 **[↑CI-3]** J. Raad, W. Beek, F. van Harmelen, N. Pernelle et F. Saïs. Detecting Erroneous Identity Links on the Web Using Network Metrics. *The 17th International Semantic Web Conference (ISWC 2018), Monterey, CA, USA, October 8-12, 2018*, volume 11136 de *Lecture Notes in Computer Science*, pages 391–407, 2018.

2017 **[↑CI-2]** J. Raad, N. Pernelle et F. Saïs. Detection of Contextual Identity Links in a Knowledge Base. *The 2017 Knowledge Capture Conference (K-CAP 2017), Austin,TX, USA, December 4-6, 2017*, pages 8 :1–8, 2017.

[↑CI-1] D. Symeonidou, L. Galárraga, N. Pernelle, F. Saïs et F. M. Suchanek. VICKEY : Mining Conditional Keys on Knowledge Bases. *The 16th International Semantic Web Conference (ISWC 2017), Vienna, Austria, October 21-25, 2017*, volume 10587 de *Lecture Notes in Computer Science*, pages 661–677, 2017.

Autres articles dans des conférences ou workshops à comité de lecture

2019 [↑CO-7] M. Atencia, J. David, J. Euzenat, L. Ibanescu, N. Pernelle, F. Saïs, E. Thiéblin et C. Trojahn. Discovering Expressive Rules for Complex Ontology Matching and Data Interlinking. *The 14th International Workshop on Ontology Matching co-located with the 18th International Semantic Web Conference (OM@ISWC 2019), Auckland, New Zealand, October 26, 2019*, volume 2536 de *CEUR Workshop Proceedings*, pages 199–200, 2019.

[↑CO-6] N. Ballier, T. Gaillat, M. Zarrouk, A. Simpkin, M. Bouyé et B. Stearns. A Franco-Irish project for the automatic identification of criterial features in learners of English. *The 27th EUROCALL conference (EUROCALL 2019), Université de Louvain la Neuve, Belgium, August 28-31, 2019*, Louvain la Neuve, Belgium, 2019.

[↑CO-5] T. Gaillat, N. Ballier, M. Zarrouk, A. Simpkin, B. Stearns et M. Bouyé. Investigating Criterial Features of Learner English and Predicting CEFR Levels in French Learners of English. *The 27th EUROCALL Conference (EUROCALL 2019), Louvain-la-Neuve, Belgium, August 28-31, 2019*, 2019.

[↑CO-4] A. Khajeh, N. Pernelle et F. Saïs. Generating Referring Expressions from Knowledge Graphs. *The 14th International Workshop on Ontology Matching co-located with the 18th International Semantic Web Conference (OM@ISWC 2019), Auckland, New Zealand, October 26, 2019*, volume 2536 de *CEUR Workshop Proceedings*, pages 193–194, 2019.

[↑CO-3] T. Mecharnia, N. Pernelle, L. Chibout Khelifa et F. Hamdi. Approche de prédiction de présence d'amiante dans les bâtiments basée sur l'exploitation des descriptions temporelles incomplètes de produits commercialisés. *30èmes Journées francophones d'Ingénierie des Connaissance (IC 2019), Toulouse, France, July 2-4, 2019*, pages 144–157, 2019.

2017 [↑CO-2] D. Chistaras, N. Pernelle et F. Saïs. Approche numérique pour l'invalidation de liens d'identité (owl : SameAs). *28èmes Journées francophones d'Ingénierie des Connaissances (IC 2017), Caen, France, July 3-7, 2017*, pages 50–55, 2017.

[↑CO-1] J. Raad, N. Pernelle et F. Saïs. Détection de liens d'identité contextuels dans une base de connaissances. C. Roussey, éditeur, *28èmes Journées francophones d'Ingénierie des Connaissances (IC 2017), Caen, France, July 3-7, 2017*, pages 56–67, 2017.

Édition d'ouvrages collectifs

2018 [↑ED-2] P. Chapman, D. Endres et N. Pernelle. *The 23rd International Conference on Conceptual Structures (ICCS 2018), Edinburgh, UK, June 20-22, 2018*, volume 10872 de *Lecture Notes in Computer Science*. Springer, 2018.

2017 [↑ED-1] A. Grezka et M.-H. Viguier. *Figement en mouvement et complexité en devenir : regards sur quelques problèmes liés aux structures défigées et complexes*. Numéro 2/2017 Études de linguistique appliquée. klincksieck (didier erudition jusqu'en 2003), 2017.

3.3 Domaine 4. Inscription des activités de recherche dans la société

3.3.1 Référence 1. L'équipe se distingue par la qualité et la quantité de ses interactions avec le monde non-académique.

En raison de son expertise reconnue dans la conception de modèles d'apprentissage et de gestion de connaissances dédiés aux données textuelles et aux graphes de connaissances, l'équipe est régulièrement sollicitée pour des collaborations industrielles.

Sur la période on compte 3 prestations de service réalisées pour des entreprises : Crowd Prediction (2021–2023 et 2022-2023), API Conseil (2021-2022).

