

## Document d'autoévaluation (DAE) Unité de recherche pluri-équipes

### CAMPAGNE D'ÉVALUATION 2022-2023 - VAGUE C

**REMARQUE :** Les données de caractérisation fournies dans le fichier XLS au format préconisé par l'HCERES (et donc très contraint et non modifiable) contient des éléments provenant d'extraction automatique du système d'information de l'Université de Tours (listes des thèses, budgets récurrents et des projets/financement obtenus, ...). Ces éléments ne correspondent pas exactement aux données réelles gérés en interne par le LIFAT et ses équipes (par exemple, certains doctorants ou financements obtenus ont été gérés par d'autres organismes). **Le LIFAT a exploité ses indicateurs (gérés en interne) pour produire les statistiques incluses dans ce fichier plutôt que les données de caractérisation extraites automatiquement et donc moins précises.**

## INFORMATIONS GÉNÉRALES

Nom de l'unité pour le contrat en cours : **Laboratoire d'Informatique Fondamentale et Appliquée de Tours**

Acronyme pour le contrat en cours : **LIFAT**

Label et numéro : **EA 6300 – ERL CNRS 7002**

Nombre d'équipes : **3 équipes**

Domaine scientifique principal : **ST : Sciences et Technologies**

**Panels scientifiques (dans la nomenclature du Hcéres) par ordre décroissant d'importance :**

**Panel 1 :** ST5 : Sciences pour l'ingénieur

**Panel 2 :** ST6 : Sciences et technologies de l'information et de la communication - STIC

**Panel 3 :** ST1 : Mathématiques

**Panel 4** Choisissez un élément.

Directrice / directeur pour le contrat en cours :

- **Jean-Yves RAMEL (jusqu'au 31/08/2021)**

- **Hubert CARDOT (depuis le 01/09/2021)**

**Établissements et organismes de rattachement (tutelles) :**

Liste des établissements et organismes de rattachement de l'unité de recherche pour le contrat en cours

- **Université de Tours**

- **INSA Centre Val de Loire**

- **CNRS pour l'ERL 7002 ROOT**

# TABLE DES MATIERES

INFORMATIONS GÉNÉRALES .....	1
TABLE DES MATIERES .....	2
1. Présentation générale du LIFAT .....	3
1.1. Historique .....	3
1.2. Localisation .....	4
1.3. Structuration .....	4
1.4. Thématiques scientifiques .....	5
2. Profil d'activités .....	8
3. Environnement de recherche .....	9
3.1. Implication dans les structures fédératives régionales .....	9
3.2. Implication dans les formations à la recherche .....	9
3.3. Accueil de stagiaires de Master .....	10
3.4. Production de documents pédagogiques .....	10
4. Prise en compte des recommandations du précédent rapport niveau LIFAT .....	10
4.1. Rappels des résultats et objectifs .....	10
4.2. Rappel des objectifs et stratégie de l'unité de recherche pour le contrat en cours .....	11
DOCUMENT D'AUTOÉVALUATION .....	13
1. Domaine d'évaluation 1 : Profil, ressources et organisation de l'unité .....	13
1.1. Référence 1. Profil d'activités et environnement de recherche .....	13
1.2. Référence 2. Positionnement et objectifs scientifiques .....	15
1.3. Référence 3. Fonctionnement et GRH .....	17
1.4. Synthèse de l'autoévaluation .....	22
2. Domaine d'évaluation 2 : Attractivité .....	23
2.1. Référence 1. Rayonnement scientifique .....	23
2.2. Référence 2. Accueil et formation de chercheurs (juniors et seniors) .....	26
2.4. Référence 3. Participation à des appels à projets compétitifs .....	27
2.5. Référence 4. Equipements lourds et compétences technologiques .....	29
2.3. Synthèse de l'autoévaluation .....	30
3. Domaine d'évaluation 3 : Production scientifique .....	31
3.1. Au niveau du laboratoire .....	31
3.2. Equipe BDTLN .....	33
3.3. Equipe RFAI .....	48
3.4. Equipe ROOT .....	56
4. Domaine d'évaluation 4 : Impacts sociétaux et économiques .....	68
4.1. Au niveau du laboratoire .....	68
4.2. Equipe BDTLN .....	69
4.3. Equipe RFAI .....	71
4.4. Equipe ROOT .....	73
ANNEXE A : Données LABRI - Activités à l'international du LIFAT .....	76
ANNEXE B : ANNEXE RFAI DU DAE .....	78
ANNEXE C : ANNEXE ROOT DU DAE .....	84
ANNEXE D : ANNEXE BDTLN DU DAE .....	86
<b>Portfolio LIFAT .....</b>	<b>88</b>
<b>Portfolio BdTln .....</b>	<b>91</b>
<b>Portfolio RFAI .....</b>	<b>96</b>
<b>Portfolio ROOT .....</b>	<b>101</b>

## 1. Présentation générale du LIFAT

[Le LIFAT, Laboratoire d'Informatique Fondamentale et Appliqué de Tours \(EA 6300\)](#) est le seul laboratoire d'informatique de l'**Université de Tours**. Les enseignants-chercheurs recrutés en 27ème section à l'Université de Tours sont tous rattachés au LIFAT. Le LIFAT a également comme co-tutelle l'**INSA Centre-Val de Loire** accueillant un nombre plus réduit de membres (5 MCF dont 1 HDR).

L'unité comporte au total **environ 90 chercheurs (dont 48 E/C dont 2 renouvellements en cours)** répartis dans **3 groupes de recherche** (fig 0.1 et 0.2) : Bases de données et traitement du langage naturel (BDTLN), Recherche opérationnelle, ordonnancement et transport (ROOT – ERL CNRS 7002), Reconnaissance de formes et analyse d'images (RFAI). Une spécificité importante du LIFAT est de comporter parmi ses équipes une **ERL CNRS (équipe ROOT)**.

Les recherches menées actuellement au LIFAT portent sur **plusieurs thématiques scientifiques dont les principales sont** l'analyse d'images et de vidéos, l'apprentissage machine, les systèmes dynamiques et interactifs de reconnaissance des formes, les représentations à base de graphes, la fouille et l'analyse exploratoire de données et l'évaluation de la qualité, les aspects sémantiques des données et des services, le traitement des langues naturelles, les nouveaux algorithmes de résolution de problèmes d'optimisation combinatoire et de recherche opérationnelle : complexité, méthodes arborescentes, programmation mathématique, algorithmes de graphes, métaheuristiques/matheuristiques, approximations... **Une synthèse** plus détaillée des thématiques et domaines de recherche des équipes du LIFAT est fournie en section 1.2.1.

De manière générale, nos objectifs scientifiques sont de concevoir et développer des modèles, des méthodes et des algorithmes et mettre à disposition des ressources et des logiciels, pour extraire de l'information, tirer des connaissances à partir de données, en intégrant la problématique de l'interaction homme-machine, et également de créer des algorithmes performants pour résoudre des problèmes d'optimisation, de nature combinatoire, avec la volonté d'obtenir dans un temps de calcul raisonnable, des résultats de très bonne qualité. L'un des **points remarquables du LIFAT** réside dans son **très large réseau de collaborations** régionales, nationales et internationales **avec de nombreuses autres disciplines** et équipes/unités. Ce résultat est le fruit d'une démarche réfléchie guidant nos priorités de recherche vers des **travaux trans- et pluridisciplinaires**.

Ce choix permet aux membres du LIFAT d'**avoir accès à des données, à des expertises et à des problématiques originales** (et difficiles !) ainsi qu'à des financements qui resteraient inaccessibles si le LIFAT restait cloisonné dans sa discipline.

L'**impact scientifique et socio-économique du LIFAT** en est aussi bénéficiaire avec des **contributions significatives en santé, patrimoines, linguistiques...** qui seront mentionnées tout au long de ce document.

En raison des spécificités régionales, les domaines d'application privilégiés du LIFAT sont les "arts et sciences humaines numériques" et "la santé et le handicap".

En 2016, le LIFAT s'est associé avec le Laboratoire d'informatique fondamentale d'Orléans (LIFO) pour créer la **fédération ICVL (FED 4279) qui regroupe les chercheurs en Informatique de la Région CVL**.

Un **Centre d'Expertise et de Transfert Universitaire de l'Université de Tours (CETU ILIAD3)** dédié à la valorisation de la Recherche dans le domaine de l'ingénierie logicielle en traitement d'images et numérisation 3D a été créé récemment, et est fortement adossé au laboratoire.

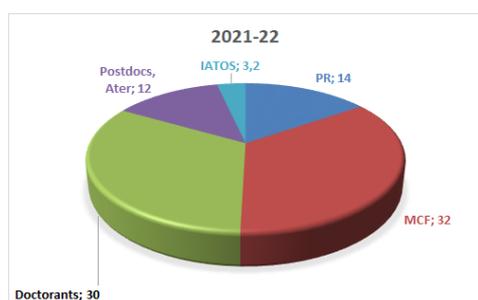


Fig. 0.1 : Répartition des effectifs LIFAT en Déc 2021 (2 renouvellements EC en cours non comptabilisés ici)

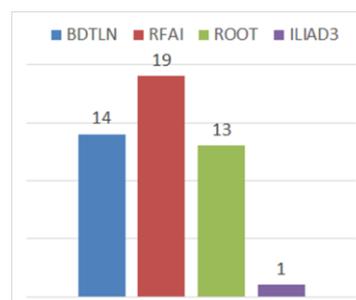


Fig. 0.2 : Répartition des effectifs permanents (EC) LIFAT + ILIAD3 en Déc 2021 (2 renouvellements EC en cours non comptabilisés ici)

### 1.1. Historique

Les faits les plus marquants dans l'histoire du Laboratoire LIFAT sont les suivants :

<b>1970</b>	moins de 10 membres, 4 permanents	création du Laboratoire d'Informatique, la même année que la création de l'Université de Tours. Deux équipes : OC (ordonnancement et conduite) et RFAI (reconnaissance des formes et analyse d'images).
<b>1996</b>	26 membres, 14 permanents E/C	Reconnaissance du Laboratoire d'Informatique par le Ministère comme Equipe d'Accueil, EA 2101.

<b>20XX</b>		Augmentation progressive de sa taille pour atteindre 60 membres répartis dans 5 équipes de taille très variables
<b>2012</b>	80 membres, 40 permanents E/C	Association avec le CNRS : l'équipe « Ordonnancement et Conduite » devient ERL-CNRS 6305, et le Laboratoire d'Informatique est l'EA 6300.
<b>2014</b>	90 membres, 46 permanents E/C	L'INSA Centre Val de Loire devient tutelle du laboratoire
<b>2014</b>		Démarrage du projet de fédération informatique avec le LIFO
<b>2016</b>	80 membres, 46 permanents E/C	Restructuration du LIFAT (entre 2013 et 2016) : Fusion des petites équipes avec RFAI - L'équipe OC devient l'équipe ROOT ERL CNRS 7002
<b>2016</b>		Création de la <a href="#">fédération ICVL</a> (Informatique Centre Val de Loire regroupant LIFAT et LIFO)
<b>2017</b>		Evaluation HCERES - Création du <a href="#">CETU ILIAD3</a> (centre d'expertise et de transfert universitaire)
<b>2021</b>	90 membres, 48 permanents E/C	Préparation évaluation HCERES en 2022

## 1.2. Localisation

Le LIFAT se situe principalement sur deux sites géographiques :

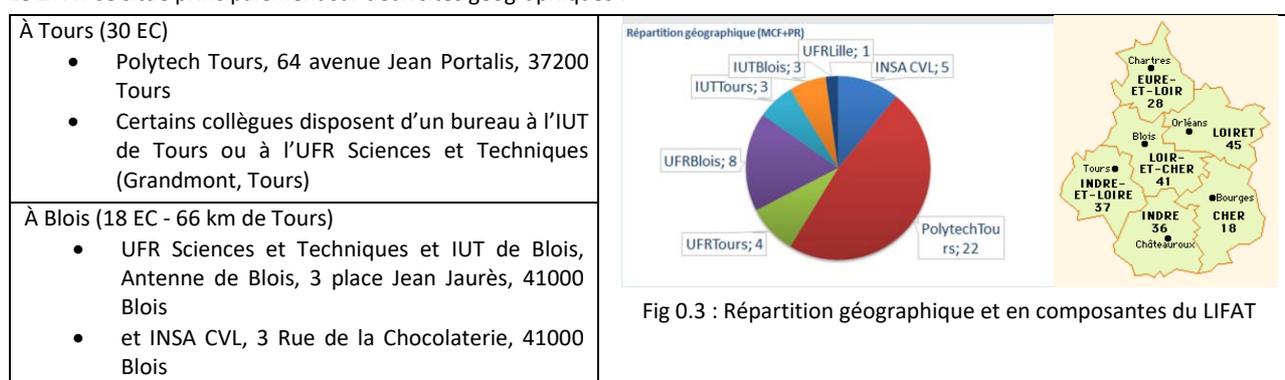


Fig 0.3 : Répartition géographique et en composantes du LIFAT

Comme le montre la figure 0.3, les membres du LIFAT sont répartis dans différentes composantes de l'Université de Tours (ou INSA CVL) situées à Tours ou Blois avec **une moitié des membres rattachés à Polytech Tours**. Le LIFAT (secrétariat, antenne financière) est hébergé au sein de Polytech Tours. Des bureaux pour l'accueil des E/C des autres composantes/établissements sont mis à disposition à Polytech Tours.

## 1.3. Structuration

Lors du précédent contrat (2011-2016), une restructuration du LIFAT a abouti à une structuration en trois équipes internes : BdTIn, RFAI et ROOT ; l'équipe ROOT ayant obtenu le label ERL CNRS 7002.

- **BdTIn (Bases de données et traitement des langues naturelles, 14 EC, 16 doctorants)** : les recherches portent sur la représentation de données, connaissances et services, et l'extraction de connaissances par des solutions centrées utilisateur. Elles se déclinent selon (1) la structure des données, (2) le type de traitement et (3) l'application visés. Côté structure de données, cela va des ressources en langues naturelles aux bases de données relationnelles et multidimensionnelles, en passant par les formats spécialisés des ressources spatio-temporelles, semi-structurées des services et données du Web et de la représentation des connaissances. Concernant les traitements, ils visent d'une part la constitution de ressources et connaissances respectant des critères de qualité et d'autre part des méthodes et algorithmes d'analyse et extraction de connaissances. Les applications quant à elles sont tournées vers la société : environnement, emploi, journalisme, patrimoine culturel, santé et handicap. Les travaux de l'axe Exploration et analyse interactive de données s'illustrent dans l'extraction interactive de motifs, la proposition de langages pour l'exploration de données, la génération de recommandations et la définition de mesures. Ceux de l'axe Traitement automatique des langues et interactions sont connus pour produire des applications et des ressources linguistiques diverses, avec de fortes collaborations internationales, pour la détection d'éléments d'intérêt (entités nommées ou polylexicales) et la mise en relation des mots d'intérêt (coréférence, relations temporelles et prédicats). Une autre contribution de cet axe porte sur un système d'aide à la saisie de textes pour des personnes en situation de handicap. L'axe Intelligence des données et services travaille au développement de solutions s'appuyant sur les standards du Web pour l'interopérabilité sémantique, le calcul de similarités sémantiques et la découverte, la composition et le déploiement de services et d'objets. La fouille du Web des données, menée avec l'axe Exploration au cours du contrat, a produit des résultats publiés dans les meilleures conférences du Web sémantique.
- **RFAI (Reconnaissance des formes et analyse d'images, 19 EC, 6 doctorants)** : Les domaines d'intérêt et de compétences de l'équipe RFAI relèvent de l'analyse d'images, de la visualisation de données et de la reconnaissance des formes. Aussi bien la visualisation de données que la reconnaissance des formes sont des traitements de haut niveau nécessitant

d'effectuer (au préalable) des traitements bas niveau pour adapter ou préparer les données initiales. L'originalité des travaux de l'équipe provient notamment du fait que nous considérons que l'ensemble de ces étapes sont très interdépendantes, particulièrement lorsqu'il s'agit d'analyser des flux de données hétérogènes ou multimodaux (données temporelles, 3D, vidéos) pour détecter ou reconnaître des formes ou motifs singuliers soit automatiquement (reconnaissance des formes) soit interactivement (fouille visuelle). Les deux domaines applicatifs privilégiés de l'équipe concernent actuellement les données et images biomédicales - Santé et Handicap - ainsi que les données du patrimoine (images de documents, de bâtiments anciens ou de vestiges, de sculptures, numérisés) - Arts et Humanités -. L'équipe RFAI est parmi les rares équipes en France à travailler depuis plusieurs années à la mise en place de méthodes interactives d'analyse d'images, de reconnaissance des formes et de fouille visuelle de données. Avec au moins quatre permanents sur le domaine, l'équipe RFAI est l'une des équipes comportant le plus de personnes à travailler dans le domaine « Graph Matching et Graph Neural Networks » en France et à l'international. Également, les collaborations avec ROOT ont fait émerger des travaux singuliers autour de la combinaison entre méthodes d'optimisations et apprentissage machine. Plusieurs contributions significatives ont été produites par l'équipe dans ces différents domaines durant le contrat

- **ROOT (ERL CNRS Recherche Opérationnelle, Ordonnancement et Transport, 13 EC, 8 doctorants)** : Les activités scientifiques de l'équipe s'inscrivent dans le domaine de la Recherche Opérationnelle avec des contributions à la résolution de modèles génériques d'ordonnancement et/ou de transport, ainsi que des contributions au champ des méthodes. Les recherches menées au sein de l'équipe se sont inscrites aussi bien dans un cadre académique que dans le cadre de la résolution de problèmes plus appliqués issus de défis sociétaux dans le milieu de la santé, la production industrielle, la mobilité et le développement durable et les systèmes de calcul haute performance. L'équipe a été associée au CNRS (Institut des Sciences de l'Information et de leurs Interactions, INS2I) en tant qu'Equipe de Recherche Labellisée (ERL OC 6305) en 2012 et a été de nouveau associée en 2016 en tant que ERL ROOT 7002.

## 1.4. Thématiques scientifiques

Trois thèmes transversaux avaient été identifiés en 2016 pour inciter aux collaborations inter-équipes et focaliser les activités du LIFAT sur des domaines bien ancrés dans le paysage régional. Ils correspondaient à des **domaines d'application** dans lesquels toutes les équipes étaient actives et sources potentielles de collaborations et financements pour le LIFAT.

- **Thème « Santé et handicap »**

Toutes les équipes du LIFAT ont des collaborations fortes avec le CHRU de Tours ou d'autres organisations dans le domaine de la santé et du handicap. Ce thème fédérateur existe depuis de nombreuses années. Le domaine de la santé ouvre des champs d'études scientifiques et d'applications très vastes en informatique et en Intelligence artificielle. D'une part, les problématiques d'extraction de connaissances, aide aux diagnostics, de prédictions, d'organisations sont bien identifiées et d'une importance indéniable pour contribuer au bien-être voire à l'amélioration de la prise en charge des patients. D'autre part l'optimisation de la prise en charge des patients et de l'utilisation des dispositifs médicaux lourds devient aussi cruciale.

Sans pouvoir être exhaustif, voici les projets majeurs auxquels le LIFAT a contribué ou contribue actuellement :

- Une collaboration de longue date existe entre le LIFAT (équipe RFAI + [CETU ILIAD3](#)), l'équipe [INSERM iBrain](#) et l'équipe PRC de l'UMR INRAE-UT-CNRS-IFECE CNRS située à 25 km de Tours concernant la mise en place de techniques d'intelligence artificielle (Analyse d'images IRM 3D, fouille visuelle 3D, apprentissage sur graphes) pour la mise en place de méthodes et modèles pour l'analyse de cerveaux humains et animaux (neuro-anatomie et neuro-imagerie). Plusieurs thèses et projets ANR (Fibratlas I et II) et APR (NeuroGeo, Neuro2co, Fibravasc...) se sont succédés et ont notamment permis la création de la [plateforme logicielle SILA3D](#) aujourd'hui opérationnelle et dont divers modes de valorisation/transfert sont en cours d'étude avec l'aide de C-VALO et des services Partenariat de l'université et du CHRU.
- Toujours avec l'équipe iBrain, le projet ANR SIRCUS vient d'être accepté en 2021 (malgré un très fort taux de sélectivité). Ce projet met en collaboration l'équipe iBrain, le LIFAT (RFAI) et le CHRU de Tours dans le cadre du [GIS Autisme](#) afin de faire progresser la compréhension actuelle des mécanismes physiopathologiques des Troubles du Spectre Autistique et de leur diagnostic précoce avec le but ultime de proposer un soutien et une thérapie précoce aux patients. Le projet vise à produire de nouvelles méthodes d'analyse des données d'eye tracking, grâce à la participation d'experts en intelligence artificielle, et devrait aboutir à une interface logicielle accessible à d'autres chercheurs et cliniciens.
- Le LIFAT (RFAI) est également impliqué dans un projet international en oncologie porté par l'équipe INSERM 1069 et financé par le Studium afin d'accueillir pendant 1 an à Tours, David Crottès, chercheur à Los Angeles. L'objet du projet, concernant le volet informatique, est de proposer un modèle numérique permettant de prédire, d'après des mesures faites par le biologiste et par apprentissage, dans quelle mesure des familles de cellules réagissent ou non à des molécules utilisées en chimiothérapie. Le résultat opérationnel vise à relier la réponse au calcium d'une famille de cellules (facile à mesurer) au type de cancer, et à la capacité d'une molécule à enrayer un cancer au sein de cette famille. Ceci dans un contexte de médecine personnalisée.
- L'équipe ROOT travaille de manière soutenue depuis plusieurs années sur la « Résolution de problèmes intégrés dans un contexte de production et livraison de chimiothérapies » en collaboration avec le CHRU de Tours et la société Eticsys. Ces travaux ont donné lieu au projet APR Distri-chimio et 1 thèse.
- De manière connexe, le projet ANR JJC PONoSAD (2013-2017) s'intéressant à la planification et l'optimisation pour les soins à domicile et porté par un membre de l'équipe ROOT a donné lieu à une thèse et un prototype logiciel de gestion et

- de planification de rendez-vous en temps réel pour les coordinateurs de soins mais aussi les intervenants à domicile via leur smartphone. Ce prototype (source et documentation) est disponible sur [le site web du projet \(en open source\)](#).
- Des travaux autour des problématiques de déploiement et redéploiement d'ambulances pour les systèmes pré-hospitaliers Nord-Américain ont été menés en collaboration avec des chercheurs Québécois (du CIRRELT) venus en visite à Tours (Angel Ruiz et Valérie Bélanger).
- Le projet APR SQUALD démarré en 2021 impliquant des membres de BDTLN et ROOT porte sur le suivi de qualité de vie de patients d'affections de longue durée. Ce projet, porté par un membre de BdTIn, vise à développer une approche d'accompagnement continu des patients hors murs en cohérence avec leur parcours de soins. Pour cela il s'agit d'associer de nombreux intervenants pour articuler leurs offres pour en offrir une vue unifiée au patient d'une part et d'autre part de coordonner les informations de suivi du patient pour le soulager dans la réalisation de ses soins de support.
- La thématique de l'aide aux personnes en situation de handicap a continué dans l'équipe BdTIn avec le projet PREDICT4ALL, financé par la fondation Paul Bennetot, sur le développement d'un module de prédiction de mots, en partant des propositions du système Sibylle développé antérieurement. Ces travaux ont mené également à la coordination d'un dépôt de projet ANR (AAC4All) qui a été accepté et débute en 2022.
- L'équipe a également mené des travaux en collaboration avec Hélène Blasco du CHRU, dans le cadre du groupe HumanInTheLoop du RTR DIAMS (financement de 2 stages de Master 2, l'un au LIFAT et l'autre au LIFO), sur des analyses de données de patients atteints de sclérose latérale amyotrophique, en vue d'en modéliser les évolutions. Ces travaux ont donné lieu en 2022 au dépôt d'un projet région et à un partenariat étendu (CHRU, BDTLN et RFAI au LIFAT, LIFO, ATOS, ainsi que d'autres acteurs régionaux).

Les membres du LIFAT contribuent également à l'animation scientifique locale (participation au conseil d'administration de la [SFR FED Neuro-imagerie fonctionnelle](#) regroupant des unités localisées sur Tours, Orléans et Poitiers, participation au comité scientifique de la plateforme d'imagerie PIXANIM) et nationales (création en 2014 du groupe de travail ROSA-Recherche Opérationnelle et Santé du GDR RO).

#### • **Thème « Arts et Sciences Humaines »**

L'environnement régional (châteaux de la Loire, patrimoine riche, vignes et vins, ...) a naturellement conduit à proposer ce thème comme axe fédérateur pour le LIFAT. [Le projet Intelligence des Patrimoines \(iPat\)](#) porté par le Centre d'Etudes Supérieures de la Renaissance (CESR UMR-CNRS 7323) a été, il y a plusieurs années, le résultat d'une démarche de promotion du territoire (patrimoines culturels et naturels, gastronomiques, artisanaux, agricoles, etc.) alliant la recherche scientifique et le monde socio-économique de la Région Centre-Val de Loire. iPat regroupait la plupart des organismes de recherche et des établissements d'enseignement supérieur (U. de Tours, U. d'Orléans, CNRS, INSA Centre-Val de Loire, BRGM, INRAE, ESCEM) de la région CVL. En 2016, iPat a donné lieu à un dépôt de projet ARD (Ambition Recherche Développement) financé par la région sur un thème devenu un de ses domaines de spécialisation intelligente. Dans ce cadre, de nombreux projets de recherche rattachés à iPat ont été soutenus par la Région.

Dans ce contexte, le LIFAT a contribué aux thématiques suivantes :

- De plus en plus de reconstitutions 3D de tombeaux, statues et reliefs sont présentées dans des expositions en réalité virtuelle exploitant les dernières avancées en matière de visualisation 3D et d'immersion des spectateurs. Les [projets Valmod, Visit et Sculpture 3D se situent dans ce contexte et exploitent les](#) compétences conjointes du LIFAT et du CETU ILIAD3 en fouille de données et visualisation 3D pour permettre aux historiens du CESR et aux responsables de lieux touristiques de s'approprier de manière autonome ces nouvelles technologies 3D.
- L'expertise du LIFAT dans le domaine de l'analyse et de l'indexation de documents anciens est également mise à contribution pour différents projets d'aide à la transcription d'ouvrages anciens ([projet BVH](#) du CESR, dépôt ANR JCJC en cours), de projets de détection d'objets d'intérêt (archéologiques) à partir de données LIDAR ou d'étude de l'évolution des territoires par vectorisation de plans cadastraux (projets VECCAR et SOLIDAR en collaborations avec le laboratoire CITERES).
- Le projet APR [SmartLoire](#) (2017-2020 + nouveau dépôt en cours) porté par l'équipe BDTLN et impliquant l'équipe ROOT traite de l'organisation de séjours de tourisme culturel en Région Centre-Val de Loire par composition de services Web, ce qui amène les participants côté BdTIn à collaborer avec des historiens de l'art, des géographes et des structures liées au tourisme et patrimoine. L'équipe ROOT a été responsable d'un lot traitant de la conception d'itinéraires touristiques personnalisés et optimisés (recrutement d'un postdoc de 9 mois, 4 publications en conférences). Côté BdTIn ce projet a financé deux thèses en co-tutelles, en cours, et une thèse soutenue en 2020 l'a pris comme domaine d'application d'un système de mashup intelligent de services et données pour la recommandation de points d'intérêt touristiques et patrimoniaux.
- Initié au contrat précédent, l'investissement de l'équipe BdTIn dans des collaborations sur la question de l'interopérabilité sémantique des ressources numériques élaborées par nos collègues en histoire de l'art, en archéologie ou en anthropologie s'est prolongé et a pris de l'ampleur, à travers un projet régional (postdoc d'un an, création d'un outil d'alignement de vocabulaire, une publication en conférence) et un projet ANR commun entre CITERES-LAT et le LIFAT (recrutement d'un ingénieur et d'un postdoc d'un an, réalisation d'outils d'exploitation de graphes de connaissances et d'ontologies, 4 publications en conférences et 1 en revue), puis amenant la participation du LIFAT à deux projets H2020, ainsi qu'à la création à la MSH Val de Loire d'un pôle Web sémantique au sein de son atelier numérique, création soutenue par la TGIR HumaNum. Ces développements sont également à l'origine de la création d'une série de workshops internationaux sur le Web sémantique pour le patrimoine culturel.
- Le LIFAT est associé au projet de la MSH pour le contrat à venir, en particulier dans son projet d'axe « Modèles, modélisation, simulation » associant divers laboratoires de SHS, ce qui ouvre a priori de nouvelles perspectives de collaborations régionales.

- Le LIFAT a accueilli une artiste Laurence Dreano en son sein pendant plusieurs mois en 2019 avec l'aide du CETU ILIAD3 dans le cadre de l'action "Résidence d'artiste" organisée par l'université de Tours. L'objectif était d'introduire, de combiner [numérique et virtuelle au sein de sculptures](#).

- **Thème « Big data et Calcul Haute Performance »**

Ce thème faisait partie des axes mis en avant en 2016 notamment du fait de l'existence de la fédération régionale [CaSciModOT](#). Le LIFAT est impliqué dans cette structure d'échange autour de la simulation, de la modélisation, de l'algorithmique, et du traitement de données massives. CaSciModOT est maintenant un cluster de la Région Centre avec un comité de pilotage constitué de représentants de laboratoire. Le LIFAT, via son représentant a co-organisé la 34e journée du projet CaSciModOT à Tours en 2021.

Quelques travaux ont été réalisés dans cet axe concernant des approches de résolution de problèmes d'ordonnement multi-agents dans des systèmes distribués (1 thèse, quelques publications en conférences) et la mise en place de méthodes d'optimisation de la gestion des ressources sur une plate-forme de type « Big Data » basée sur le logiciel Hadoop (1 thèse CIFRE avec Cyres Groupe, quelques publications en conférences).

Durant le contrat, l'« Apprentissage et le Deep Learning » a constitué une évolution devenue incontournable et génératrice de nouveaux axes fédérateurs et de spécialisation. Le « **virage du Deep Learning** » (que l'on a vécu durant ce contrat) a été saisi comme une opportunité par le LIFAT compte tenu des domaines de spécialisation historiques du LIFAT (Traitement Automatique du Langage, Machine Learning, Computer Vision, Recherche Opérationnelle, Data-Mining). Ce qui aurait pu être un bouleversement apparaît plutôt comme une opportunité ayant permis d'augmenter encore les projets et collaborations inter-équipes. Deux **nouveaux axes de spécialisation** pour le laboratoire ont ainsi émergé durant ce contrat quinquennal notamment grâce à l'intensification des **financements de thèses inter-équipes** mis en place par l'équipe de direction :

- **Combinaison Apprentissage Machine (AM) et Recherche Opérationnelle (RO)** : En RO, les problèmes d'optimisation doivent être formalisés et bien structurés. Un algorithme d'optimisation exploite ces informations structurées pour les résoudre. Au contraire, l'Apprentissage Machine (AM) est un domaine de recherche traitant de la conception d'algorithmes permettant de résoudre des problèmes au moyen d'approches statistiques. Les algorithmes d'AM extraient des informations à partir d'exemples pour faire des prédictions, ce qui constitue une différence majeure par rapport aux algorithmes de RO. Le cœur des algorithmes d'apprentissage réside dans leur capacité à apprendre et à généraliser à partir d'informations "non structurées" ou "non formalisées". Dans ce contexte des travaux récents du LIFAT concernent, d'une part, l'AM pour améliorer les algorithmes issus de la RO pour la résolution de problèmes combinatoires, et d'autre part, la RO pour améliorer les modèles d'AM dans le domaine de la vision par ordinateur. Cette dualité forme un bouclage de pertinence qui revêt un caractère novateur. Un doctorant et 4 permanents travaillent sur cette thématique. 2 projets ANR portant sur ces sujets ont été soumis en collaboration avec Politecnico de Turin et le CIRRELT au Canada mais ils n'ont malheureusement pas été retenus. Les verrous attaqués dans l'axe RO pour l'AM sont par exemple l'intégration d'algorithmes de RO dans les architectures de deep learning, posant la question épineuse de la différentiation des fonctions discrètes. Dans le second axe « AM pour la RO », le verrou visé est l'utilisation d'un prédicteur statistique pour guider l'exploration d'une méthode de résolution d'un problème combinatoire. A ce jour, ces travaux ont donné lieu à 2 publications dans des conférences francophones (ROADEF). Plus largement, les collaborations d'ordre similaire entre ROOT et RFAI ont donné lieu à plusieurs co-encadrements de thèses (Darwiche, Lacharme), soumissions de projets (ANR LOR) et publications (voir publications inter-équipes RFAI-ROOT).
- **Vers plus de sémantique et d'explicabilité dans les systèmes d'aide à la décision** (ontologies, ex-AM, Human in the loop). Le LIFAT et la MSH CVL se sont fortement rapprochés durant ce contrat. Les échanges ont abouti à la mise en place d'un axe « Sciences des données et interaction humaines » dans le projet MSH-CVL pour le prochain contrat, ce qui ouvre a priori de nouvelles perspectives de collaborations régionales. A ceci s'ajoute un groupe de travail « *human in the loop* » de la fédération ICVL ayant pour objectif de créer une synergie LIFO-LIFAT autour de cette thématique. Des collaborations émergentes entre l'équipe BDTLN et RFAI dans le cadre de l'ARD JUNON (projet Ambition Recherche Développement porté par le BRGM et largement financé par la région pendant 5 ans (2021-2026) sur un thème devenant un de ses domaines de spécialisation intelligente) visent aussi à produire des systèmes de prédiction et de simulation numériques (jumeaux numériques) facilement utilisables et appréhendables par des usagers non experts en sciences des données ou apprentissage machine. Ces nouveaux travaux se sont traduits et se traduisent par :
  - la multiplication des collaborations inter-équipes : 4 thèses en co-encadrement BDTLN—RFAI et 2 projets régionaux. Ces premières collaborations ont abouti à 4 publications communes en conférences internationales et nationales.
  - la multiplication des collaborations pluridisciplinaires concrétisées par des financements obtenus au niveau régional, national et européens déposés avec des partenaires du secteur SHS ou de la santé. Ces projets ont donné lieu à des publications communes dans des revues et dans des conférences internationales et nationales.
  - Le montage de l'action HELP du GdR Madics qui vise à créer une communauté nationale autour des questions de l'explicabilité, avec la particularité de proposer un regard croisé informatique / sciences humaines et cognitives. L'action HELP s'intéresse notamment à l'explication de pipeline de fouille de données de bout-en-bout, en ne se focalisant pas uniquement sur les modèles prédictifs mais en identifiant les verrous liés à l'explicabilité dès le prétraitement des données jusqu'à l'interaction avec l'utilisateur final.

De par ces actions et résultats obtenus, **la pluridisciplinarité et la prise en compte des dimensions humaines dans la mise en place des systèmes d'aide à la décision devient un axe de spécialisation du LIFAT que nous pensons mettre encore plus en avant dans l'avenir.**

## 2. Profil d'activités

Notre **mission** en tant qu'enseignants-chercheurs est multiple : produire de nouvelles connaissances, enseigner, diffuser les connaissances en liaison avec l'environnement économique, social et culturel, et contribuer à la coopération internationale. Pour remplir cette mission, nous nous fixons comme priorités la formation des doctorants, la participation à des projets de grande ampleur, la publication dans les meilleures revues ou conférences de nos domaines et l'ouverture à l'international. De plus, nous inscrivons nos actions dans des réponses aux défis sociétaux, identifiés dans le cadre du Programme européen pour la recherche et l'innovation, et dans la Stratégie Nationale de la Recherche (SNR). Nous précisons, dans le tableau ci-dessous, la répartition globale moyenne en pourcentage (**pour un total en ligne de 100 % donc uniquement sur la part d'activité « recherche »**) de l'activité des membres permanents (EC) de l'unité de recherche, l'unité de décompte étant le temps consacré aux missions suivantes :

Activités VS % 0 à 100 sur ces 7 items	BDTLN	RFAI	ROOT	LIFAT (moyenne)
Administration de la recherche (responsabilité de pilotage de la recherche (VP, Direction d'Institut, DAS, ...) participation à des instances d'évaluation (CNU, CoNRS, CSS...), responsable de volet IdEx, direction de projets -ANR, Horizon Europe, ERC, CPER-, responsabilités éditoriales dans des revues ou collections nationales et internationales)	23 [0-46]	20 [0-60]	15 [0-30]	<b>19 [0-60]</b>
Dissémination de la recherche (partage de connaissances avec le grand public, médiation scientifique, interface science/société)	3 [0-5]	3 [0-10]	4 [0-5]	<b>3.3 [0-10]</b>
Encadrement de la recherche (implication au niveau D ainsi que dans le suivi de projets post-doctoraux)	29 [15-45]	27 [5-60]	30 [15-40]	<b>28.7 [5-60]</b>
Contribution à l'adossement d'enseignements innovants à la recherche (EUR, SFRI, etc.)	2 [0-10]	3 [0-5]	5 [0-10]	<b>3.3 [0-10]</b>
Expertise technique (pouvoirs publics aux niveaux national et régional, entreprises, instances internationales (FAO, OMS, ...))	3 [0-10]	7 [0-25]	8 [0-10]	<b>6 [0-25]</b>
Recherche académique	36 [3-60]	30 [10-50]	27 [10-50]	<b>31 [3-60]</b>
Valorisation, transfert, innovation	4 [0-10]	10 [0-45]	11 [10-25]	<b>8.3 [0-45]</b>

La plupart des doctorants déclarent un temps plein en recherche académique (leurs réponses n'ont pas été intégrées dans les statistiques fournies (concernant les permanents seulement). Plusieurs personnes déclarent, dans chaque équipe, entre 30% et 60% en administration de la recherche (direction de laboratoire, ...) ainsi 20% du temps est globalement consacré à l'administration de la recherche au sein du LIFAT. On note des écarts importants sur de nombreuses lignes soulignant la diversité des profils entre membres du LIFAT, ou du moins la diversité de perception de ces profils (tout le monde n'ayant pas la même interprétation).