Au niveau doctoral, on dénombre 5 thèses sur des financements industriels dont 3 thèses CIFRE avec les entreprises Pixalione (2016-2019, thèse soutenue en 2018) Thalès DMS (2021-2024, thèse en cours), Iniciativas (2021-2024, thèse en cours). Une thèse CIFRE avec la BPCE et l'équipe A3 a par ailleurs démarré en avril 2023 suite à

un co-encadrement d'un stage Master de recherche avec cet établissement en 2022. En outre, une thèse s'est déroulée au CEA et a été financée à 100% par cet établissement (thèse soutenue en 2021).

La quatrième thèse industrielle (soutenue en 2021) a été réalisée dans le cadre d'un projet FUI pour un financement LIPN de 213k€ (2017-2020). Il s'agit du projet FUI PCUv2 dont le consortium comprenait, outre le LIPN, les entreprises Smile, Proxem, Armadillo et Wallix. Ce projet a aussi permis le financement d'un postdoc.

3.3.2 Référence 2. L'équipe développe des produits à destination du monde culturel, économique et social.

L'équipe développe aussi des recherches à destination du monde culturel, social ou de la santé. Un travail en collaboration avec un psychologue de l'université de Galway sur la détection des émotions chez les personnes atteintes d'autisme est en cours. L'enjeu est de permettre le développement d'applications personnalisées pour la communication et les interactions sociales de ce public spécifique. Du fait d'une expression particulière de l'émotion par cette population, les systèmes actuels ne sont pas efficaces. Centrée sur l'interaction personne-ordinateur, au lieu de l'interaction classique personne-thérapeute, un jeu de données multimodales comprenant 57012 exemples annotés qui permettra d'entraîner des modèles d'apprentissage, est en cours d'annotation et sera disponible pour les chercheurs, notamment les chercheurs dans le domaine de la santé.

Une autre action développée par l'équipe s'intéresse au développement d'une plateforme d'accès à des textes philosophiques (https://tal.lipn.univ-paris13.fr/bni/). Ce travail s'inscrit dans le champ des humanités numériques et associent plusieurs éditeurs pour fournir les fonds documentaires. La plateforme a une ambition de démocratisation de l'accès aux philosophes et à leurs oeuvres. Elle offre des fonctionnalités d'accès riches par l'exploitation de réseaux sémantiques. Deux types de réseau sont construits : des grilles de lecture par auteur qui présentent une vision sypnoptique des oeuvres et des thésaurus qui recensent les termes du domaine permettent ainsi des navigations à ces différents niveaux. D'autres fonctionnalités vont enrichir les possibilités de navigation dans les oeuvres. Actuellement, la plateforme indexe les oeuvres de Adam Smith et Francis Hutcheson qui seront complétées par de nouveaux auteurs ainsi que par des oeuvres qui traitent des auteurs (préfaces, correspondance...). Elle a été présentée à la fête de la sciences en 2017.

Chapitre 4
Trajectoire de l'équipe RCLN

Sommaire

4.1 SWOT 329
4.2 Trajectoire 331
4.2.1 Langues peu dotées 331
4.2.2 Algorithmes pour l'analyse syntaxique et sémantique de textes . . 332
4.2.3 Exploitation conjointe des modèles de langue et des graphes de connaissances 333
4.2.4 Fouille de règles pour l'enrichissement des graphes de connaissances 334
4.2.5 Stratégie partenariale locale, nationale et internationale, et Stratégie sociétale 335

Nous présentons tout d'abord une analyse SWOT afin d'évaluer nos forces, faiblesses et opportunités, puis présentons nos objectifs pour la prochaine période.

4.1 SWOT

- **Forces et Atouts**

  — *Complémentarité des compétences* : les compétences des membres de l'équipe

sont particulièrement complémentaires à la fois en terme d'objectifs puisqu'elles vont de l'analyse syntaxique à la construction et l'enrichissement de graphes de connaissances et en terme de méthodes utilisées (linguistique de corpus, apprentissage profond, apprentissage symbolique et optimisation combinatoire). Les recrutements sur la période (1 fin 2019 et 1 en 2020) ont permis de consolider l'axe Acquisition de connaissances, et cela permet d'envisager encore plus solidement des perspectives de recherche articulant traitement automatique des langues et graphes de connaissances.

— *Visibilité* : le positionnement original de l'équipe sur les problèmes difficiles en langue, sur les langues sous-dotées ou langues de spécialités ainsi que ces compétences sur la construction et l'enrichissement de graphes de connaissances lui confère une bonne visibilité nationale et internationale qui est attestée par sa participation active dans des instances nationales, des comités de programme de conférences internationales de haut niveau et des organisations de workshops internationaux.

— *Présence de nombreux jeunes chercheurs* : malgré le faible nombre d'HDR, et l'absence de Master de l'USPN portant sur les thématiques de RCLN, le nombre de doctorants (co-)encadrés est plus élevé que sur la période précédente (17 à l'ED Galilée, et 26 co-encadrements au total sur la période), avec des recrutements d'étudiants au niveau local, national et international. Une attractivité certainement due en partie à la participation de l'équipe au labex EFL.