Bien que conscient de l'impact sur leurs activités scientifiques, les membres du LIFAT n'ont pas pu s'empêcher d'intensifier encore plus qu'avant leur implication dans des responsabilités collectives et administratives durant ce contrat. Cette surcharge de travail est bien souvent compensée par les collègues en réalisant un nombre d'heures déraisonnable de travail. Cet état de fait est la seule solution possible pour maintenir le niveau d'excellence du LIFAT au fil des contrats.

La crise sanitaire n'a pas simplifié le contexte durant les dernières années et amène à se questionner sur **le temps durant lequel une telle situation pourra être supportée par les collègues** (Voir section XXX, l'augmentation significative des demandes de départ en disponibilité – voir aussi les statistiques sur l'intérêt des jeunes pour commencer une carrière dans l'enseignement supérieur).

Ci-dessous, une liste non exhaustive des charges lourdes assurées par des membres du LIFAT durant le contrat :

- Présidence de l'Université de Tours (1 an et demi)
- Vice-présidence chargée du partenariat et de l'innovations de l'université de Tours (2 ans),
- Vice-présidence chargée du numérique de l'université de Tours (4 ans)
- Direction de Polytech Tours et vice-direction chargée des relations entreprises (6 ans).
- Direction de l'école doctorale MIPTIS Orléans-Tours (6 ans)
- Direction du département Formation par Alternance de Polytech Tours (6 ans),
- Plusieurs directions de formations et directions des études « lourdes » (6 ans).
- Co-direction (avec une ou un membre du LIFO) de la Fédération ICVL (depuis sa création, et avant : portage du projet de création)
- ...

### 3. Environnement de recherche

#### 3.1. Implication dans les structures fédératives régionales

Durant ce contrat, l'explosion du domaine de l'intelligence artificielle a bouleversé la recherche en informatique et a donc impacté les activités du LIFAT. Néanmoins, le LIFAT était tout à fait prêt, voire précurseur pour ce qui concerne les activités associées à l'exploitation de technique d'apprentissage automatique (machine Learning) dans divers domaines des sciences des données comme l'analyse d'images (équipe RFAI = couplage CV et ML), le traitement des langues et des données (BDTLN = couplage TAL, BIGDATA et ML) et l'optimisation (ROOT = couplage RO et ML).

Ainsi le LIFAT a été instigateur ou fortement impliqué dans plusieurs actions de structuration de la recherche régionale durant ce contrat quinquennal :

- La mise en place de la [fédération ICVL](#) (LIFAT/LIFO) a été terminée avec succès en 2017
- Le [RTR DIAMS](#) est un réseau de chercheurs créé en 2019 (financé pour 4 ans) en Région Centre-Val-de-Loire qui a pour ambition de structurer l'ensemble de la **communauté scientifique régionale concernée par les thématiques "numérique et données"**, de favoriser ainsi le montage de projets de recherche, de faire émerger de cette structure les thématiques les plus dynamiques, matures et prometteuses, pour le montage de projets internationaux. Le [RTR DIAMS](#) a été initié par la fédération ICVL et est depuis dirigé par un binôme EC LIFO + EC LIFAT.
- Le **CETU ILIAD3** a été mis en place en 2017 afin de faciliter le transfert des recherches (prototypage) et des compétences (formations, expertises) du LIFAT vers le monde socio-économique (voir Portfolio).
- Comme déjà mentionné, le LIFAT et la **MSH CVL** se sont rapprochés pour mettre en place un projet d'envergure autour de **l'IA et l'Humain** concrétisé au sein du projet Valopat financé dans le cadre du CPER prévoyant la mise en place d'un environnement immersif 3D (salle + casques de réalité augmentée) devant permettre d'innover dans le domaine de la fouille visuelle et interactive de données (125 K€)
- Le LIFAT est impliqué dans les **conseils scientifiques ou organes directeurs** des structures régionales suivantes : cluster Cascimodot, plateforme d'imagerie [PIXANIM](#), [SFR FED Neuro-imagerie](#), le COPIL [IA-Loire Valley](#), ...
- Enfin, notons que les activités du LIFAT s'inscrivent pleinement dans le **projet LOIRE-VAL-HEALTH** actuellement en cours de dépôt par **l'Université de Tours pour le PIA4** (cf WP2 sur la création de modèles humains et animaux en imagerie médicale et le WPX sur les aspects fouille de données médicales).

#### 3.2. Implication dans les formations à la recherche

Le LIFAT est rattaché à l'Ecole Doctorale 551 MIPTIS « Mathématiques, Informatique, Physique théorique, Ingénierie des systèmes » commune à Orléans et à Tours. Quatre unités de recherche sont dans cette ED : les deux laboratoires d'informatique (LIFAT et LIFO), le laboratoire de Mathématique (UMR Institut Denis Poisson Orléans-Tours) et des équipes de l'institut PRISME d'Orléans. Cette ED a un directeur et un directeur adjoint, l'un à Tours, l'autre à Orléans. Le Pr. H. Cardot, membre du LIFAT a assuré la direction de cette ED durant tout le contrat quinquennal.

Les membres du LIFAT interviennent dans plusieurs formations d'ingénieurs, à l'Université de Tours et à l'INSA Centre Val de Loire à Blois. Il s'agit de formations d'ingénieurs Bac+5, où les étudiants sont recrutés à Bac+3 ou post-bac via le Parcours des écoles d'ingénieurs Polytech pour [Polytech Tours](#), et post-bac pour [l'INSA CVL](#) :

- la spécialité « Informatique » du diplôme de Polytech Tours. Un **parcours de spécialisation en Intelligence Artificielle** a été mis en place (couvrant plus de 200h d'enseignement/ étudiant).
- la spécialité « Informatique industrielle » (en apprentissage) de Polytech Tours
- les spécialités « Sciences et Technologies Pour l'Ingénieur » et « Génie des systèmes Industriels » de l'INSA CVL

La spécialité d'ingénieur en Informatique de Polytech Tours inclut dans le parcours de formation de ses élèves ingénieur **un projet Recherche et Développement** qui se déroule deux jours par semaine lors de la dernière année de formation (70 élèves ingénieurs par année en moyenne). La très grande majorité des sujets proposés sont en relation directe avec les activités des enseignants-chercheurs de Polytech Tours et du LIFAT. Ce travail constitue une réelle initiation à la recherche en lien avec les activités du laboratoire. **Ce travail de recherche permet aux élèves-ingénieur de postuler sur des offres de thèses au LIFAT** ou dans des laboratoires hors université de Tours, faisant valoir outre leur grade de master, une première expérience de recherche. Ces projets sont également l'occasion d'initier des projets communs avec le CESR de Tours en faisant collaborer chaque année en binôme un étudiant du Master Humanités Numériques et 1 étudiant en informatique.

Dans ce cadre, **les doctorants du LIFAT ont l'opportunité d'encadrer** un à deux élèves-ingénieurs par an sur des thématiques en relation directe avec leur thèse, leur permettant d'avoir une première expérience d'encadrement de niveau Master2. Notons que régulièrement les élèves-ingénieurs sont associés à des communications en collaboration avec l'encadrant de ce travail, qu'il soit doctorant ou enseignant-chercheur.

Les membres du LIFAT ont défini et assurent également le parcours Big Data Management and Analytics (BDMA) de la mention Informatique du Master « Science, Technologies, Santé » de l'université de Tours. Le Master BDMA est à la Faculté des Sciences et Techniques à Blois. Faisant suite au Master Erasmus Mundus IT4BI et lui-même Master Erasmus Mundus de 2017 à 2018, son

programme recouvre les aspects du Big Data et de l'informatique décisionnelle et est couplé à une approche éthique indispensable à la formation de professionnels et de chercheurs. Il forme aux entrepôts, analyse de données, statistiques décisionnelles et recherche opérationnelle, à la fouille de données avancée, aux principes et méthodes issues du Web (Cloud computing, services Web, données et connaissances du Web), au traitement des images, des langues et à l'ingénierie des connaissances. Le M2 comporte un programme de séminaires initiant à la recherche et un projet transverse dans lequel cette initiation est appliquée. De septembre à février, 2 jours par semaine y sont consacrés. Les diplômés occupent des fonctions d'ingénieurs en informatique et un nombre non négligeable d'entre eux s'engagent en thèse (12 diplômés du Master de Blois ont été ou sont doctorants au LIFAT).

Des membres du LIFAT ont défini les contenus des parcours IDCP et MNCP de la mention Humanités Numériques du Master « Sciences Humaines et Sociales » de l'université de Tours et d'autres interviennent dans ses enseignements depuis sa création en 2018, ou participent à son comité de pilotage. Un poste de MCF 27 (enseignements au CESR – recherche au LIFAT) a été demandé sans succès depuis la mise en route de ce master pour absorber la surcharge de travail (administrative et pédagogique) engendrée.

Notons également, que le LIFAT a organisé et assuré la mise en place de **2 enseignements optionnels de la faculté de Médecine** d'introduction à l'IA accueillant pendant 2 x 20h une quinzaine d'étudiant en formation de médecine ou de pharmacie.

Toutes ces formations d'ingénieurs et de Master sont conçues et administrées par des enseignants-chercheurs membres du LIFAT. A l'INSA-CVL, ce sont les membres du LIFAT qui gèrent la filière informatique (filiale transverse sur les cinq années).

L'offre de formation Master en Informatique de l'Université de Tours est donc riche. La contrepartie est un certain manque de lisibilité. Une réflexion s'est engagée pour que l'offre de formation soit plus lisible. De nouvelles maquettes sont en cours de préparation.

### 3.3. Accueil de stagiaires de Master

Le LIFAT accueille chaque année des étudiants en stage. Ces étudiants reçoivent systématiquement une gratification, conformément à la législation si le stage dépasse deux mois. Ces indemnités (environ 5 par an) proviennent souvent de financements obtenus en répondant aux **appels des structures régionales (ICVL, DIAMS, Cascimodot) et nécessitant une collaboration** avec d'autres organismes de recherche régionaux. D'autres peuvent provenir de projets financés (type ANR, contrat industriel, projet Région), ou être pris sur les budgets de l'équipe concernée. Le nombre de stagiaires de Masters accueillis au **LIFAT varie donc entre 8 et 23 stagiaires (M1+M2) accueillis chaque année avec une moyenne de 12 par an.**

### 3.4. Production de documents pédagogiques

Les enseignements de Master ou de formation d'ingénieurs donnent lieu à la création de documents pédagogiques, déposés sur une plateforme de cours en ligne (CELENE).

Des ressources pédagogiques associées aux formations, options IA ouverts à d'autres composantes (CESR, Fac de Médecines, SFR Neuro-Imagerie, Journée INRAE, ...) et de vulgarisation scientifique mises en place et exploitées durant ce contrat pour diverses actions de formations ou diffusion de la culture scientifique sont disponibles sur le [site web du LIFAT](#) (École d'été, workshops, ...).

## 4. Prise en compte des recommandations du précédent rapport niveau LIFAT

### 4.1. Rappels des résultats et objectifs

Rappelons tout d'abord que le **LIFAT avait obtenu une excellente évaluation lors de la dernière visite** :

Extrait retour évaluation HCERES : « *La demande d'association du laboratoire au CNRS est pleinement justifiée. Pour le CNRS, il s'agit d'un laboratoire qui mène une recherche de qualité, entretient des relations fructueuses avec les sciences sociales et le secteur santé, et possède un bon ancrage régional et industriel. L'association avec le CNRS ouvrirait notamment la possibilité d'accueillir sur place des personnels en délégation* »

Extrait retour section 7 du CNRS : « *La section aurait souhaité envoyer un message d'encouragement en donnant un avis favorable à l'association de l'unité au CNRS. Cependant, le CNRS ne souhaite pas actuellement associer de nouvelles unités. Dans ce contexte, il faudrait un niveau scientifique et un rayonnement encore plus élevé pour que la section défende le projet, malgré la politique scientifique de l'organisme.* »

Les recommandations et objectifs du LIFAT fixés pour la période 2016-2021 sont rappelés ci-après. Dans les grandes lignes, il s'agissait d'améliorer la communication interne (Blois-Tours) et externe (site web, diffusion de ressources, ...) et le rayonnement à l'international du LIFAT. Il s'agissait également de maintenir le nombre et la qualité des publications dans des revues internationales référencées et de renforcer les collaborations inter-équipes. Sur le plan scientifique, les équipes déclineront chacune dans leur partie, les objectifs qui leur sont propres, relatifs à leurs thèmes de recherche et aux défis qu'elles abordent dans leur domaine. Les prises en compte des recommandations faites aux différentes équipes sont donc présentées en section D3

## 4.2. Rappel des objectifs et stratégie de l'unité de recherche pour le contrat en cours

Les constats faits lors de la campagne et la prise de fonction de la nouvelle équipe de direction du LIFAT (JY Ramel, A. Giacometti) en 2017 ainsi que les remarques formulées par le comité d'évaluation et par le Comité National ont servi de feuille de route pour définir la stratégie du LIFAT jusqu'en fin 2021. Celle-ci s'est concentrée sur plusieurs volets, synthétisés dans le tableau suivant.

Volet	Stratégie mise en place, actions entreprises
<p>L'accès à la production scientifique du laboratoire mérite d'être facilité.</p> <p>De même, la diffusion des principaux logiciels et leur accès devraient être améliorés</p>	<p>Systématisation de l'utilisation de HAL (<a href="#">collections LIFAT</a>, <a href="#">BDTLN</a>, <a href="#">ROOT</a>, <a href="#">RFAI</a>), incitation mise en ligne du texte intégral</p> <p>Les sites web <a href="#">LIFAT</a> et équipes (<a href="#">RFAI</a>, <a href="#">BDTLN</a>) sont maintenu à jour donnant accès aux productions du LIFAT. Un <a href="#">Gitlab lifat</a> a été mis en place.</p> <p>L'analyse proposée section 3.1 démontre une amélioration sur ce point</p>
<p>L'effort d'amélioration de la qualité de la production doit être poursuivi</p>	<p>L'objectif fixé est que chaque doctorant termine sa thèse avec au moins une publication en revue (acceptation finale éventuellement après la soutenance).</p> <p>Un référencement des conférences et journaux de référence (SJR, QUALIS, ...) a été mis en place dans chaque équipe (voir bilan scientifique de chaque équipe)</p> <p>Un budget LIFAT permet de faire appel à un éditeur privé pour la correction des articles soumis en revue avec un doctorant comme co-auteur (utilisable à une seule phase du processus de soumission).</p> <p>Les analyses proposées section 3 démontrent une amélioration sur ce point</p>
<p>Le rayonnement international est encore à améliorer dans certaines thématiques.</p> <p>Des pistes restent à approfondir pour améliorer encore la visibilité du laboratoire, dont l'organisation de grands événements et l'invitation de chercheurs de renom.</p>	<p>Le LIFAT a nommé un responsable « Activités Internationales ». Le LIFAT fait partie des laboratoires les mieux dotés par le <b>BQI (Bonus Qualité International)</b> mis en place par l'université pour inciter au rayonnement international. La dotation dépend des données saisies dans la base LABRI de l'université (CF Annexe A)</p> <p>Les participations à des projets européens et internationaux, les dépôts de nouveaux projets, et l'accueil de doctorants en cotutelle se sont intensifiés</p> <p>Le LIFAT a poursuivi sa politique d'invitation de professeurs étrangers, avec en moyenne deux invitations par an pour une durée d'un mois minimum chacune.</p> <p>Le LIFAT a organisé les conférences internationales VEROLOG16, ICPRS19, GBR19, Matheuristics18. La crise sanitaire a ensuite rendu difficile l'organisation d'événements scientifiques.</p> <p>Une analyse plus complète est disponible en sections 2.1, 2.2 et 2.3.</p>
<p>La communication interne et externe du laboratoire est perfectible.</p> <p>Il serait bon d'organiser un séminaire régulier bien visible, au niveau du laboratoire</p> <p>La diffusion des séminaires d'équipe devrait être élargie.</p>	<p>Le site web LIFAT comporte des rubriques <a href="#">News</a>, <a href="#">Positions</a>, ... exploitées pour la diffusion d'informations et d'annonces. Une <a href="#">plaquette</a> + diverses actions de communication ont été opérées.</p> <p>Un séminaire LIFAT annuel au vert a été instauré - Ouverture des séminaires équipe à tout le laboratoire.</p> <p>Une analyse plus complète est disponible en section 1.2.2.</p>
<p>Il faut veiller à ce que tous les projets partenariaux présentent un réel intérêt scientifique. La valorisation industrielle de certains développements effectués au CETU pourrait être envisagée</p>	<p>La collaboration entre le CETU ILIAD3 et le LIFAT est tout à fait opérationnelle et permet aux chercheurs du LIFAT de se concentrer sur les aspects recherche tandis que le CETU ILIAD gère la valorisation et le transfert des résultats vers le monde socio-économique.</p> <p>Voir section 2.5.1 et portfolio</p>
<p>Les efforts visant à la coopération entre les équipes méritent d'être poursuivis, voire renforcés.</p>	<p>Un gros effort de la direction et des équipes a été fait pour renforcer les coopérations inter-équipes : dépôt d'un projet fédératif façonnable-IA (APR 2019), renforcement des prises de décision au niveau du conseil de laboratoire (profilage des postes et des bourses de thèse, ...), instauration d'une bourse de thèse inter-équipe annuelle, diffusion élargie des séminaires équipe, news sur site web LIFAT (séminaires, workshops, journées scientifiques, soutenances, ...), mise en place d'un séminaire annuel au vert,... Les collaborations sont maintenant très nombreuses.</p>

<p>La durée normale des thèses devrait être plus souvent respectée. Il convient enfin d'éviter la dispersion des formations, notamment en master</p> <p>Il serait bon que les doctorants soient plus systématiquement incités à suivre des écoles thématiques.</p>	<p>Un comité de suivi scientifique de la thèse a été ajouté au comité imposé par l'ED MIPTIS. Les doctorants proviennent plus souvent de nos formations de master locales (BDMA, Polytech Tours)</p> <p>Plus de 20 participations de doctorants à des écoles d'été</p> <p>Malgré toutes ces actions de suivi mises en œuvre, la durée des thèses n'a que très peu diminué (autour de 43 mois en moyenne au lieu de 44 au précédent contrat) ; la crise sanitaire a impacté en partie cette durée moyenne.</p> <p>Voir analyse détaillée section 2.3.2</p>
<p>L'axe de recherche sur la science des données, commun aux 3 équipes, pourrait être mieux exploité en externe dans l'affichage des spécialités de laboratoire, et en interne dans les collaborations inter-équipes...</p>	<p>La direction du LIFAT s'est engagée dans de nombreuses actions ayant comme objectifs de faire connaître les compétences en sciences des données au niveau régional (RTR DIAMS, fédération ICVL, IA-Loire-Valley, collaboration MSH-CVL, CA + école d'été SFR Neuro-imagerie, séminaire cosmetic-valley, formation IA avec la faculté de médecine, CPER Valopat, PIA4, ARD IPAT/Junon, ...)</p> <p>Les membres du LIFAT ont participé à de nombreuses tables rondes de diffusion scientifique, actions Centre-science, bulletin d'associations (AFIA, GDR, EGC, ...). Le LIFAT est maintenant très fortement identifié en Région et nationalement, et très/trop sollicité (CF nombre de projets et thèses CIFRE sur le contrat par exemple).</p>
<p>Les points à améliorer dépendent malheureusement peu des efforts du laboratoire : la bi-localisation du laboratoire entre Blois et Tours est évidemment source de complications</p>	<p>Le renforcement de la cohésion entre équipes, composantes et sites géographiques a été l'un des principaux objectifs de l'équipe de direction lors de son mandat. Plusieurs actions ont été mises en place pour rapprocher les chercheurs (EC, doctorants, postdoc) des 2 sites Blois/Tours : co-direction bi-site du LIFAT, incitation au rapprochement des formations niveau master, renforcement des prises de décision au niveau du conseil de laboratoire, séminaire annuel au vert, ...). Ces actions ont eu de fortes répercussions.</p>
<p>Surcharge déraisonnable des enseignants-chercheurs en tâches d'enseignement et d'administration est un handicap sérieux.</p> <p>Le renfort d'un personnel BAP E en développement logiciel permettrait une meilleure diffusion et valorisation de l'importante production logicielle des équipes.</p>	<p>Les charges administratives supportées par des EC LIFAT se sont encore amplifiées sur ce contrat (cf section 2.). Une <b>incitation à la demande de CRCT, de délégation ou de décharge</b> a été faite (2 demandes LIFAT par an) : au moins un EC a bénéficié d'un CRCT ou d'une décharge de 6 mois chaque année durant le contrat.</p> <p>La création du CETU ILIAD3 permet <b>d'auto-financer</b> depuis 2017 <b>1 CDI IGR temps plein + d'autres temps pleins IGE ou Techniciens en CDD</b> en permanence. Ceci est très positif mais ne couvre pas le besoin initial d'un personnel dédié au LIFAT n'ayant pas à s'autofinancer.</p> <p>Des demandes de sortie de réserve (fonds de roulement) pour l'embauche d'un IGE ont été faites et refusées à 2 reprises sur le contrat malgré un budget disponible. Le dépôt du projet APR Façonnable-IA pour l'embauche d'un IGR LIFAT afin de mieux capitaliser les productions n'a pas été retenu.</p>

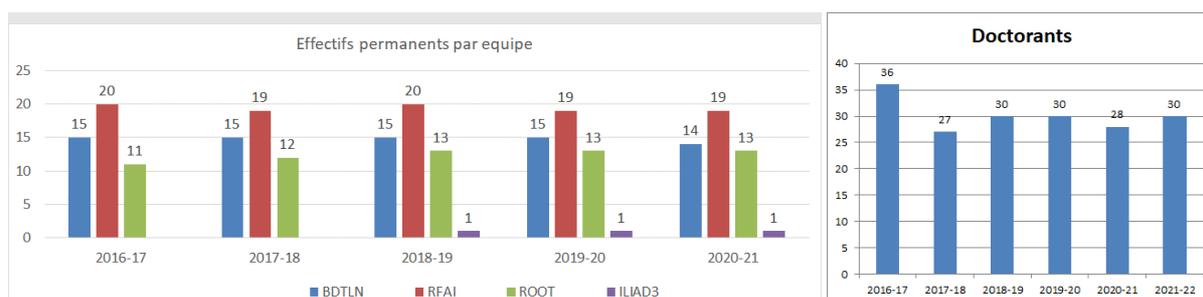
# DOCUMENT D'AUTOÉVALUATION

## 1. Domaine d'évaluation 1 : Profil, ressources et organisation de l'unité

### 1.1. Référence 1. Profil d'activités et environnement de recherche

#### 1.1.1. Moyens humains

Au cours de la période 2016/2021, les effectifs en enseignants-chercheurs ont très peu évolué. Ils sont restés stables autour de **46 Enseignants-Chercheurs (+ 2 renouvellements en cours)**. De même, les effectifs en doctorants sont également stabilisés autour de **30 doctorants par an**.



- Fig. 1.1 Evolution des effectifs LIFAT (a/permanents – b/doctorants)

Une analyse plus fine sur les recrutements, sur le suivi de carrière et le déroulement des doctorats est proposée en section 2.3.1. Concernant la répartition Femme/Homme (Fig 1.3), le LIFAT n'est pas très bien positionné avec environ 22% de femmes dans son effectif global ; les chiffres étant encore moins favorables lorsque l'on regarde les membres permanents et les grades : 1 professeure seulement sur 14 mais recrutée non au LIFAT mais dans une université de Lille (bien qu'elle soit rattachée désormais au LIFAT) et 6 (5 depuis un départ pour une promotion PU en septembre 2021) maîtresses de conférences durant le contrat sur 32. Sur les 6 recrutements de MCF au cours du contrat, il y a eu 1 femme récemment, mais une autre a quitté le LIFAT pour prendre un poste de PU à Saclay donc il n'y a toujours que 7 permanentes sur 46 fin 2021, soit 15%, ce qui est bien en dessous de la moyenne de 24% de femmes enseignantes chercheuses en informatique dans les universités françaises (CNU 27), moyenne recensée en 2020-21 par le MESR (<https://data.esr.gouv.fr/FR/>). Nous sommes conscient.e.s que des actions doivent être mises en place sur ce plan mais il est difficile de faire évoluer les statistiques lorsqu'aucun poste n'est ouvert (0 recrutement de professeur sur le contrat).

Le LIFAT a participé (subvention) avec polytech Tours et l'association femmes et mathématiques à la journée intitulée "filles, maths et informatique : une équation lumineuse" qui a eu lieu le 14/12/2021. De plus, une conférence sur les biais inconscients liés au genre dans l'environnement professionnel est planifiée pour notre prochain séminaire (13/07/2022).

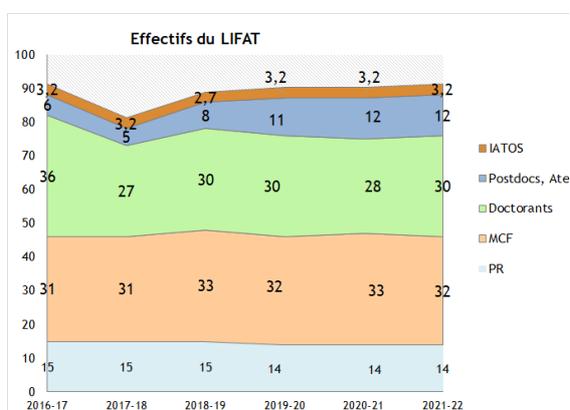


Fig. 1.2 : Evolution des effectifs toutes catégories

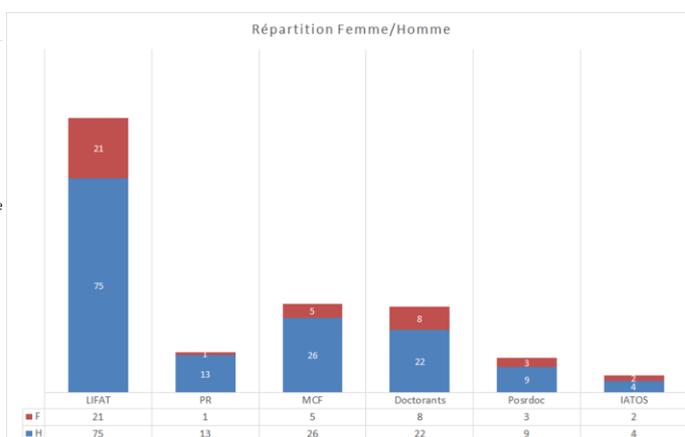


Fig. 1.3 : Répartition Femme/Homme par grade

#### 1.1.2. Moyens financiers

Nous distinguons dans notre gestion deux sortes de crédits financiers : les crédits « non fléchés », récurrents et indépendants des projets de recherche (le financement attribué par la Commission Recherche de l'Université de Tours, par le CNRS pour l'ERL, par

l'INSA-CVL, le « Bonus Qualité International », le reversement du Préciput par l'Université, le reversement d'une part de la DGF de Polytech Tours) et les crédits « fléchés », non récurrents, plus ponctuels, liés à nos projets de recherche.

- Les crédits « non fléchés » ou récurrents subissent un prélèvement pour la politique du laboratoire, et les crédits restants sont affectés aux équipes selon des clés de répartition (sauf le budget du CNRS qui est uniquement affecté à l'ERL) votées en AG. Ces budgets sont restés stables pendant la période 2016-2021.

Les crédits « fléchés » sont issus de l'action spécifique d'une équipe ou de plusieurs, et ils sont affectés entièrement aux équipes concernées. Ces budgets sont davantage fluctuants et les succès ne sont pas les mêmes d'une année à l'autre, mais on note globalement une augmentation du nombre de projets acceptés et des financements qui les accompagnent (en grande partie pour financer du personnel).

Le budget récurrent est d'environ 90 k€ par an et représente donc environ 10% du budget total du LIFAT.

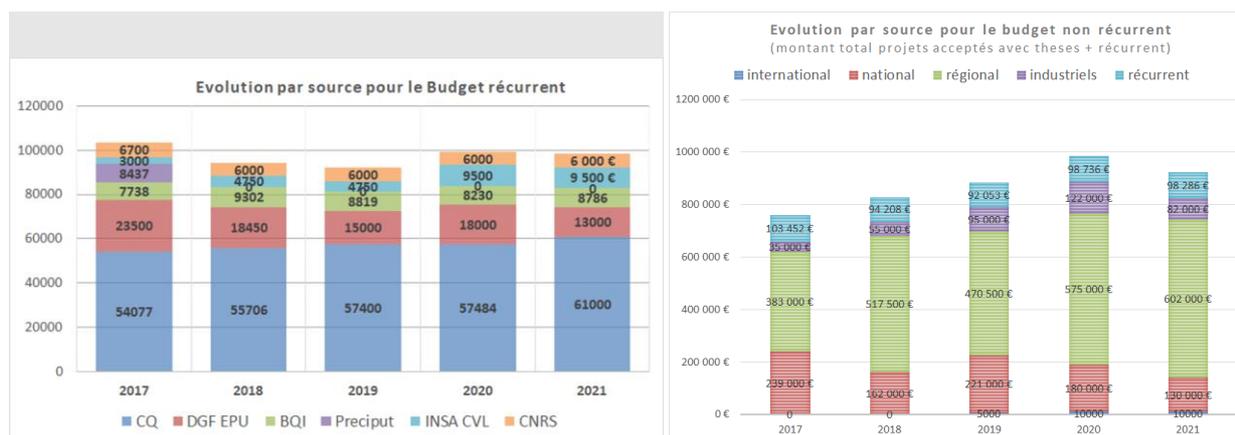


Fig 1.4 : Evolution du budget récurrent et « fléché » du LIFAT

#### Evolution du Budget Récurrent :

- Financement récurrent à peu près constant (règles de répartition établies par la Commission Recherche de l'université de Tours).
- Suppression du Préciput reversé par l'université depuis 2018 et diminution mineure du reversement de la part de Polytech Tours (sur sa « Dotation Globale de Fonctionnement ») depuis 2020.
- Notons le « Bonus Qualité International » (BQI), qui réussit très bien au LIFAT (budget alloué par l'Université sur la base de l'activité à l'international de l'unité dans l'année précédente).

#### Evolution du budget global (pour chaque année est indiqué le montant total correspondant aux projets exécutés dans l'année) :

- Nous notons une augmentation du budget obtenu en 2018, et une part de plus en plus importante des financements de la Région Centre-Val de Loire.
- Il est à noter que les salaires correspondant aux bourses institutionnelles (1 bourse Région + 1 bourse ministérielle/établissement obtenues en moyenne chaque année par le LIFAT) sont intégrés **dans la part du budget fléché** et présentée dans les statistiques (bourses Régions intégrées dans le budget Régional et bourses ministérielles intégrées dans le budget National – Fig. 1.4).

Les clés de répartition des budgets (montants prélevés et clés de répartition) sont votées chaque année en Assemblée Générale. Pour les crédits non fléchés, des sommes sont prélevées pour la politique du laboratoire :

- Le financement du CDI à mi-temps du secrétariat du Laboratoire
- L'appel à un éditeur professionnel pour la correction d'articles en anglais des doctorants
- Subvention à la fédération ICVL pour l'incitation aux échanges LIFAT-LIFO (missions essentiellement)
- Réalisation d'un séminaire annuel LIFAT et autres exposés scientifiques

Le montant total de ces **prélèvements pour le fonctionnement du LIFAT représente moins de 40% (35 k€/an environ)** du budget « non fléché ». Les sommes restantes sont réparties de la façon suivante :

- Un pourcentage (80%) est réparti entre les équipes au prorata de leurs effectifs.
- Le reste (20%) est réparti équitablement entre les équipes.

Les crédits fléchés (projets) sont pour la plupart « justifiables » sur le plan comptable. Autrement dit, les administrations qui financent demandent une justification précise des dépenses (projets Région, FEDER, ANR, prestations). En cas de problèmes pour justifier les dépenses, les crédits alloués peuvent devoir être à rembourser. Dans le cas de crédits justifiables, ceux-ci sont donc affectés intégralement au projet, sans faire l'objet de prélèvement par le laboratoire. Il en est de même lorsque les crédits fléchés sont « non

justifiables » (CIFRE, projets ponctuels) sur lesquels aucun prélèvement laboratoire n'est fait (sauf lorsque la réglementation l'impose, projet ANR par exemple).

Le suivi du budget est réalisé par le secrétariat du LIFAT, en collaboration étroite avec l'antenne financière de Polytech Tours. Notons que compte tenu de la complexité croissante des procédures administratives et comptables, ce suivi demande une technicité de plus en plus élevée et de plus en plus de travail.

### 1.1.3. Moyens mutualisés

Plusieurs types de moyens sont mutualisés : des personnes, des budgets (les deux étant liés), ainsi que des financements de thèses. Le personnel mutualisé est le secrétariat du laboratoire, dont une partie correspond effectivement à une mutualisation de moyens financiers (cf. explications précédentes). Une partie du budget est également mutualisée (cf. explications précédentes).

Nous insisterons ici sur une autre mutualisation qui a été mise en place au LIFAT : la mutualisation des financements de thèses. En moyenne, le laboratoire bénéficie chaque année de deux bourses de thèse, en provenance de la Région et/ou du Ministère. La politique qui a été mise en place pour l'attribution de ces bourses jusqu'en 2021 est une horloge avec quatre bénéficiaires : chaque année une équipe à tour de rôle plus « le laboratoire » bénéficient d'un financement de thèse. Le sujet attribué à une équipe est donné par l'équipe, selon les priorités qu'elle définit. **Le sujet « laboratoire » doit être un sujet de thèse inter-équipes**, avec un co-encadrement dans deux équipes. **Ainsi 5 thèses inter-équipes ont été mises en place entre 2016 et 2021.**

Même si les arbitrages ne sont pas faciles à faire, cette mutualisation de certains financements de thèses a permis un fort rapprochement entre les équipes du laboratoire ainsi que l'émergence d'axes fédérateurs et spécialisant (cf section 1.3).

Au niveau du matériel, le LIFAT dispose largement des moyens nécessaires à la réalisation de ses recherches notamment au travers de mutualisation d'équipements et de plateformes obtenus au travers de différents financements ou accords avec d'autres entités locales et composantes d'enseignement (cf section 2.5).

## 1.2. Référence 2. Positionnement et objectifs scientifiques

### 1.2.1. Positionnement

Les thématiques et domaines de recherche des équipes du LIFAT sont synthétisés dans le tableau suivant. Des détails supplémentaires sont fournis dans les parties réservées à la description plus précise des activités de chacune des équipes.