— *Projets et contrats de recherche* : le positionnement de l'équipe à l'interface de l'informatique et de la linguistique ont permis de développer de nombreuses actions au sein du labex EFL, et de porter ou de participer à un nombre conséquent de projets académiques et industriels (3 ANR, 1 ITN H2020, 3 projets bilatéraux, 1 FUI, 3 thèses CIFRE, contrats d'expertise auprès d'industriels).

— *Valorisation logicielle et mise à disposition de ressources* : la présence d'un IR CNRS et la participation active de deux des membres RCLN à la Dev team du LIPN permet de faciliter la réalisation et la mise à disposition de la communauté scientifique de nombreux logiciels et prototypes de recherche ainsi que de nombreuses ressources (Graphes de connaissances, benchmarks, ressource morphologique).

— *Parité* : Sur 15 membres permanents, 7 membres sont des femmes (47%), et sur la période parmi les 23 doctorants et post-doctorants recrutés, 10 sont des femmes (43%).

- **Faiblesses**

— *Faible nombre d'HDR* : seuls les 3 PR de l'équipe sont habilités, et l'un d'entre eux a pris d'importantes fonctions au sein du CNRS depuis le début de la période. Cependant, deux habilitations de maîtres de conférences sont en cours de rédaction et leur soutenance est prévue fin 2023.

— *Absence de chercheur* : depuis 2009, l'équipe ne compte plus de chercheur permanent, ce qui limite sa capacité d'action dans un contexte où les charges et responsabilités au niveau local et national sont lourdes.

Trajectoire de l'équipe RCLN

- **Opportunités**

— *Collaborations au sein du LIPN* : la présence au LIPN d'une équipe de recherche spécialisée en Apprentissage (A3) et en Optimisation Combinatoire (A0C) permet à l'équipe de mener des recherches en collaborations avec des membres de ces deux équipes. Ces collaborations ont été renforcées durant la période par trois co-encadrement de thèses dont l'une a démarré début 2023, 9 co-publications ainsi que des participations et nouveaux dépôts de projets communs (1 FUI, 1 ANR en cours, 2 ANR déposées en 2023). La constitution d'un axe transverse Science des

Données au LIPN va sans doute nous amener à pérenniser ces collaborations et explorer de nouvelles thématiques de recherche communes.
— *Projets internationaux* : les collaborations internationales actuelles de l'équipe devraient à terme déboucher sur des projets internationaux plus nombreux.
— *Nouvelles technologies* : l'apparition de nouvelles technologies et de nouvelles ressources telles que les grands modèles de langues (e.g. BERT, GPT) font émerger de nouveaux usages des outils de TAL et de nouvelles thématiques de recherche pour lesquelles les compétences complémentaires de l'équipe en TAL et en graphes de connaissance peuvent être mises à profit.

- **Menaces**

— *Labex EFL* : il existe une incertitude quant au renouvellement du Labex après 2024 et cela pourrait entraîner une diminution importante des ressources de l'équipe. Cependant RCLN participe très activement à la proposition de renouvellement du Labex qui sera déposée fin 2023.
— *Recrutement et encadrement des doctorants* : l'absence de master adossé aux thématiques de l'équipe complique le recrutement des doctorants et le faible nombre d'HDR incite les MCF de l'équipe à mener des co-encadrements en collaboration avec d'autres laboratoires.

4.2 Trajectoire

introduction

Depuis la création de l'équipe et jusqu'aux années 2000, l'équipe RCLN a essentiellement développé des méthodes de TAL basées sur des modèles symboliques, et des recherches en ingénierie des connaissances. A partir des années 2010, les approches développées ont évolué vers l'apprentissage automatique que ce soit pour l'analyse ou l'acquisition automatique de connaissances. Les progrès réalisés dans le domaine de l'apprentissage profond au cours des quinze dernières années ont entraîné une amélioration indéniable des performances dans le traitement du langage naturel (NLP) ou dans celui des graphes de connaissance, qu'il s'agisse de tâches telles que la traduction automatique, l'analyse syntaxique, la génération automatique de textes ou la complétion de graphes. Depuis la dernière période, l'équipe RCLN s'est en grande majorité tournée vers des recherches basées sur ces méthodes. Dans le précédent rapport HCERES, les perspectives mentionnaient (1) le développement d'approches pour l'analyse syntaxique et sémantique, et en particulier l'utilisation des réseaux de neurones, et des méthodes issues de l'optimisation combinatoire, (2) l'annotation sémantique et exploration textuelle et (3) l'acquisition de connaissances à partir de textes. Nous avons pu développer des approches innovantes pour chacune de ces thématiques et avons appliqué certaines d'entre elles au thème transverse concernant la littérature scientifique. En revanche, le départ du professeur responsable du deuxième axe transverse proposé concernant la lecture automatique nous a amené à abandonner cet axe de recherche.

Nos objectifs de recherche actuels concernent essentiellement le traitement automatique de langues sous-dotées, les algorithmes pour l'analyse syntaxique et sémantique de textes, l'exploitation conjointe de textes et de graphes de connaissances et la fouille de règles pour l'enrichissement de graphes de connaissances.