	BdTIn	RFAI	ROOT
<b>Cœur de l'activité</b>	Langages et algorithmes pour l'analyse exploratoire, la fouille de données et l'évaluation de qualité Aspects sémantiques des données et des services : définition, découverte et exploitation Ressources pour le TAL, découverte et mise en relation d'éléments d'intérêt en document	Traitement et analyse d'images et de vidéos Représentation à base de graphes Apprentissage automatique Systèmes dynamiques et interactifs de reconnaissance des formes et de fouille visuelle de données.	Nouveaux algorithmes de résolution de problèmes d'optimisation combinatoire. Recherche Opérationnelle : - Complexité, Méthodes arborescentes, Programmation mathématique, Algorithmes de graphes, Métaheuristiques/Matheuristiques, Approximations, etc.
<b>Objectifs scientifiques, problèmes académiques</b>	- Systèmes interactifs d'analyse, de fouille et de recommandation - Définition et exploitation de la sémantique des services et données du Web - Traitement automatique des langues et interaction.	- Analyse d'images et vidéos - Fouille visuelle de données - Reconnaissance des formes	- Ordonnancement - Transport et logistique - Problèmes complexes intégrés
<b>« Défis sociétaux »</b>	- Santé (intégration de données, services et modélisation) et handicap (aide à la saisie) - Arts et sciences humaines (interopérabilité sémantique, fouille de trajectoires, exploration et analyses de données)	- Santé (imagerie médicale) et handicap (autisme, déficience visuelle) - Arts et sciences humaines (numérisation et valorisation du patrimoine)	- Santé (organisation des systèmes de santé) - HPC - Mobilité et développement durable - Production industrielle

BdTln	RFAI	ROOT
	- Environnement (agriculture numérique, jumeaux environnementaux)	
<b>Positionnement vis-à-vis de l'écosystème associé (« concurrence »)</b>	<p>- L'humain au centre (de l'analyse des données, ...)</p> <p>- L'attention portée à la connaissance liée aux objets d'étude, à la sémantique et à la qualité de ces objets : données, services, scénarii d'analyse et d'exploration, mesures et algorithmes</p>	<p>L'écosystème recherche dans lequel évolue RFAI est large (CV, ML) et hautement concurrentiel en ce moment, les points différenciant de RFAI sont principalement :</p> <p>- la prise en compte de l'humain (interactivité, explicabilité) dans les méthodes conçues</p> <p>- l'usage de représentation de type graphe et la combinaison d'approches de recherche opérationnelle et d'apprentissage automatique</p> <p>- le travail sur les flux de données et les données séquentielles ou temporelles</p>
		<p>La résolution de problèmes issus de défis sociétaux (santé notamment).</p> <p>La résolution de problématique d'ordonnancement. L'hybridation de techniques d'optimisation (RO-IA, algorithmique modérément exponentiel, etc.)</p>

A ces domaines se rajoutent les thématiques donnant lieu à des collaborations inter-équipes qui ont été explicitées dans la section Thématiques scientifiques de la partie Présentation générale du LIFAT :

- Combinaison de méthodes d'**Apprentissage Machine et de Recherche Opérationnelle**
- Introduction de plus de **sémantique et d'explicabilité dans les systèmes d'aide à la décision**
- **Developpement d'approches pluridisciplinaires** notamment au travers de collaborations avec différents partenaires dans les domaines de la "Santé et du Handicap" et de l'Art et des Sciences Humaines

### 1.2.2. Animation scientifique

Depuis 2019, un **séminaire au vert d'une journée est organisé par le laboratoire** chaque année en juillet. Cet événement est l'occasion, pour l'ensemble des membres du LIFAT d'avoir un panorama sur les différentes activités dans les équipes (recherches et projets en cours, ...) et d'échanger durant un moment de convivialité (activités ludique organisée en fin de journée). Une table-ronde est généralement prévue pour échanger sur la stratégie scientifique et les actions pouvant être suivies au niveau LIFAT. Les doctorants sont également impliqués par la présentation de poster synthétisant leurs travaux de recherche.

Les séminaires scientifiques organisés par les équipes LIFAT sont fréquents. Ils sont organisés en interne au niveau des équipes, de façon régulière selon un calendrier fixé à l'avance et aussi selon les opportunités (invitations de professeurs étrangers, visites, etc.)

Les annonces de séminaires sont diffusées à tous les membres du LIFAT et ouverts à tous (personnes extérieures au LIFAT également).

Le tableau ci-dessous synthétisent les activités d'animation scientifiques proposés aux membres du LIFAT durant ce contrat quinquennal.

Niveau Laboratoire LIFAT	BDTLN	RFAI	ROOT
<p>1 à 2 conférences invitées par an proposées par les chercheurs en visite longue au LIFAT</p> <p>1 séminaire au vert annuel en juillet</p> <p>1 séminaire ICVL annuel commun avec le LIFO en janvier</p> <p>1 journée Doctorants annuelle en avril</p> <p>Et aussi :</p> <ul style="list-style-type: none"> <li>- 4 journées du RTR DIAMS</li> <li>- 3 workshops internationaux,</li> <li>- Conférence EGC 2021</li> <li>- 2 expositions, 1 journée SFR...</li> </ul> <p><a href="#">Voir site Web</a></p>	<p>1 exposé scientifique par mois au début de chaque réunion d'équipe</p> <p>Une fois par an, demi-journée en juin à Blois, ouverte aux étudiants et public externe : de 6 à 8 exposés</p> <p>« Apéros info » organisés en soirées de 2016 à 2020, ouverts aux étudiants et public externe (stoppés momentanément par COVID, repris en 2022) : 3 à 4 épisodes par an</p>	<p>L'équipe RFAI a organisé environ 40 séminaires scientifiques entre 2016 et 2021 (avec une moyenne annuelle de 10 séminaires. L'équipe se réunit en moyenne 1 fois par mois.</p>	<p>L'équipe ROOT organise depuis 2017 des réunions d'équipe une fois par mois et des brainstormings ou séminaires une à deux fois par mois.</p>

Chaque année depuis 2005, les membres du LIFAT et du LIFO se retrouvent à l'occasion d'une « Journée Informatique Région Centre (JIRC) », organisée à tour de rôle sur les différents sites (Blois, Bourges, Orléans, Tours). Ces journées comportent un exposé scientifique par un extérieur, un ou plusieurs exposés scientifiques des membres des laboratoires, et un temps est gardé pour des échanges informels. En règle générale, un conférencier extérieur au LIFAT et au LIFO est invité.

Chaque année, au mois de novembre, les doctorants du LIFAT et du LIFO se réunissent lors d'une journée organisée conjointement par l'ED MIPTIS et le RTR DIAMS (à Tours, Blois ou Orléans). Cette journée est l'occasion pour les doctorants de présenter leurs sujets de thèse (exposés courts pour les doctorants qui démarrent, plus longs pour les doctorants aguerris), d'échanger sur les fonctionnements des unités et de mieux se connaître. Chaque réunion regroupe une quarantaine de participants. Cette manifestation est organisée par les doctorants, avec le soutien financier et logistique du laboratoire d'accueil (pour ce qui concerne l'accueil) et du laboratoire de départ (pour ceux qui se déplacent). Quelques enseignants-chercheurs assistent à ces séminaires.

### 1.2.3. *Projet scientifique LIFAT pour l'avenir*

En plus des projets scientifiques de chaque équipe, le LIFAT encourage les travaux inter-équipes grâce en particulier à des bourses de thèses. Pour lancer ces travaux, nous avons mis pendant plusieurs années la moitié des financements académiques (établissement et Région), nous venons de réduire à un tiers de ces financements car les équipes ont maintenant pris l'habitude de mieux travailler ensemble. Les principales thématiques inter-équipes sont :

- Science des données
- Apprentissage pour l'optimisation
- Optimisation pour l'apprentissage
- Temporalité
- Explicabilité

Au niveau applicatif, trois axes transversaux seront particulièrement adressés :

- Santé et handicap
- Arts et sciences humaines numériques
- Environnement et développement durables

Le premier "Santé et handicap" est depuis longtemps un axe transversal applicatif fort et de nouveaux projets continuent d'arriver.

Le deuxième "Arts et sciences humaines numériques" est une évolution de l'axe "humanités numériques" de ces dernières années. Ainsi nous souhaitons élargir nos partenariats avec les laboratoires SHS avec lesquels nos relations se renforcent.

Le troisième "Environnement et développement durables" est un nouvel axe applicatif dans lequel nous souhaitons nous investir.

Le LIFAT se sent fortement concerné par l'application et l'impact de ses recherches sur la société et l'environnement. Deux projets de recherches regroupant un grand nombre de chercheurs viennent d'être lancés ou proposés. Le projet régional Junon (porteur BRGM) où le LIFAT participe à la conception de nouveaux services numériques pour améliorer le suivi et la compréhension de l'environnement, pour une meilleure gestion des ressources naturelles. De même, dans le cadre du projet ANR MEMO (porté par l'IGN en deuxième phase) le LIFAT participe à la définition de nouveaux outils de surveillance d'écosystème via l'intégration de données enrichie sémantiquement issue de sources complémentaires.

En ce qui concerne l'impact de notre recherche sur la société. L'un des sujets principaux concernant le numérique nous semble être l'inversion des courbes inquiétantes de l'impact des activités humaines sur la planète, les espèces animales et végétales, l'eau, l'air, le climat. Le LIFAT prend en compte ce problème dans ses activités de recherche. Par exemple, via la définition de méthode d'échantillonnage représentatif de gros volume de données. Ainsi, des analyses énergivores peuvent être lancées sur des jeux de données de taille beaucoup plus faible. Finalement, des réflexions sur nos activités nous ont poussés à changer nos comportements. Par exemple, en favorisant largement les déplacements en train en France et en Europe, nous contribuons à la baisse de nos sources d'émission de CO2 (le poste déplacement est la plus grosse source d'émission de CO2 pour l'université de Tours).

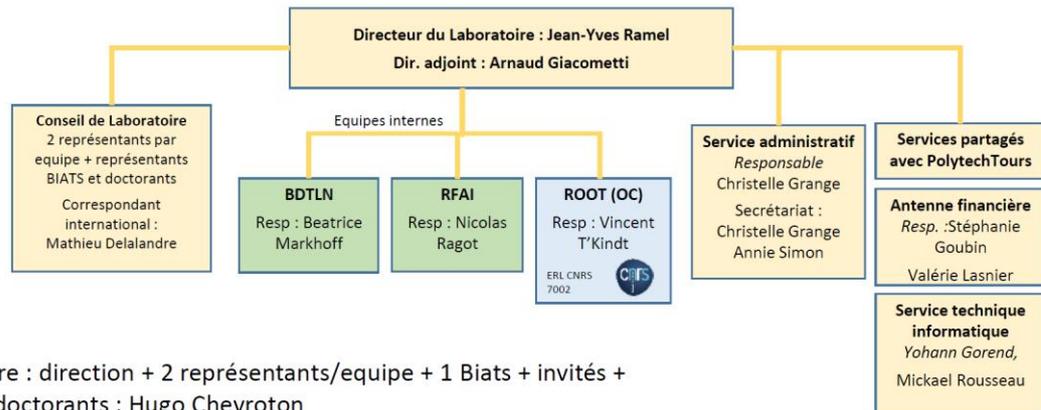
## 1.3. Référence 3. Fonctionnement et GRH

### 1.3.1. *Instances de pilotage*

#### Pour la direction

Le fonctionnement du LIFAT est régi par des statuts et par un règlement intérieur ([disponibles sur le site web](#)). Le laboratoire est dirigé par un Directeur (associé à un directeur adjoint) élu en Assemblée Générale pour une durée égale à celle du contrat ministériel. Il est assisté pour la conduite de la vie scientifique, technique et administrative, d'un Conseil de Laboratoire.

L'organigramme présenté Fig. 1.5 décrit l'organisation majoritairement présente entre 2016 et 2021. Notons néanmoins que JC Billaut a assuré la direction du LIFAT jusqu'à fin juin 2018 (codirection assurée par JY Ramel), que Arnaud Giacometti a démissionné de la direction adjointe en décembre 2020 (suite à son élection à la présidence de l'université de Tours) et que H. Cardot a pris la succession de JY Ramel à la direction le 1er septembre 2021.



Conseil de Laboratoire : direction + 2 représentants/équipe + 1 Biats + invités + 1 représentant des doctorants : Hugo Chevroton

Fig. 1.5 : Organigramme et organisation interne du LIFAT

Le responsable de l'équipe ROOT est également **directeur de l'ERL CNRS ROOT durant la période 2016-2020. A ce titre, c'est lui qui est en contact avec les représentants locaux et nationaux de l'INS2I et non le directeur du LIFAT.** La mise en place d'une **gestion séparée entre l'ERL ROOT et le LIFAT** (budget, logiciel de gestion, missions, ...) génère une surcharge de travail non négligeable.

Le Conseil de Laboratoire comprend le Directeur et son adjoint, deux représentants par équipe de recherche, un représentant des personnels BIATSS et un représentant des doctorants, et de personnalités invitées (le Directeur du LIFO, le Directeur de l'Ecole Doctorale, ...). Il se réunit toutes les six semaines environ. Il est en outre ouvert à tous, et tout membre du laboratoire peut y participer et intervenir (sans prendre part au vote s'il y a lieu). Le Conseil du Laboratoire a un rôle consultatif et de proposition. Il est consulté par le Directeur sur :

- la sélection du sujet de thèse inter-équipe chaque année,
- la définition des profils de postes EC ouverts aux recrutements
- la coordination des programmes de recherche
- les demandes de crédits et de personnels, transmises par chaque équipe de recherche, auprès des divers organismes de tutelle,
- la politique des publications ou communications diffusées au nom du laboratoire,
- la répartition des moyens (financiers, matériels et humains) au sein du laboratoire,
- l'établissement annuel de la liste des membres du laboratoire,
- la tenue à jour du Rapport d'Activité du laboratoire et du site Internet sur la base des rapports d'activité fournis par les équipes.

En outre, les communications en interne transitent par le Conseil de Laboratoire, et il arrive régulièrement que des consultations et décisions se fassent par email sur des questions urgentes non planifiées.

Tout membre du Laboratoire peut demander au Directeur ou aux Directeurs Adjointes l'inscription d'un problème particulier à l'ordre du jour d'une séance du Conseil du Laboratoire.

Une **Assemblée Générale** ordinaire réunit tous les membres du Laboratoire une fois par an en décembre. Lors de cette AG, le bilan de l'activité de l'année écoulée est présenté (production scientifique, soutenances de thèse, budgets), les clés de répartition budgétaire sont votées, et les informations importantes sont présentées et discutées. L'Assemblée Générale ordinaire se termine par un buffet. Une Assemblée Générale extraordinaire a eu lieu en 2014 pour voter la création de la Fédération LI-LIFO, et une en 2015 pour voter le porteur du projet du LIFAT et les statuts de la fédération.

### Pour les moyens financiers

Toute la **gestion des budgets (LIFAT + ERL CNRS + CETU ILIAD3) est assurée avec le soutien du secrétariat et de l'antenne financière de Polytech Tours.** Les budgets relatifs aux projets sont gérés par les porteurs des projets, les budgets (« non fléchés ») des équipes sont gérés par les responsables des équipes, et le budget non fléché du laboratoire est géré par le secrétariat du LIFAT, après une mise en place avec le Directeur, conformément aux décisions votées lors de l'Assemblée Générale.

#### 1.3.2. *Politique de recrutement et d'accueil*

### Politique de recrutement des enseignants-chercheurs

La politique de recrutement d'enseignants-chercheurs affichée au LIFAT se résume aux deux points suivants.

- Pour le profil des postes : durant ce contrat, le profilage des postes EC a été décidé par le conseil de laboratoire. Un appel à profil est diffusé, des propositions émanant des équipes sont remontées et présentées en conseil de Laboratoire qui sélectionne le profil retenu par vote. Cela étant dit, le conseil tient compte de l'historique (renouvellement d'un départ dans une équipe spécifique) et veille à maintenir un juste équilibre pour le développement des équipes.
- Pour le profil des candidats : le comité de sélection recrute le meilleur candidat, conformément à la norme, sans qu'il ne soit formulé de consigne particulière relative à l'origine du candidat (ni préférence locale, ni interdiction). Pour les recrutements de Maîtres de Conférences, nous notons que la qualité des dossiers des candidats s'est très largement améliorée ces dernières années, et qu'un séjour post-doctoral à l'étranger d'un ou deux ans est de plus en plus la norme

requis pour avoir un dossier compétitif. Naturellement, ceci donne assez peu de chances à un candidat local qui aurait fraîchement soutenu sa thèse.

Pour les recrutements de Maître de Conférences, il ressort de cette politique un recrutement essentiellement externe (5 sur 6 recrutements depuis 2016 sont externes, sachant que le seul recrutement 'local' correspond à un candidat qui avait effectué un séjour post-doctoral de deux ans à l'étranger après sa thèse).

En termes de recrutements d'enseignants-chercheurs, de 2016 à 2021, **le LIFAT n'a procédé à aucun recrutement de Professeurs (et a même perdu un poste à la suite d'une demande de mise en disponibilité) alors que le nombre d'HDR est passé de 5 à 10 (5 soutenances + 1 recrutement HDR + un départ vers un poste de professeur à Paris Saclay)**. Cet état de fait constitue un risque / une menace (dans l'analyse SWOT) pour le LIFAT dans l'avenir.

Similairement à la tendance nationale, 6 EC (5 MCF + 1 PR) ont quitté le LIFAT au travers de 4 demandes de mise en disponibilité + 1 départ en retraite + 1 promotion PR. Ces départs ont tous juste été en partie contrebalancés par 6 recrutements MCF effectués sur le contrat.

Sur la fig. 1.1, on peut noter la tendance à aller vers un rééquilibrage des effectifs entre équipes (2016 vs 2021).

Le LIFAT recrute et accueille chaque année de nombreux ATER (de 4 à 10) pour couvrir les besoins en enseignement en informatique dans les différentes composantes.

### Place des BIATSS

La place des personnels administratifs est prépondérante dans l'unité. Notre secrétariat a fonctionné pendant plus de dix ans avec une seule personne en poste, aidée par des personnes à mi-temps sur CDD, qu'il fallait régulièrement former et remplacer. Avec le soutien de l'Université, nous avons pu recruter une personne en CDI (décision votée en conseil de laboratoire puis validée par l'Université en septembre 2015). Le budget est pris sur la part récurrente du budget du laboratoire. Ce recrutement passe les effectifs du secrétariat à 1,5 permanents ETP de façon pérenne (auxquels on ajoute deux dixièmes correspondant à des personnes en poste à Blois, et deux mi-temps correspondant au personnel technique, en poste à la DSI et rattachées à Polytech Tours). En tout, ce sont donc 2,7 personnes ETP qui constituent le personnel BIATSS permanent du LIFAT.

La complexité croissante de l'activité de secrétariat et la multiplication des budgets à gérer (comptes fédération ICVL, comptes CETU IliAD3, ...) rend cette tâche de plus en plus difficile. Les effectifs en personnel administratif permanents sont restés constants depuis 2015. Le fait que ce type de personnels (secrétariat, personnel technique) soit géré administrativement (demande de postes, promotions, ...) par les composantes d'enseignement (responsables administratifs) et non pas par les directeurs de Laboratoire pose problème. Ainsi, depuis juin 2021, le LIFAT ne dispose plus d'un secrétariat stable ce qui pose d'énormes difficultés. Cet état de fait constitue également un risque figurant dans l'analyse SWOT du LIFAT.

Concernant le suivi de carrière de ces personnels, depuis 2014, le Directeur du laboratoire fait passer les entretiens professionnels au personnel BIATSS du laboratoire chaque année (CDI BIATS, CDD IGR, IGE, Tech). Le personnel BIATSS du laboratoire suit régulièrement des formations : Geslab et Reseda CNRS 2016 et 2017, Formation SIFAC (2018), formation AiTour (2019), formation GBCP CNRS (8 heures, 2021), élaboration d'un bilan financier d'un projet de recherche – confirmés (une journée, 2016), inscription à une formation 'se réconcilier avec l'orthographe' (3 h 2017).

Malgré plusieurs tentatives, il n'a pas été possible de mettre en place un support pérenne de type BAP E en développement logiciel, diffusion et valorisation sur ce contrat.

### Recrutement des doctorants

Les effectifs des doctorants sont restés stables pendant la période. **Nous avons en moyenne 30.2 doctorants travaillant au LIFAT (alors que nous avons 26,5 doctorants lors du précédent contrat)**.

L'amélioration du déroulement du doctorat passe par une amélioration du recrutement des doctorants et par une amélioration du suivi du déroulement de la thèse.

La politique de recrutement du LIFAT vise à garantir la qualité des doctorants et leur adéquation aux thématiques de recherche des équipes. Les équipes sont les premiers acteurs de la sélection des candidats, en s'assurant de la formation initiale du candidat et du niveau de ses résultats universitaires.

- Pour les financements de thèse de la Région Centre-Val de Loire ou pour les Allocations de Recherche, les candidats de la « short-liste » sont auditionnés par un jury constitué par l'École Doctorale, qui effectue un classement. Ces doctorants bénéficient alors d'un Contrat Doctoral.
- Pour les financements de thèses CIFRE, les candidats sont choisis en concertation avec l'entreprise (parfois c'est le département Ressources Humaines de l'entreprise qui effectue le choix final).
- Pour les autres financements de thèses (projets ANR, projets Région, bourses de pays étrangers, etc.), la sélection se fait généralement sur entretien et après une étude confiée au candidat. Les annonces des offres de thèses se font à l'échelle nationale et parfois internationale.

Tous les doctorants sont inscrits à l'École Doctorale et suivent des modules d'enseignement. Tous les doctorants, ainsi que leur(s) encadrant(s), le directeur d'unité et le Président de l'Université signent une « Charte des thèses » spécifiant l'encadrement, la durée de la thèse, les procédures de médiation, les règlements, etc.

La figure 1.6 illustre les types de financement des thèses sur le contrat. On note un nombre **très important de thèses CIFRE (13)**. Le taux de **doctorants étrangers est de 60%** (sur les thèses soutenues depuis 2016 ou en cours), en provenance de pays très variés (Algérie, Vietnam, Palestine, Mauritanie, Colombie, Tunisie, Sénégal, Inde, Liban, Brésil, Mexique, Gabon, Albanie, ...). Ce chiffre montre notre aptitude à attirer des doctorants venant de l'étranger. **Environ 22% (54% des doctorants français) de nos doctorants proviennent assez souvent de nos formations de master locales (BDMA, Polytech Tours).**

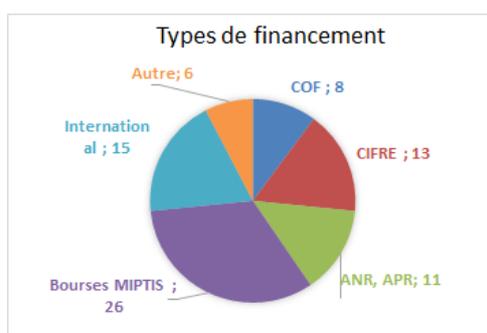


Fig 1.6 : Répartition par type de financement des thèses

Le LIFAT n'admet pas de doctorant s'il ne bénéficie pas d'un financement. Pour certains doctorants étrangers, bénéficiant d'une bourse de leur pays, il est souvent nécessaire de trouver un complément de bourse pour leur donner la possibilité d'avoir un niveau de vie correct.

Les doctorants élisent un représentant des doctorants, qui participe (et vote) à chaque conseil de laboratoire.

#### Suivi des doctorants

Afin de motiver les doctorants et – le cas échéant – de mettre en lumière les problèmes le plus tôt possible, le LIFAT a mis en place en septembre des **comités de suivi scientifique de thèses en plus de celui proposé par l'ED MIPTIS qui ne concerne pas les aspects scientifiques**. Le comité de suivi de thèse d'un doctorant est une structure relativement souple, composée :

- du ou des encadrant(s) du doctorant,
- d'un collègue enseignant-chercheur extérieur à l'établissement (sans condition sur son statut),
- du directeur du laboratoire ou de son représentant.

Le comité de suivi de thèse MIPTIS se réunit une fois par an (donc deux fois en tout) et le LIFAT demande à ce qu'il y ait en plus au moins un (2 conseillés) comité de suivi scientifique en 2<sup>e</sup> année de thèse. Il est demandé au doctorant de fournir un rapport écrit plus une présentation des premiers articles soumis ou publiés et un projet d'article pour une revue.

Ce qui est demandé au membre extérieur :

- regarder les éléments qui lui sont envoyés,
- participer à la présentation orale (éventuellement en visio) et envoyer un avis écrit dans lequel il donne un avis sur la poursuite de la thèse.

Ce dispositif prend de plus en plus d'envergure et se traduit parfois par des soutenances à mi-parcours publiques.

#### Thèses et HDR soutenues

Au cours de la période 2016-2021, **40 thèses ont été soutenues (soit une moyenne de 6.6 thèses soutenues par an - 7.8 au précédent contrat)**. La **durée moyenne de ces thèses est d'environ 43 mois**, soit 3 ans et 7 mois. Sur les 79 doctorants qui se sont inscrits en thèse depuis 2016 (30 thèses toujours en cours), nous déplorons 9 abandons, soit un **taux d'abandon de 11.4%**.

L'évolution notable durant ce contrat est l'augmentation du nombre de thèses inter-équipes (6 depuis 2016) témoignant de la volonté d'intensifier les collaboration inter-équipes ; choix qui semble avoir été très opportun aujourd'hui car créateur de spécialisation scientifique originales pour le LIFAT (voir section 1.2)

Pour améliorer la production scientifique des doctorants et par suite la qualité des thèses, l'objectif demandé (et rappelé régulièrement) aux doctorants est la publication d'un article en revue en fin de thèse. Pour soutenir cet effort, un budget de 2 k€ est réservé pour faire appel à un éditeur professionnel pour corriger l'anglais des articles cosignés par un doctorant et soumis en revue. Bien entendu, compte tenu des délais de publications, celle-ci peut arriver quelques mois voire un ou deux ans après la soutenance de la thèse. Les statistiques de publications des doctorants sont fournies Fig.3.2.

Une enquête a été effectuée pour étudier le devenir des anciens doctorants. Les résultats sont fournis Fig.1.9.

Notons que **5 HDR ont été** soutenues entre 2016 et 2021.

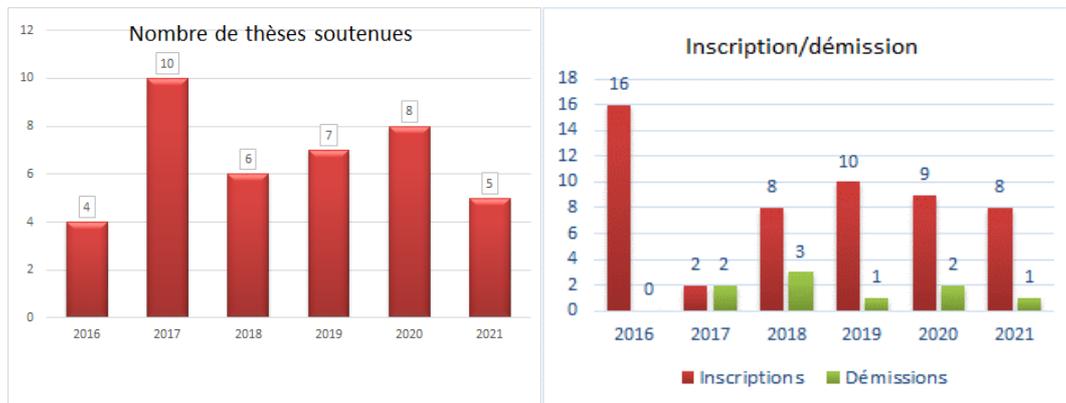


Fig 1.7 : Nombre de thèses soutenues, inscriptions/démissions

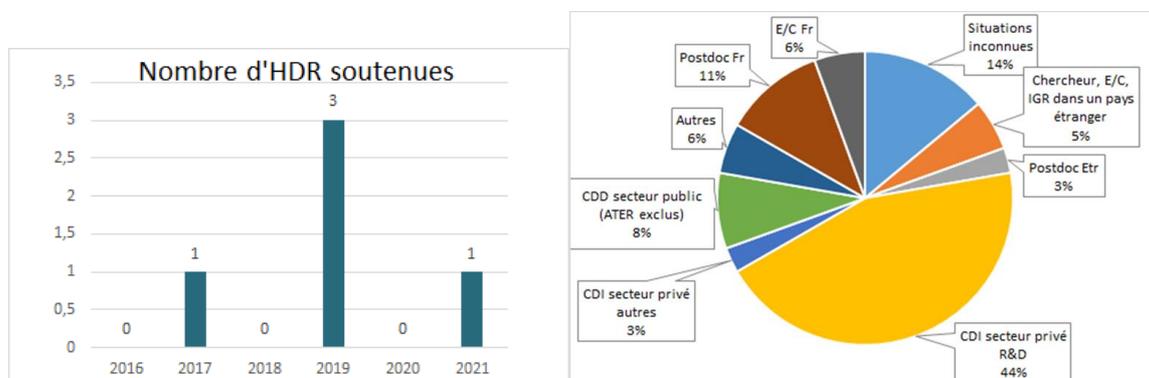


Fig. 1.8 : Nombre d'HDR soutenues

Fig. 1.9 : Devenir des docteurs

### 1.3.3. Parité, bien-être au travail, Hygiène et sécurité ...

**Le LIFAT est soucieux du bien-être** de ses membres (la parité est abordée en section 1.1.1.) et différentes procédures et démarches qualité ont été mises en place pour aller dans ce sens. Chaque année en octobre une réunion conviviale d'accueil des doctorants est organisée afin de présenter et d'expliquer le fonctionnement du laboratoire, de l'école doctorale et le déroulement des thèses. Un entretien professionnel (avec le directeur du LIFAT) a lieu chaque année pour tous les personnel BIATS (incluant IGR, IGE, techniciens, Postdoc, ...).

Des moments conviviaux sont organisés (sauf interdiction covid) régulièrement au niveau LIFAT et dans les équipes (repas après les AG, séminaires au vert, séminaires d'équipes, repas de fin d'année, galette, ...)

Le LIFAT est confronté à des problèmes de sécurité, notamment pour ce qui concerne le « travail isolé ». Lors des périodes de congé et de fermeture du bâtiment, les personnes désirant accéder au bâtiment doivent indiquer au préalable les périodes précises pendant lesquelles elles envisagent de venir (jour et horaires de présence), afin d'être certain qu'elle ne sera pas seule. Un agenda partagé est également mis en place, afin que chacun soit certain qu'il n'est pas seul dans le bâtiment.

**La crise sanitaire du COVID19** a été correctement surmontée par les membres du LIFAT et aucun problème notable n'est à mentionner (**hormis la forte réduction de mobilités sortantes chez les doctorants**). La plupart des activités du LIFAT ont pu être réalisées en télé-travail (conseils de laboratoire, séminaires, réunions encadrants-doctorants, gestion administrative, ...) et la plupart des serveurs de calcul sont accessibles à distance. Le LIFAT s'est conformé aux consignes ministérielles et à celle de l'université de Tours (notamment pour ce qui concerne les confinements et plan de retour sur site). Avec l'aide du représentant des doctorants au conseil de laboratoire, une attention particulière a été portée sur l'isolement possible de certains doctorants (notamment doctorants étrangers tous juste arrivés en France).

Notons tout de même que la gestion de la crise a eu un impact sur la charge de travail administrative des collègues (budget à revoir, mise en place des consignes, organisation des soutenances, gestion des stages, communication et prise de décision moins efficace en distanciel, ...) et sur le rayonnement (ré-organisation des événements, annulation/report des visites internationales entrantes et sortantes...).

#### 1.3.4. Open data, intégrité scientifique, éthique, ...

En termes **d'intégrité scientifique et d'éthique**, nous pouvons mentionner une forte importance apportée à ces aspects au sein du laboratoire comme en témoigne la volonté de mettre en place une **charte éthique** dans le cadre de la fédération ICVL avec nos collègues du LIFO. Un article expliquant la démarche est disponible sous [HAL](#). La question de la reproductibilité des expériences est aussi souvent abordée et certains collègues prônent l'utilisation de cahier de laboratoire (notamment par les doctorants). Les signatures de complaisance des publications sont donc proscrites, et l'ordre des signatures vise prioritairement à mettre en avant le doctorant (dans la mesure où sa contribution est significative), puis les contributeurs en fonction de leur niveau d'implication. Cette considération de l'ordre n'est toutefois pas adoptée par toutes les équipes, pour les collègues en Bases de Données il est fréquent que ce soit plutôt systématiquement l'ordre alphabétique qui soit appliqué. Dans le cas des signatures avec des collaborateurs d'autres disciplines, les règles sont également adaptées.

**L'utilisation de HAL pour la gestion des publications du LIFAT** est effective et généralisée depuis 2016. HAL constitue l'unique source de référence bien que aucun personnel ne soit dédié ou puisse apporter une aide sur cette tâche. Il en est de même pour **le site web et le GitLab du LIFAT** d'ailleurs. Chaque équipe a donc nommé un référent HAL et un référent KSUP (site web LIFAT) et le directeur doit s'acquitter lui-même de toutes ces tâches pour le niveau laboratoire. Une incitation à la mise en ligne des textes intégraux des articles a été réalisée et a porté ses fruits puisqu'environ 80% des articles scientifiques sont en ligne (open-access).

### 1.4. Synthèse de l'autoévaluation

	<b>Forces</b>	<b>Faiblesse</b>
<b>Interne</b>	Compétences scientifiques des équipes très complémentaires apportant une originalité Très bonne implantation et reconnaissance régionale Intégrité scientifique, stabilité organisationnelle Pas de problème de budget et de ressources matérielles	Manque de personnels de recherche et support Charges administratives lourdes impliquant un manque de temps pour la recherche des seniors <b>Fatigue, sollicitations trop nombreuses</b> Dispersion géographique et inter-composantes
	<b>Opportunités</b>	<b>Menaces</b>
<b>Externe</b>	Positionnement scientifique sur un domaine en pleine expansion Expériences de la pluridisciplinarité CETU ILIAD3 Rattachement/label CNRS apporté par l'ERL ROOT	Forte concurrence (de grosses structures) Secrétariat du laboratoire instable Départ de collègues moteurs (HDR, dispo), <b>Découragement, tension, burn-out, démission</b>

## 2. Domaine d'évaluation 2 : Attractivité

L'attractivité du LIFAT est bonne depuis le précédent contrat. Elle s'est même améliorée avec la participation à 2 projets H2020 et l'implication dans différents réseaux COST européens. Chaque équipe est impliquée dans l'animation nationale de la recherche au travers de GDR, associations savantes et projets ANR. Le rayonnement au niveau régional a déjà été présenté dans la section 0.3.

Les sections ci-dessous reprennent les **données marquantes uniquement en synthétisant les analyses** plus détaillées produites par chaque équipe.

### 2.1. Référence 1. Rayonnement scientifique

#### 2.2.1. Collaborations suivies avec d'autres laboratoires

Le LIFAT a un certain nombre de collaborateurs privilégiés, i.e. des unités de recherche avec lesquelles les collaborations sont régulières, à la fois au niveau régional, national et au niveau international. Ces relations ont permis de développer des travaux de recherche conjoints qui ont été publiés dans des revues et conférences internationales. Plus de détails sont fournies **dans les parties décrivant chaque équipe**. Globalement, les collaborations internationales privilégiées sont à noter avec les unités suivantes. Ces collaborations sont attestées par de nombreux séjours/visites entrants et sortants ainsi que par des publications conjointes.

- Equipe BDTLN : USTBH Alger (Algérie), Université de Ioannina (Grèce), Université d'Uruguay (Uruguay), Université de Sfax (Tunisie), Université de Belgrade (Serbie), Université de Dusseldorf (Allemagne), Université de Varsovie (Pologne), Universités de Bozen-Bolzano et de Bologne (Italie), Université Gaston Berger de Saint-Louis (Sénégal), Universités de Tunis et de Sfax (Tunisie), University College Dublin (Irlande).
- Equipe RFAI : Université de Mascara (Algérie), Universités Milan et Salerne (Italie), Université de Hong Duc et FICT (Vietnam), Université de Tarragone (Espagne)
- Equipe ROOT : université de Laval au Canada, Politecnico di Torino en Italie, HEC Montréal au Canada, université de Antioquia en Colombie, université de St Louis aux USA, école d'Ingénieurs de Monastir en Tunisie, école Polytechnique de Nouakchott en Mauritanie

Sur le plan national, les collaborations sont souvent liées à des projets ANR ou autres projets nationaux. Les collaborations privilégiées sont réalisées avec les unités suivantes :

- Equipe BDTLN : IGN, GREYC (Caen), Telecom Paris, IRIT (Toulouse), CRESTIC (Reims)
- Equipe RFAI : LITIS (Rouen), GREYC (Caen), LIRIS (Lyon), LIPADE (Paris V)
- Equipe ROOT : LAAS (Toulouse), LARIS (Angers), HeuDiaSyC (Compiègne), CERMICS (Paris) LIMOS (St Etienne et Gardanne)

Sur le plan régional et local, les collaborations privilégiées sont à noter avec les unités suivantes : LIFO (Orléans), CESR (Tours), LLL (Orléans-Tours), CITERES (Tours), INSERM U930 IBrain (Tours), CHRU Tours, PRC (UMR INRAE-CNRS-UT Nouzilly), IRBI (Tours), IDP (Orléans), BRGM (Orléans) et PRISME (Bourges).

#### 2.2.2. Conférences invitées

Certains membres du LIFAT sont régulièrement invités pour détailler leurs expertises durant des écoles d'été, des conférences, des journées de GDR ou pour animer des tables rondes.