Trajectoire de l'équipe RCLN

4.2.1 Langues peu dotées

Nous souhaitons poursuivre nos recherches sur la définition d'approches de TAL dédiées aux langues peu dotées afin de mettre à disposition des ressources numériques pour ces langues et parce que ces langues posent souvent des questions intéressantes

sur le plan linguistique et sur leur traitement automatique. Le traitement de la variation linguistique représente un défi important à la fois en linguistique et en TAL, en particulier pour les langues sous-dotées, et les modèles développés dans le cadre d'une langue et de ses variations peuvent servir de référence pour l'annotation d'autres langues peu documentées. Ainsi, malgré une tradition écrite multi-séculaire, l'arménien manque cruellement de ressources numériques. La langue arménienne comprend l'arménien classique (5e-10e siècle après J.-C.), l'arménien moyen (11e-17e siècle) et l'arménien moderne (17e siècle jusqu'à aujourd'hui) avec ses deux normes : l'arménien moderne oriental et l'arménien moderne occidental. Outre les deux normes, le continuum linguistique arménien comprend divers dialectes, ainsi que des formes vernaculaires mais toutes les variétés utilisent un unique alphabet. Nous sommes partenaires du projet ANR DALHI (Digitizing Armenian Linguistic Heritage : Armenian Multivariational Corpus and DataProcessing) qui vient de débuter et qui vise à construire pour la première fois une plateforme numérique linguistique libre d'accès, pour l'ensemble des variations de la langue arménienne. Six corpus annotés concernant chacun différentes variétés de cette langue sont en cours de constitution par des membres de l'INALCO et de spécialistes de l'arménien et l'équipe RCLN se focalisera sur la définition de méthodes pour le traitement des langues avec des données rares, afin de proposer des alternatives aux tendances axées sur les données quantitatives. Les développements récents dans le domaine de l'apprentissage profond en TAL ont montré l'intérêt des modèles linguistiques pré-entraînés, basés sur les plongements de mots contextuels, tels que BERT et m-BERT. Ces modèles sont entraînés sur une grande quantité de textes non annotés, en tenant compte du contexte bidirectionnel. Les modèles de langue multilingues tels que m-Bert présentent deux avantages : ils évitent de former un modèle pour chaque variété et tirent parti des similitudes structurelles pour apprendre des caractéristiques interlinguistiques, ce qui peut être considéré comme une forme de transfert des langues riches vers les langues plus pauvres en ressources. Les modèles qui ont été développés pour l'arabe standard serviront de base pour le transfert vers des dialectes moins dotés et nous étudierons en particulier le caractère indispensable d'une modélisation explicite de la morphologie en la comparant à une modélisation implicite par les modèles de langues (multilingue) utilisés pour le transfert. Nous appliquerons ces approches à la langue arménienne pour étudier le caractère généralisable de nos conclusions mais aussi pour développer, efficacement, des modèles pour l'arménien pour laquelle nous collectons actuellement les données dans le cadre du projet DALiH.

4.2.2 Algorithmes pour l'analyse syntaxique et sémantique de textes

Trajectoire de l'équipe RCLN

Les approches basées sur de l'apprentissage profond sont à la fois gourmandes en données, sont difficilement appliquées aux langues à faibles ressources, et sont gourmandes en calculs puisqu'elles nécessitent un nombre sans cesse croissant de paramètres pour améliorer les modèles précédents. En outre, les manipulations symboliques telles que la logique et les structures, qui sont de nature combinatoire, sont encore mal prises en compte par la nature continue des réseaux neuronaux. À titre d'exemple, même les modèles récents d'analyse syntaxique des constituants ou des dépendances nécessitent encore un algorithme combinatoire pour construire un arbre syntaxique valide, puisque la solution du classifieur seul peut générer des graphes non valides. Enfin, ces modèles sont monolithiques et manquent d'un mécanisme d'accord entre les tâches en plus du partage des paramètres. Par exemple, l'analyse syntaxique avec des grammaires lexicalisées (divisées en choix de la structure de parangon correcte pour chaque mot dans le contexte, et en construction d'une analyse syntaxique), l'extraction de relations (identification d'entités et de prédicats, et établissement de liens entre eux) peuvent naturellement être exprimées comme des sous-tâches qui doivent s'accorder. Le plus souvent, ces sous-tâches pourraient être simples à résoudre de manière

isolée, mais la recherche d'un accord ajoute de la complexité qui rend les tâches globales difficiles ou irréalisables. Nous proposons de poursuivre notre collaboration avec l'équipe AOC pour étudier comment faire prédire à un réseau neuronal une solution approximative d'un processus d'optimisation. Par exemple, les algorithmes d'optimisation itératifs couramment utilisés dans le cadre lagrangien peuvent être remplacés par des algorithmes amortis et semi-amortis, d'une manière similaire à l'amortissement des distributions de propositions dans les méthodes variationnelles probabilistes [1] et plus généralement dans l'optimisation continue [**?**].

4.2.3 Exploitation conjointe des modèles de langue et des graphes de connaissances

Les grands modèles de langues tels que la famille des GPT-x et les technologies associées offrent des possibilités sans précédent à la société dans son ensemble, et plus particulièrement aux chercheurs : elles peuvent permettre de communiquer des résultats scientifiques de manière plus efficace, via une interaction homme-machine, ce qui peut libérer du temps pour se concentrer par exemple sur le développement d'idées novatrices. Toutefois, en dépit de l'engouement qu'elles suscitent, ces technologies ont de nombreuses limites : ces modèles ne peuvent pas être librement adaptés par les individus en fonction de leurs besoins, les textes générés peuvent contenir des informations incorrectes ou biaisées, elles s'appuient sur des modèles de langues de très grande taille, qui sont difficiles à constituer, et enfin, il existe des préoccupations éthiques quant à leur utilisation abusive pour générer de la fausse science et du plagiat.