Interventions en tant que conférenciers invités de membre de l'équipe BDTLN :

- H2020 ARIADNE (workshop final, interne au projet), « Semantic data integration », Florence, décembre 2016
- Workshop Action COST EnetCollect, "Crowdsourcing for language learning", Bolzano, 2017
- Ecole d'été Université de Sophia, « Multiword Expressions in a Nutshell », 2018
- Conférence invitée Atelier CAJOLE de CORIA-TALn, « De la qualité des données à la qualité d'analyse », mai 2018
- Ecole thématique CNRS DONIPAT (Données interopérables pour le patrimoine), « des ontologies, des triplestores », octobre 2019
- Conférence invitée Université de Genève, « Without lexicons, multiword expression identification will never fly », 2019
- Ecole d'été Université de Pavie, « Multiword expressions – the Achilles' heel of natural language processing », 2019
- Conférence invitée à la conférence Handicap 2020, « Communication alternative et augmentée et prédiction de mots : tour d'horizon des technologies sous-jacentes », 2020
- Conférence invitée à la conférence internationale ADBIS/TPDL/EDA 2020, « From source data to data narratives : accompanying users in the way to interactive data analysis », août 2020
- Conférence invitée aux Journées d'étude annoter l'oral (U. Paris Sorbonne), « Prédire ou expliquer la langue : lorsque les spécificités de l'oral et les contraintes applicatives du TAL nous imposent la modestie en matière d'annotation », 2021
- Conférence invitée à EUROPHRAS 2021, « Literal Occurrences of Multiword Expressions : Rare Birds That Cause a Stir », Louvain, septembre 2021
- Ecole d'hiver e-EGC 2022, « Exploratory data analysis: from insights to storytelling », Janvier 2022

Interventions en tant que conférenciers invités de membre de l'équipe RFAI :

- Conférence Débat « [Cosmetic Expert : Intelligence Artificielle & Industrie du Futur](#) : Et si on collaborait... » Webinaire Cosmetic valley - décembre 2020
- Table ronde « [Numérisation et mémorisation du patrimoine](#) », à Tours, Semaine de la mémoire - B2V, septembre 2021.
- Conférence sur l'intelligence artificielle et la médecine, Colloque Croix-Rouge formation, IRFSS CENTRE-VAL DE LOIRE, Tours, novembre 2018
- Conférence « Deep Learning & Graph Representation for Image Analysis ». 10èmes Journées Scientifiques et Techniques du Réseau des Microscopistes INRAE, sept 2021 – Tours
- Conférence "Semantic & interaction: the meeting points between Document Image Analysis and Computer Vision", International Workshop IDAKS 2019. Sienna Italie, Oct 2018.
- Conférence "From image to content". DARIAH Summer school 2017 : Bibliotheca Digitalis, Le Mans, juillet 2017
- Conférence/Débat Avenir de la communauté Ecrit et Documents. SIFED 2017. Nantes Juin 2017.

Interventions en tant que conférenciers invités de membre de l'équipe ROOT :

- Tutoriel et TP en programmation par contrainte, école d'été Jeunes Chercheurs du GDR RO en 2021 à Paris -
- Tutoriel "Integrated Scheduling and routing", International spring school on Integrated Operational Problems - GDR RO en 2018 à Troyes
- Exposés liés aux projets ANR ATHENA et e-VRO, école d'été Jeunes Chercheurs du GT2L en 2017 à Gardannes
- Cours en Ordonnancement, école Jeunes Chercheurs du GDR RO en 2016 à Angers.

### 2.2.3. *Participation à des réseaux nationaux et internationaux*

#### Réseaux internationaux

Les membres du LIFAT se sont impliqués dans de nombreux réseaux sur le plan international :

- IC1207 COST action PARSEME (2013-2017), *Chair of the management committee* du réseau européen PARSEME (*Parsing and Multi-Word Expressions*)
- CEA-MITIC, Centre d'Excellence Africain (CEA) en Mathématiques, Informatique et Technologies de l'information et de la Communication (MITIC), logé à l'UGB (Saint-Louis, Sénégal), financement Banque Mondiale (2015-2018), Membre du Conseil Scientifique.
- TD1210 COST action KNOWeSCAPE (2013-2017), Représentante suppléante de la France au Management Committee. En collaboration avec 26 pays européens et 3 hors Europe
- Working Group PMS : membre de l'advisory board pour l'organisation de PMS Durant les conférences EURO.
- TC-15 de l'IAPR : Principal Chair depuis Juin 2021, Newsletter Editors depuis 2015 (Technical Committee #15 of the International Association for Pattern Recognition).
- Réseau [IDAKS](#) (Institute for Document Analysis Knowledge Science) : membre invité
- UniTwin CS-DC de l'UNESCO : Depuis 2015, représentant local

#### Réseaux nationaux

Les membres du LIFAT participent à des réseaux nationaux, soit en qualité d'élus (CNU, associations), soit en qualité de membres de bureaux ou membres de GDR-CNRS avec une responsabilité :

- CNU 27<sup>ème</sup> section (au moins un représentant par équipe au CNU pendant la période, actuellement trois membres du laboratoire sont au CNU) :
  - Y. Kergosien membre élu du CNU depuis oct. 2015
  - A. Soukhal membre élu du CNU depuis oct. 2015
- Sociétés savantes
  - Association Evolution Artificielle (EA), Présidence
  - Institut Fédératif de Recherche sur les Aides Techniques pour personnes Handicapées (IFRATH), membres du conseil d'administration, vice-président chargé de la prospective, Président 2016-2019.
  - Association Française pour la Reconnaissance et l'Interprétation des Formes (AFRIF) secrétaires, membres du CA
  - Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF) : vice-présidence
- Groupes de Recherche
  - GDR-CNRS MADICS : responsables d'actions prospectives
  - GDR-CNRS MAGIS : responsable d'action prospective
  - GDR-CNRS RO : directrice adjointe, membre de son conseil scientifique, responsables d'axes
  - Groupe de Travail ROSa du GDR RO : co-fondateur et co-animateur de ce Groupe de Travail
  - Association GRCE (Groupe de Recherche en Communication Ecrite), secrétaire, membre du CA.
  - Groupe de Travail Visualisation d'informations, interactions, fouille de données (GT-VIF) de l'association EGC

#### 2.2.4. Organisation de conférences, colloques, journées, séminaires

Deux conférences internationales (Matheuristics 2018 et GBR 2019) ont été organisées à Tours et à Blois pendant la période, d'autres manifestations ont été co-organisées par les membres du LIFAT. Les plus significatives sont listées ci-dessous :

Au niveau international

- [Matheuristics 2018](#) (30 participants internationaux, recherche opérationnelle)
- [GBR 2019](#) (50 participants internationaux, Reconnaissance de formes et graphes)
- [ICPRS 2019](#) (50 participants internationaux, Reconnaissance des formes et images)
- [Track Video Processing for Human Behavioral Analysis de l'ACM](#) Symposium on Applied Computing (SAC 2020 et SAC 2021)
- 13th Workshop on Multiword Expressions (MWE 2017)
- Second International Workshop on Semantic Web for Cultural Heritage (SW4CH 2017)
- Nineteenth International Workshop On Design, Optimization, Languages and Analytical Processing of Big Data (DOLAP 2017)
- Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)
- Third International Workshop on Semantic Web for Cultural Heritage (SW4CH 2018)
- First international workshop on Qualitative Aspects of User-Centered Analytics (QAUCA 2019)
- Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019)
- Joint Workshop on Multiword Expressions and Electronic Lexicons (MWE-LEX 2020)
- Second international workshop on Qualitative Aspects of User-Centered Analytics (QAUCA 2020)
- Geographical Information Analysis Track (GIA) @ SAC2020
- Geographical Information Analysis Track (GIA) @ SAC2021
- International Joint Workshop on "Semantic Web and Ontology Design for Cultural Heritage" (SWODCH 2021)
- International Workshop On Design, Optimization, Languages and Analytical Processing of Big Data (DOLAP 2021)

Au niveau national :

- [Ecole d'été France Excellence Chine 2017](#)
- [EGC 2022](#)

De nombreuses journées de GDR ou projets ANR ont également été organisées par des membres du LIFAT (voir description de chaque équipe).

#### 2.2.5. Participation à des comités scientifiques, direction de collections, travaux d'édition

##### Comités scientifiques

Les membres du LIFAT participent aux comités scientifiques des grandes conférences de leur domaine, au niveau national et international. Cela peut consister en une participation active à la sélection des articles présentés (travaux d'arbitrage), en l'organisation de tracks ou de sessions dans un domaine précis : Extending Database Technology (EDBT), International Conference on Conceptual Modelling (ER), IFIP TC13 Conference on Human-Computer Interaction (Interact), European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD), International Conference on Very Large Databases (VLDB), Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), International Semantic Web Conference (ISWC In-Use), Extended Semantic Web Conference (ESWC Ressources), The Web Conference (Semantics and Knowledge), International Conference on Data Warehousing and Knowledge Discovery (DAWAK), International Workshop on Data Warehousing and OLAP (DOLAP), International conference on modeling, optimization and simulation (MOSIM), Project Management and Scheduling (PMS), 2CoDIT, ICFDNS, Matheuristic, SCALCOM, VECoS, Multidisciplinary International Scheduling Conference Théory & Applications (MISTA), Congrès national de la société française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF), Conférence francophone Gestion et Ingénierie des Systèmes Hospitaliers (GISEH), International Conference on document Analysis and Recognition (ICDAR), International Conference on Pattern Recognition (ICPR), Workshop on Document Analysis System (DAS) et Graph Based Representation (GBR), Colloque International Francophone sur l'Écrit et le Document (CIFED), Extraction et Gestion de Connaissances (EGC), Congrès national sur la Reconnaissance de Formes et l'Intelligence Artificielle et Perception (RFIAP), etc.

Pour beaucoup de conférences, la participation d'un membre à un comité scientifique est renouvelée tacitement d'une édition de la conférence à l'autre

Les membres du LIFAT participent à des comités éditoriaux de revues scientifiques internationales :

- *European Journal of Operational Research* (EJOR) – Elsevier depuis 2015, Editorial board.
- *Journal of Scheduling* (JoS) - Springer depuis 01/2015, Associate editor.
- *Decision Support Systems* – Elsevier depuis 2015, Editorial Board.
- *Algorithms* (Q3) depuis 2021, Editorial Board.
- *OMEGA* (Q1) jusqu'en 2017, Editorial Board.
- *IARIA Journal - Journal On Advances in Internet Technology* depuis 2016, Editorial Board.
- Elsevier Internet of Things depuis 2020, Editorial Board.
- *MDPI Journal of Imaging* depuis 2018, Editorial Board.
- Elsevier Pattern Recognition Letters depuis 2016, Editorial Board.
- *IEEE Transactions on Emerging Topics in Computing* depuis 2019, Editorial Board.

- Data and Knowledge Engineering, Elsevier, depuis 2018, Editorial Board
- Infotheca, depuis 2017, Editorial Board
- Language Science Press, depuis 2019, Advisory Board

Les membres du LIFAT participent à la relecture d'articles pour des revues internationales :

- Journaux de rang A ou A\* selon CORE : Decision and Support Systems, ECML PKDD (en 2016 pour la version revue de la conférence), IEEE Transactions on Knowledge and Data Engineering, Information Systems, Practice and Experience, Theoretical Computer Science, 4OR, Computers & OR, Annals of OR, Discrete Applied Mathematics, EJOR, IJPE, IJPR, JORS, Omega, Operations Research, Transportation Science, Transportation Research, Pattern Recognition (A\*), IEEE Transactions on Pattern Analysis and Machine Intelligence (A\*), IEEE Transaction on Image Processing (A\*), Data Mining and Knowledge Discovery (A).
- Journaux de rang B selon CORE : ACM Transactions on Speech and Language Processing, Information Sciences, Pattern Recognition Letters (B), Visual Computer (B), Literary and Linguistic Computing (B).

### 2.2.6. Participation au programme d'Investissements d'Avenir, à des travaux d'expertise, jurys, etc.

Les participations du LIFAT à différents projets investissement d'avenir sont décrites en section 2.4.1

Les membres du LIFAT sont régulièrement sollicités pour participer à des évaluations scientifiques. Il peut s'agir d'expertises de projets à un niveau international (projets européens, normalisation AFNOR XO3A, miroir du working group ISO/TC37/SC4/WG2, ...) ou à un niveau national (ANR, Crédit Impôt Recherche, CIFRE pour l'ANRT, ...).

Participation à des comités d'expertise HCERES comme représentant CNU 27.

## 2.2. Référence 2. Accueil et formation de chercheurs (juniors et seniors)

La mobilité internationale a été fortement impactée par la crise sanitaire du COVID19 interdisant les déplacements durant presque 3 ans depuis 2019. L'impact le plus important pour le LIFAT a concerné les mobilités sortantes des doctorants qui n'ont pas pu avoir lieu. Une analyse de l'accueil et la qualité de l'environnement de travail du personnel LIFAT (doctorants, EC et Biats) hors mobilité a déjà été fournie section 1.3.2. Les données de mobilité du LIFAT sur le contrat sont fournies dans l'annexe A : Extraction de la base LABRI de l'université de Tours sur les activités à l'international de ses unités de recherche afin de leur attribuer un **BQI (Bonus Qualité International) dans leur budget annuel**. Le LIFAT se situe souvent parmi les laboratoires **les mieux dotés de l'université**.

### 2.3.1. Mobilité entrante

#### Accueil de chercheurs juniors

L'accueil d'étudiants doctorants et Postdoc s'est poursuivi de manière similaire au précédent contrat avec **en moyenne 9 postdocs par an au LIFAT** (voir fig. 1.2). L'accueil de Postdocs est directement lié à l'obtention de financements lors des projets APR ou ANR (ou de postes d'ATER considérés dans le décomptage des Postdocs fig. 1.2).

Comme le montre le tableau des mobilités entrantes fourni en annexe A, le cumul des semaines d'accueil de chercheurs étrangers juniors (doctorants + post-doctorants) au LIFAT (avec interruption covid pendant 3 ans) est de **238 semaines soit environ 47,6 mois**.

#### Accueil de chercheurs seniors

Le laboratoire a intensifié les invitations de chercheurs étrangers seniors (voir liste complète en annexe A). Ces invitations se font principalement grâce à un « nombre de mois » mis à disposition par l'Université pour inviter des collègues étrangers en séjours d'au moins un mois. Ces invitations permettent d'établir des collaborations pérennes avec des collègues étrangers. Ces invitations depuis 2016 (avec interruption covid pendant 3 ans) permettent de totaliser **47 semaines soit environ 9,5 mois d'invitations de Professeurs étrangers dont certains de très forte renommée**. Ces visites donnent systématiquement lieu à des publications cosignées ou à des co-encadrements de thèses. Le LIFAT apprécie tout particulièrement ce dispositif, offert par l'Université de Tours, qui sert énormément dans sa politique d'ouverture à l'international.

### 2.3.2. Mobilité sortante

#### Chercheurs juniors

Le LIFAT a mis en place **10 thèses en co-tutelles depuis 2016** durant lesquelles il est prévu, pour la plupart, 50% du temps dans chacun des 2 laboratoires (voir liste annexe A).

Dans le cadre de la formation des doctorants, la mobilité à l'international est fortement encouragée (en dehors des participations aux conférences internationales). **La crise sanitaire a eu un très fort impact sur cette possibilité** habituellement offerte. Sur la période 2016-2021, nous notons que **2 mobilités sortantes** (12 semaines) indiqués dans le tableau fourni en annexe LABRI. **Une vigilance spéciale devra donc être portée sur ce point dans l'avenir** pour relancer ce type d'échanges.

Notons néanmoins que l'organisation d'une mobilité à l'international est fonction du statut du doctorant (plus difficile pour un doctorant en CIFRE, parfois difficile pour un étudiant étranger qui est déjà en mobilité) et de nos collaborations à l'international (tous les sujets de thèse ne sont pas nécessairement propices à une collaboration internationale).

### Chercheurs seniors

Concernant les mobilités sortantes, le LIFAT a accentué significativement la mobilité sortante de ses chercheurs en particulier en ce qui concerne les séjours de longue durée. En plus des déplacements habituels lors des conférences internationales, les EC du LIFAT ayant bénéficié d'invitations dans des laboratoires à l'étranger pour une durée supérieure à 2 semaines sont listés dans le tableau des mobilités sortantes fourni en annexe A. Un cumul de **109 semaines soit 21 mois de mobilité est à noter pour les EC LIFAT depuis 2016** malgré la crise sanitaire.

Des mobilités ont également été réalisées au **niveau national** notamment en organisant des séjours de doctorants dans des laboratoires partenaires et par des visites de longues durées de collègues ayant obtenu des CRCT.

## 2.4. Référence 3. Participation à des appels à projets compétitifs

### 2.4.1. Participation à des projets

Nous présentons le bilan sur les participations à des projets selon qu'ils se situent sur un volet international, national ou régional. La figure 2.1 synthétise les soumissions et acceptations. Le LIFAT a participé à la soumission de **37 projets régionaux, 45 projets nationaux et 16 projets internationaux**. Sur ce contrat, on peut noter un léger déséquilibre entre les équipes au niveau des taux d'acceptation sur les projets nationaux et internationaux.