Au cours des dernières années, nous avons assisté à l'émergence de graphes de connaissances (KG) qui décrivent explicitement les connaissances de la recherche dans le but de permettre à des systèmes intelligents de soutenir et accélérer les recherches scientifique. Ces ressources caractérisent des entités telles que les tâches, les méthodes, les techniques d'évaluation, ou des objets d'études telles que des protéines ou des produits chimiques), leurs relations, et les acteurs impliqués ou documents pertinents (e.g. chercheurs, organisations, articles). Leur utilisation peuvent permettre d'aider à construire des systèmes intelligents pour expliquer, conforter ou réfuter des assertions générées automatiquement ou pour fournir une aide à la rédaction intelligente de texte. Certains de ces graphes sont alimentés par crowd-sourcing (par exemple, ORKG, UMLS, Nanopublications), tandis que d'autres sont générés automatiquement à partir du texte et des métadonnées des articles de recherche tels que AI-KG que nous avons co-développé, CSO, ou encore TKG. Comme beaucoup d'autres KG, ceux qui décrivent les entités liées à la recherche souffrent d'incomplétude. Aussi nous poursuivrons nos recherches sur les problématiques de complétion automatique de graphes. Le problème de l'incomplétude des graphes de connaissances est généralement traité par des techniques de prédiction de liens ou de classification de triplets, qui ont donné de bons résultats dans plusieurs domaines. Ces méthodes utilisent souvent des modèles de plongement de graphes (e.g. TransE, RotatE, ComplEx), des caractéristiques basées sur les chemins, ou des réseaux de neurones graphiques (GNN). Nous étendrons les travaux de recherche actuels sur les GNN, en examinant des modèles plus avancés tels que les Graph Attention Networks (GAT) [5] et leurs variantes. Par exemple, la particularité de prédire les relations dans un graphe tel que AI-KG est que les connexions de ce graphe ont tendance à se concentrer autour de certains noeuds importants (graphes en étoile). Par conséquent, il existe un certain biais en faveur des noeuds ayant un degré plus élevé qu'il faut pouvoir contrôler. Pour limiter les risques d'utilisation abusive des modèles génératifs dans la production de contenu académique, le contenu généré doit être expliqué. L'explicabilité est un moyen de valider le contenu généré par des recherches dans des sources de connaissances. Dans le cas de textes scientifiques, cette méthode peut être mise en oeuvre en reliant le texte à d'autres travaux, avec de multiples avantages : l'explication de la suggestion elle-même, une suggestion de citation, et également un aspect didactique permettant d'apprendre qui a travaillé sur des sujets

similaires. Nous souhaitons mettre en oeuvre des modèles de langage qui combinent des informations lexicales et sémantiques, en tirant parti de graphes de connaissances linguistiques bidirectionnels profonds, et en utilisant conjointement les graphes de citations et les graphes de connaissances scientifiques.

Face à la prolifération des contenus générés automatiquement et aux risques d'inconduite scientifique que cela représente pour le monde académique, il est essentiel de développer des méthodes de détection efficaces pour disqualifier les textes et en particulier les articles de recherche générés automatiquement. Nous avons récemment proposé un benchmark qui s'inscrit dans ce cadre et qui inclut des textes générés automatiquement mais partiellement manipulés. Nous avons considéré la détection comme une tâche de classification binaire et testé divers algorithmes de classification. Les résultats montrent que les modèles de classification récents existants ont pu fournir une précision maximale de 70,2 % sur notre ensemble de données. Ce résultat montre que ce problème reste ouvert. Aussi, nous envisageons de développer des méthodes originales orientées vers la détection de textes et de fragments générés automatiquement, en tirant éventuellement parti des connaissances pour détecter les contradictions et les phrases hors contexte.

4.2.4 Fouille de règles pour l'enrichissement des graphes de connaissances

Découvrir des règles de la logique du premier ordre dans de grands graphes de connaissances (KG) a été abordé dans différents travaux qui diffèrent selon les biais de langage utilisés, ou les stratégies d'exploration de l'espace des règles [2, 4, 3]. Ces règles peuvent permettre d'expliciter des motifs implicites dans les graphes, de détecter des données erronées ou de faire de la complétion de graphes. Bien que les approches existantes soient efficaces, elles ne sont pas en mesure de découvrir des règles qui impliquent des contraintes définies sur des valeurs numériques ou temporelles. Or, nous pensons que ces contraintes peuvent être très pertinentes dans des domaines tels que les sciences sociales, la santé publique ou la littérature scientifique en général. Nous nous intéresserons à la découverte de règles impliquant des contraintes numériques telles que des intervalles dans de grands graphes de connaissance. Nous chercherons à sélectionner efficacement les prédicats numériques et les contraintes numériques en utilisant des techniques supervisées qui distinguent le mieux les prédictions correctes et non correctes d'une règle détectée par un système de fouille de règles efficace existant.

De plus, nos recherches sur la fouille de règles dans les graphes de connaissances se sont récemment focalisées sur l'extraction de règles causales. Ces approches ont montré que l'effet causal peut être évalué pour des graphes de connaissances par une estimation des contre-factuels en se basant sur des individus similaires. Cependant, nous avons montré que les méthodes de calcul des similarités de deux descriptions impactent fortement le nombre et la qualité des règles causales découvertes (i.e. une similarité trop contrainte réduit fortement le nombre de règles mais garantit une meilleure qualité de la règle) [CI-29]. Nous étudierons comment des approches de fouille de règles causales peuvent se baser sur des profils générés par méthodes énumératives non supervisées définies pour les graphes attribués. Nous étudierons en particulier comment ces méthodes peuvent être utilisées pour évaluer l'influence des messages alimentaires de différentes natures sur le comportement alimentaires d'élèves (collaboration qui vient de débuter avec le LEPS de l'USPN et l'équipe A3), et donc dans un cadre où l'on dispose à la fois des descriptions des élèves, des relations qu'ils entretiennent et des descriptions des messages alimentaires qu'ils reçoivent.

Trajectoire de l'équipe RCLN

4.2.5 Stratégie partenariale locale, nationale et internationale, et Stratégie sociétale

Nous souhaitons poursuivre et renforcer nos partenariats avec l'équipe AOC et A3 au sein du laboratoire. Pour cela, nous favorisons le dépôt de projets communs et le co-encadrement de thèses, et participerons activement à l'axe transverse Sciences des données afin de faire émerger de nouvelles thématiques de recherche communes.

La convention LabEx avec l'IDEX prendra fin en décembre 2024. Les opérations financées par le LabEx et portées par des membres de l'équipe RCLN ont clairement contribué à soutenir les thématiques de recherche de l'équipe (recrutements de doctorants, de post-doctorants, financement de bourses de mobilité, de matériels, nouvelles collaborations nationales et internationales). Aussi, l'ensemble des partenaires est pleinement investi pour répondre à l'appel à soumission de nouveaux projets LabEx de l'IDEX Paris Cité. Sur le plan thématique, le projet de l'axe Analyse sémantique computationnelle (axe 5) dans lequel nous sommes impliqués consiste à tirer parti du nouveau paradigme marqué par les évolutions technologiques récentes que sont les modèles de langue dont la dernière incarnation est chatGPT. L'axe souhaite étudier comment les modèles de langue (en général) peuvent contribuer à apporter de nouvelles connaissances, à étudier ces nouveaux modèles pour mieux les comprendre et à développer des activités applicatives tant en direction de la société en général que pour la communauté en sciences du langage, activités telles que l'extraction d'information, la vérification de l'information générée, l'étude et la conception de systèmes d'apprentissage pour les langues rares et peu dotées. Aussi, les perspectives de l'équipe en NLP s'inscrivent pleinement dans ce nouveau projet.

En janvier 2023, l'USPN vient de déposer un projet d'alliance européenne. L'objectif du projet UNINOVIS - Data for L.I.F.E. est de créer une alliance de sept établissements d'enseignement supérieur publics répartis dans toute l'Europe (University Sorbonne - Paris Nord, Technical University of Applied Sciences - Würzburg-Schweinfurt - Germany, University of Campania Luigi Vanvitelli - Caserta - Italy, University of Applied Sciences - Tampere - Finland, University of Applied Sciences - Kaunas - Lithuania, University of Malaga - Spain, University of Tirana - Albania). Cette alliance souhaite combiner son expertise afin d'améliorer les capacités et les compétences des européens pour comprendre, analyser, gérer, utiliser et protéger les données, et ainsi contribuer à certains des défis majeurs de l'Europe du 21e siècle. UNINOVIS vise à devenir l'université européenne leader dans l'apprentissage des sciences des données dans tous les domaines, de la médecine aux biotechnologies, de l'ingénierie à la linguistique, des affaires aux sciences sociales, de l'économie aux sciences humaines. L'équipe RCLN s'est impliquée dans le montage du projet et s'investira dans les actions pédagogiques menées en collaboration avec les autres universités afin que ses compétences en IA et en NLP puissent être mises à profit afin que les étudiants de tout domaine puissent mieux comprendre et manipuler les données issues des nouveaux outils de NLP ou des graphes de données. De même, la problématique des nouveaux usages des modèles de langue pour la littérature scientifique, nous a conduit à être partenaire d'un projet Horizon CL4 2023 qui est en cours d'évaluation.

Nous formons une équipe de nature interdisciplinaire puisqu'elle inclut des spécialistes en linguistique de corpus et des informaticiens. Nous souhaitons non seulement travailler sur des applications qui bénéficient à des experts issus d'autres domaines que l'informatique, comme nous le faisons en partie (linguistique, santé publique, droit, environnement), mais aussi renforcer les aspects interdisciplinaires de nos recherches en collaborant plus intensivement avec des experts de domaine, comme c'est le cas pour la collaboration initiée avec LEPS sur des thématiques de science de l'éducation et de nutrition.

Références complémentaires

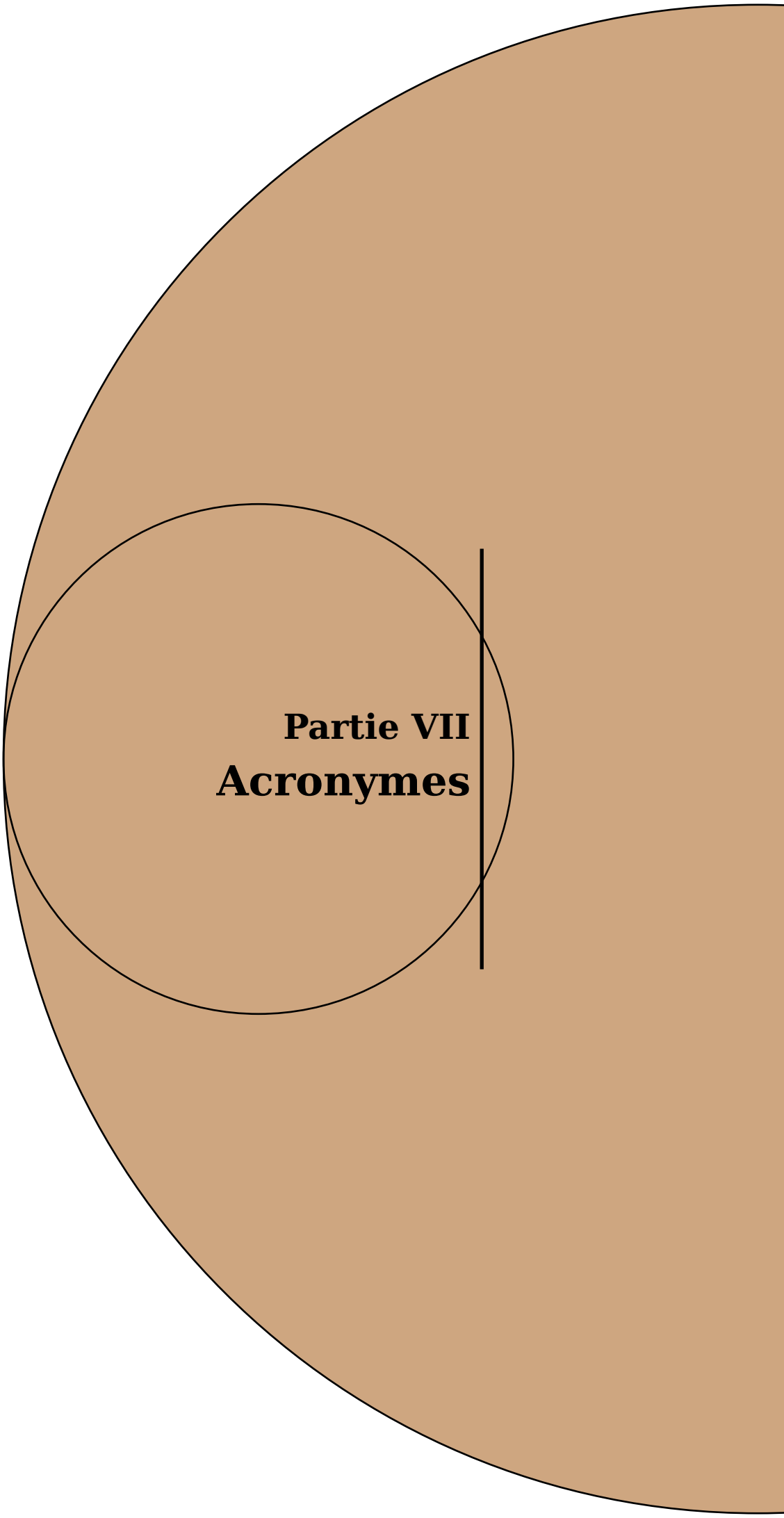
[1] D. P. Kingma and M. Welling. Auto-encoding variational bayes. In Y. Bengio and Y. LeCun, editors, *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014.

[2] J. Lajus, L. Galárraga, and F. M. Suchanek. Fast and exact rule mining with AMIE 3. In A. Harth, S. Kirrane, A.-C. N. Ngomo, H. Paulheim, A. Rula, A. L. Gentile, P. Haase, and M. Cochez, editors, *The Semantic Web - 17th International Conference, ESWC 2020, Heraklion, Crete, Greece, May 31-June 4, 2020, Proceedings*, volume 12123 of *Lecture Notes in Computer Science*, pages 36–52. Springer, 2020.

[3] C. Meilicke, M. W. Chekol, M. Fink, and H. Stuckenschmidt. Reinforced anytime bottom up rule learning for knowledge graph completion. *CoRR*, abs/2004.04412, 2020.

[4] S. Ortona, V. V. Meduri, and P. Papotti. Rudik : Rule discovery in knowledge bases. *Proc. VLDB Endow.*, 11(12) :1946 ?1949, aug 2018.

[5] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio. Graph attention networks. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018.



Partie VII
Acronymes

A^3 Apprentissage Artificiel et Applications

AI Assistant Ingénieur (CNRS)

ANR Agence Nationale de la Recherche

AOC Algorithmes et Optimisation Combinatoire

ASPC Alliance Sorbonne-Paris-Cité

ATER Attaché Temporaire d'Enseignement et de Recherche

BIATSS Bibliothécaires, Ingénieurs, Administratifs, Techniciens, de Service et de Santé

BQR Bonus Qualité Recherche (Université Sorbonne Paris Nord)

CALIN Combinatoire, Algorithmique et Interactions

CD Contrat Doctoral

CDSN Contrat Doctoral Spécial Normalien

CDD Contrat à Durée Déterminée

CER Comité d'Éthique de la Recherche

CID Commission Interdisciplinaire (CoNRS)

CIFRE Conventions Industrielles de Formation par la Recherche

CoNRS Comité National de la Recherche Scientifique

CNRS Centre National de la Recherche Scientifique

CNU Conseil National des Universités

Complexités Axe transverse Complexités

COMUE Communauté d'Universités et Établissements

CR Chargé de Recherche

CRCT Congé pour Recherches ou Conversion Thématique

DR Directeur de Recherche

DIM Domaine d'Intérêt Majeur

EA Equipe d'Accueil

ECOS Evaluation-orientation de la COopération Scientifique

ED École Doctorale

EDF Electricité de France

EFL Labex *Empirical Foundations of Linguistics*

ENPC École nationale des ponts et chaussées

ERC European Research Council

ESSEC Ecole Supérieure des Sciences Economiques et Commerciales

ETP personnel Equivalent Temps Plein

EUR École Universitaire de Recherche

FR Fédération de Recherche

FUI Fonds Unique Interministériel

GDR Groupement de Recherche

GDRI Groupement de Recherche International

GT Groupe de Travail

HCERES Haut Conseil de l'Évaluation de la Recherche et de l'Enseignement Supérieur

HDR Habilitation à Diriger des Recherches

IA Intelligence Artificielle

IdEx Initiative d'Excellence (appel Grand Emprunt)

IE Ingénieur d'Études (CNRS)

IEA International Emerging Actions

IG Institut Galilée

IGE Ingénieur d'Études (Université)

IHU Institut Hospitalo-Universitaire

IMJ Institut de Mathématiques de Jussieu

IML Institut de Mathématiques de Luminy - UMR 6206

IML Institut de Mathématiques de Luminy - UMR 6206

Inria Institut National de la Recherche en Informatique et Automatique

INS2I Institut des Sciences de l'Information et de leurs Interactions, CNRS

IR Ingénieur de Recherche (CNRS)

IREM Institut de Recherche sur l'Enseignement des Mathématiques

IRIF Institut de Recherche en Informatique Fondamentale - UMR 8243

IRN International Research Network

IT Ingénieurs, techniciens et personnels administratifs CNRS

IUT Institut Universitaire Technologique

IUTSD IUT de Saint-Denis

IUTV IUT de Villetaneuse

IUF Institut Universitaire de France

JCJC Jeunes Chercheuses et Jeunes Chercheurs

L2TI Laboratoire de traitement et transport de l'information, USPN - EA 3043

LabCom Laboratoire commun (Appel PIA 4)

LabEx Laboratoire d'Excellence (appel Grand Emprunt)

LaBRI Laboratoire Bordelais de Recherche en Informatique - UMR 5800

LAGA Laboratoire Analyse, Géométrie et Applications, USPN - UMR 7539

LAMA LAboratoire de MAthématiques - UMR 5127 & Université Savoie Mont Blanc

LoVE Logique et Vérification

LCR Logique, Calcul et Raisonnement

LIP6 Laboratoire d'Informatique de Paris 6

LIPN Laboratoire d'Informatique de Paris-Nord, Université Sorbonne Paris Nord - UMR 7030

LLSHS Lettres, Langues, Sciences Humaines et des Sociétés

LMFI Logique Mathématique et Fondements de l'Informatique

LMF Laboratoire de Mathodes Formelles

LoVe Logique et Vérification

MathSTIC Fédération de Recherche MathSTIC, Université Sorbonne Paris Nord - FR 3734

MCF Maître de Conférences

MPRI Master Parisien de Recherche en Informatique

MPRO Master Parisien de Recherche Opérationnelle

OCAD Optimisation Combinatoire et Algorithmique Distribuée

PAST Professeur Associé à Service Temporaire

PEPR Programmes et Équipements Prioritaires de Recherche
PHC Programme Hubert Curien
PIA Programme d'Investissements d'Avenir
Prag Professeur agrégé
PU Professeur des Universités
RCLN Représentation des Connaissances et Langage Naturel
RFSI Réseau Francilien en Sciences Informatiques
SHS Sciences Humaines et Sociales
SATT Société d'Accélération du Transfert de Technologie
UE Union européenne
UFI Unniversité Franco-Italienne
UFR Unité de Formation et de Recherche
UMR Unité Mixte de Recherche
UNIF Université numérique Paris Ile-de-France
UR Unité de recherche
USPC Université Sorbonne Paris Cité
USPN Université Sorbonne Paris Nord
USTH Université des Sciences et des Technologies de Hanoi, http://usth.edu.vn/fr/
VP Vice-Président (de USPN)