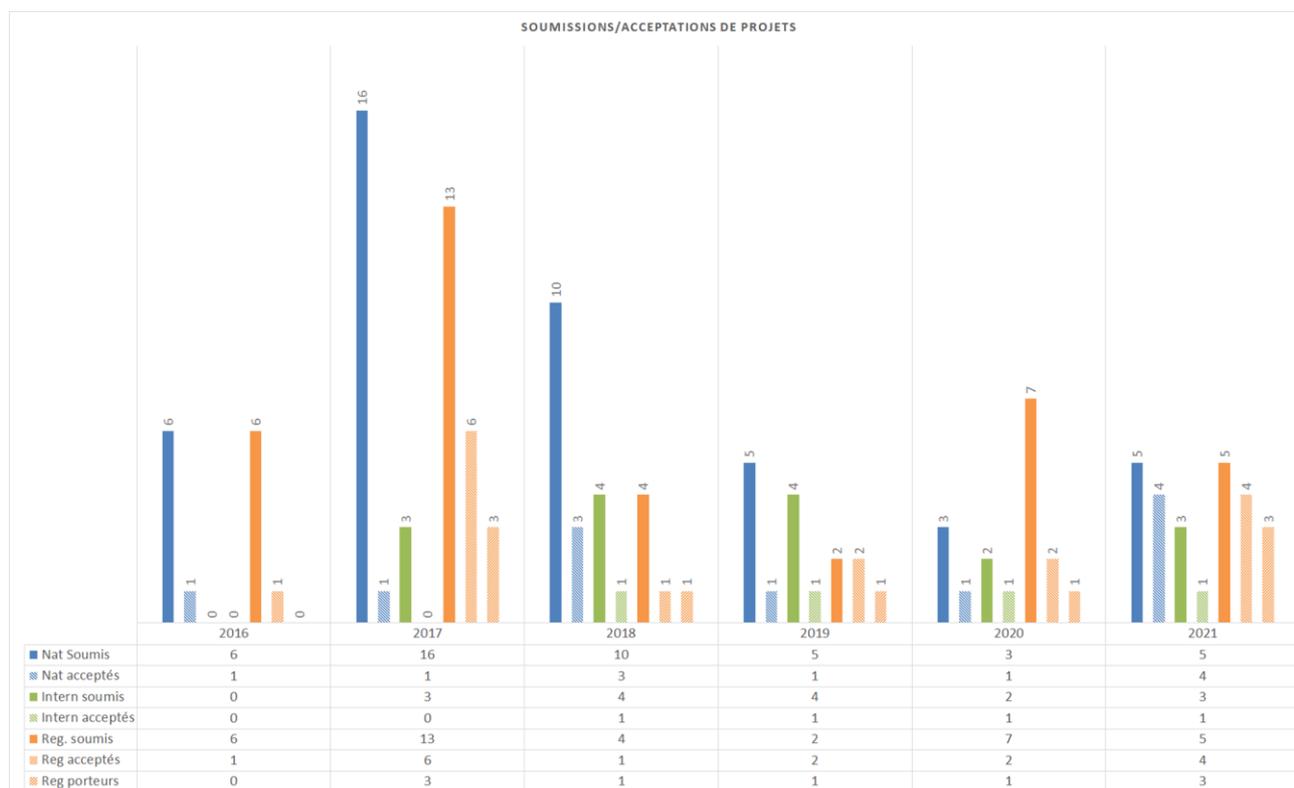


Fig. 2.1 : Taux de soumission et acceptation de projets régionaux, nationaux, internationaux

### Projets internationaux

Le LIFAT s'est clairement inscrit ces dernières années dans une politique soutenue de dépôts de projets à l'international ayant un fort taux de sélectivité (**16 participations à des dépôts**). Ces dépôts rentrent dans le cadre d'appels tels que le programme Européen Horizon 2020 ou les appels à projets bilatéraux ANR internationaux. Le tableau 2.2 présente les projets internationaux acceptés ou

d'importance pour le LIFAT (une demande d'ERC a été soumise par un membre du LIFAT en 2017). Le LIFAT compte soutenir cette dynamique et accroître ses participations aux appels à projets internationaux dans l'avenir.

ARIADNE Plus	Advanced Research Infrastructur	H2020-INFRAIA	LAT, consortium MASA	N	2018	Accepté
ANONYMOUS	Analyse de Videos	H2020, ERC	LIFAT	O	2017	non retenu
Bourse de these VIE	Project 911 - Vietnam Internation	Vied911	Université de Hong Duc	O	2019	Accepté
4CH	Competence Centre for the Cons	H2020	CITERES - LIFAT	N	2020	Accepté
NutriCancer	ML Oncologie	Studium	Inserm UMR1069 Nutriti	N	2021	Accepté
COST UNI DIVE	TALN langues rares	H2020 COST	PARSEME COST	N	2021	déposé
COST GraphML	ML sur Graphes	H2020 COST	Greyc, ...	N	2021	déposé

Fig. 2.2 : Liste des projets internationaux acceptés (+ soumission significative) sur le contrat

### Projets nationaux

Les membres du LIFAT sont impliqués dans le dépôt de très nombreux projets ANR (plus de 40), ainsi que dans des projets financés dans le cadre des « Investissements d'avenir » comme le projet **PIA4 « Loire-Val-Health »** qui vient d'être soumis par l'université de Tours.

L'équipe BDTLN a poursuivi sa participation au projet d'investissement d'avenir [Istex-biosystemique](#) en travaillant sur la détection des entités nommées dans des textes scientifiques et en annotant avec le logiciel CasEN plus de vingt millions de textes de cette bibliothèque numérique. Il y a eu ISTE 2014-2017 avec un budget LIFAT de 111 k€ (LIFAT, Consortium Couperain, Agence Abes, Université de Lorraine et CPU, CNRS). Il y a eu ensuite le projet "Biosystème" en 2017-2018 avec un budget LIFAT de 40 k€ (1 an de Postdoc + stagiaire) avec l'INRAE, ce projet a préparé l'ANR ABLISS.

**Dix projets ANR ont démarré** (25% d'acceptation) dans la période dont 1 pour lequel le **LIFAT est porteur**. Notons aussi, la présence du projet ANR Sircus (porté par INSERM-iBrain sur la détection précoce de l'autisme) sélectionné dans un contexte très sélectif. **Sur ce contrat, on note 1 seul projet ANR accepté sur 16 soumissions avec le LIFAT comme porteur** (au lieu de 3 sur le précédent contrat) au niveau national et international (soit un **taux de réussite inférieur à 6% contre 21% de réussite comme porteur au niveau régional**). Une raison de ce faible taux est la difficulté pour certaines thématiques de se faire identifier car il n'y avait pas de comité d'évaluation scientifique (CES) dans ces thématiques, le CES sur l'IA est très général et ne correspond pas à toutes nos soumissions. Nous avons l'intention de participer à un groupe de travail organisé par notre université avec des porteurs ayant eu des ANR acceptés afin de nous améliorer sur ce point.

Soumis Hors contrat	PoNoSod	Planif soins a domicile	ANR JCJC	CIRRELT Montreal	O	2013
	ATHENA	Optimisation pour l'étude intégr	ANR	AS-CNRS, LIMOS, Heudia	O	2013
	FIBRATLAS II et III	Segmentation et visualisation du	ANR	UMR 930	N	2014
	PARSEME-FR	Syntactic parsing and multiword	ANR	Alpage, LIFO, LIF, LIGM	N	2014
	DEMOCRAT	TAL - DÉScription et MODélisation	ANR	LL, LATIS, LILPA, ICAR	N	2015
	e-VRO	Electric Vehicle Routing Optimiz	ANR JCJC	Wis University (USA), Uni	O	2015
	Biosystème ISTE	Biosystème : recherche de ré	Invest. Avenir	INRA	N	2016
	1 MOBIKIDS v2	Analysis of the daily mobility and	ANR	ESO Rennes, PACTE Gre	N	2016
	2 TALAD	Analyse et traitement automatiq	ANR	AGORA, ERTIM (INALCO	N	2017
	3 ABLISS	TAL et bio-informatique	ANR	projet ANR Denis Maure	N	2018
	4 MODATHOM2	Modèle explicatif de la fabrique	ANR	LAT-CITERES, EFEO, LMJ	N	2018
5 SESAMES 2	Sémantisation Et Spatialisation d	ANR	CITERES (LAT), MAP	N	2018	
6 BI4People	BI for the people - Le décisionne	ANR	ERIC (Lyon), ELICO (Lyon	N	2019	
7 VECCAR	Analyse de cartes géographiques	MSH national	MSHCVL	N	2020	
8 AAC4ALL	Aide à la communication pour pé	ANR	LIFAT	O	2021	
9 CodeGNN	GNN Convolution et Decimation	ANR	GREYC, LITIS, LIFAT	N	2021	
10 SIRCUS	Autism + ocUlo-pupillometric	ANR PHRC	IBRAIN, CHRU, LIFAT	N	2021	
11 Materi.A.net	acquisition et classification de dc	ANR	ILIAD3 sans LIFAT	N	2021	
Lore Val Health	Modele Humain et Animal pour l	PIA4	UT, UO, INRAE, INSERM	N	2022	

Fig. 2.3 : Liste des projets nationaux acceptés (ou finalisés en bleu) durant le contrat

### Projets régionaux

Pas moins de 16 projets avec un financement régional ont démarré dans la période 2016-2021 avec des montants pour le LIFAT très variables (montant moyen autour de 80k€), allant de 10 k€ à plus **de 800 k€ sur 10 ans pour l'ARD Junon** (porteur BRGM). Les projets ARD font partie du dispositif « [Ambition Recherche Développement Centre-Val de Loire](#) ». JUNON est un ambitieux programme de développement d'un pôle de recherche numérique sur l'environnement continental (agricole, urbain, forestier et fluvial), en région Centre-Val de Loire. Ce pôle vise à concevoir des services numériques pour améliorer le suivi et la compréhension de l'environnement, pour une meilleure gestion des ressources naturelles.

Sur 37 projets déposés, les 16 projets acceptés représentent un taux de réussite de 43 %. Les partenaires de ces projets sont variés, avec toutefois une prédominance du CHRU, du CESR, de l'INRAE, du LIFO et du PRISME. **Plusieurs projets se font également en partenariat avec le CETU ILIAD3.**

1	DANIEAL 2	Détection et Analyse des Nerfs d	APR IR	Prisme, CHRU	N	2016
2	SmartLoire	Organisation de séjours de touris	APR IR	Agence Départemental	O	2017
3	DistriChimio	Ordonnancement et distribution	APR IR	CHRU, Eticsys	O	2017
4	NEURO2CO	Neeuro-biologie des	APR IR	INRA PRC, beauval	N	2017
5	FIBRAVASC	Classification automatique de pr	APR Post-doc	CHRU Tours (service d'h	O	2017
6	RAVIOLI	Reconnaissance Automatique de	APR IA	LLL, Prisme, LIFO	N	2017
7	PubArchSW	Publication de données archéolo	APR Post-doc	CITERES, ARD Ipat	N	2017
8	VISIT	Visite Interactive et Touristique	APR IR	CESR, ILIAD (pour 85K€)	O	2018
9	Equipement GPU	Serveur GPU	APR IA +FEDER	Cascimodot	O	2019
10	Ethnos	Ethique de la numérisation et de	APR IA	LIFO LIFAT	N	2019
11	DESHERBROB	Détection et intervention roboti	APR IR	PRISME, FRADEM, AUTC	N	2020
12	JUNON	Jumeau Numerique	ARD	BRGM, Univ Orléans, IN	N	2020
13	SQVALD	NTIC et IoT pour l'organisation de	APR IR	Oncocentre CHRU	O	2021
14	AssistMobDV	Assistant mobile quotidien intell	APR IA	living lab	O	2021
15	DOING	Données Intelligente	APR IA	LIFO, LIFAT	N	2021
16	FuturIndus	Industrie du Futur	APR IA	LIFAT GREMAN Calipsy	O	2021

Fig. 2.4 : Liste des projets régionaux acceptés sur le contrat

## 2.5. Référence 4. Equipements lourds et compétences technologiques.

### 2.5.1. CETU ILIAD 3

Le LIFAT est fier d'avoir largement contribué à la création du CETU ILIAD3 présenté dans le portfolio. Un Centre d'Expertise et de Transfert Universitaire (CETU) est une composante de l'université de Tours (comparable à une spin-off intégrée) dédiée à la valorisation de la Recherche dans un domaine particulier et **devant s'autofinancer** (y compris pour les salaires) en offrant des services de prestation et conseil R&D. Le CETU ILIAD3 est spécialisé dans le domaine de l'ingénierie logicielle en traitement d'images et numérisation 3D.

Composé d'ingénieurs et ingénieurs-chercheurs, il propose des services pour le compte de laboratoires de recherche (dont le LIFAT), d'organismes privés ou publics, de PME ou de grands groupes dans le cadre de collaborations public/public, privé/public ou de prestations de services incluant des développements technologiques spécifiques et innovants.

Le CETU ILIAD3 représente donc un véritable centre de compétences technologiques pour le LIFAT. Dans un sens, il peut apporter des ressources humaines de développement technologiques et des ressources matérielles en réalisant des prestations pour le LIFAT. Dans l'autre sens, il bénéficie des expertises scientifiques des membres du LIFAT qu'il peut faire fructifier pour produire des prestations ou des services se basant sur des innovations LIFAT.

Les matériels disponibles au travers de ILIAD3 sont :

- Stations de calcul graphique GPU
- Scanner Laser 3D Faro
- Scanners Optiques 3D GoScan/Faro
- Dispositifs d'interactions (Kinect, LeapMotion, etc...) et de visualisation 3D (hologramme)

### 2.5.2. Serveur de calculs CPU et GPU

Localement, les membres du LIFAT ont accès aux ressources de calcul suivantes mutualisées :

- Serveur LETO installé dans le centre de calcul régional cascimodot (cf figure 2.5)
- **Le projet Super IA déposé à la Région par le LIFAT a permis d'obtenir en 2020 un financement de 100K€ pour l'achat de 3 GPU AMD V100 - 256 Go (inclus dans Leto)**
- Projet Région Lumineux : 2 CPU : Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz, 8Go RAM, 2 GPU : NVIDIA GeForce GTX 1080 Ti pour un total de 22Go mémoire
- Serveur GPU EPU : 2 CPU : Intel(R) Xeon(R) Silver 4214 CPU @ 2.20GHz, 32 Go RAM, GPU : Quadro RTX 8000 avec 48Go mémoire
- Portable GPU RTX3000 pour démonstration LIFAT
- Salles machines Polytech :
  - 16 PC salle TP système (pour les cours de programmation GPU sur linux et Traitement des langues

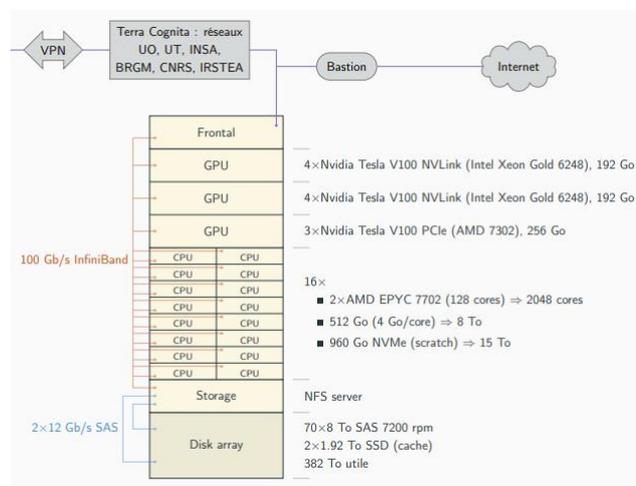


Fig. 2.5 : Serveurs de calcul Cascimodot

GPU sur Windows) : Intel Core i7-10700 (8 cœurs),  
Nvidia Geforce RTX2060 Super (8Go de mémoire)

- 16 PC Salle TP Unix B (en cours d'acquisition, pour les cours de parallélisme sur CPU) : Intel Xeon W-2245 (8 cœurs HT, 3.9Ghz), GPU : NVIDIA T1000 (4Go de mémoire)
- PC (pour projets PRD en GPU/Windows) : Intel Core i7 9700 (8 cœurs), GPU : Nvidia Geforce RTX2080 Super (8Go de mémoire)

### 2.5.3. Autres plateformes

Le LIFAT dispose également d'un accès à :

3. Plateforme d'imagerie biomédicales [PIXANIM](#)
4. Matériels d'acquisition 3D du CETU ILIAD3 et IPAT
5. Salle immersive LIFAT-MSH (financée en 2022 par le CPER Valopat (125K€))

## 2.3. Synthèse de l'autoévaluation

	Forces	Faiblesse
Interne	Participation à de nombreux projets nationaux et régionaux - Augmentation des participations à des projets internationaux sélectifs Nombreuses visites entrantes (juniors et seniors) Nombreuses visites sortantes (seniors) CETU ILIAD3	Baisse des mobilités sortantes des doctorants (liée à la crise sanitaire)
	Opportunités	Menaces
Externes	Participation au dépôt PIA4 et ARD Junon. Nombreux partenariats internationaux. Demande sociétale forte dans les domaines de la science des données. Nombreuses sollicitations d'entreprises, en particulier en TAL et IA.	Faible taux de succès comme porteur à l'ANR et au niveau international mais il a augmenté récemment (1 ANR acceptée sur 16 dépôts comme porteur – taux de réussite < 6% sur ce contrat, peut-être dû à l'étiquette EA). Demandes sociétales fortes en informatique : soit augmentation des charges d'enseignement du fait de l'ouverture de formations sans aucune création de poste ; soit enseignements d'informatique assurés par des non informaticien.ne.s ce qui empêche de demander des postes en informatique. Manque de moyens à l'université et non association à autre organisme de recherche (CNRS ou INRIA) : manque de chercheur permanent et absence totale d'ingénieur permanent en support de la recherche.

### 3. Domaine d'évaluation 3 : Production scientifique

#### 3.1. Au niveau du laboratoire

##### 3.1.1. Production d'articles

Les publications du LIFAT sont systématiquement rentrées dans HAL (le [LIFAT](#) et chaque équipe a une collection) : [BDTLN](#), [RFAI](#) et [ROOT](#). La page d'accueil [HAL-LIFAT](#) permet de visualiser les productions réalisées en collaborations internationales sur une période donnée (cf fig 3.1 pour 2016-2021)

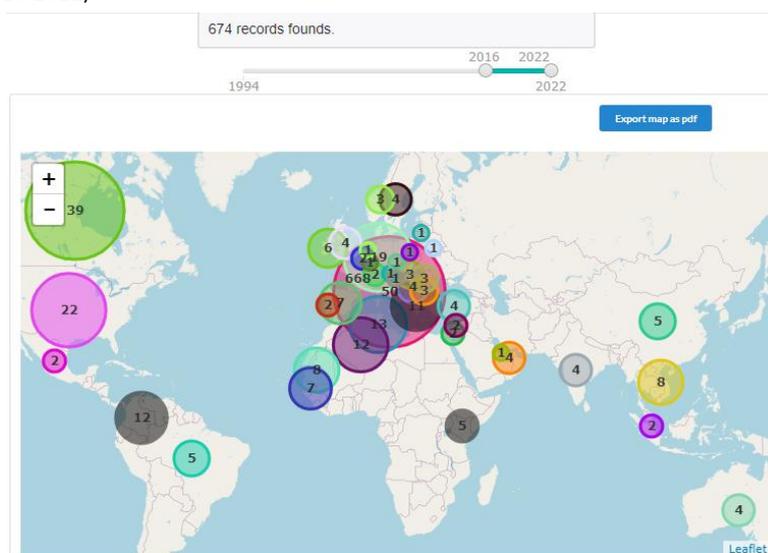


Fig. 3.1 : Visualisation HAL des articles à l'international du LIFAT

Chaque équipe détaille individuellement (dans les sections ci-dessous) la qualité de sa production scientifique. Nous proposons ici uniquement des éléments de synthèse correspondant à une vision globale de la production scientifique du LIFAT.

Globalement, la production scientifique par rapport au contrat précédent a **légèrement augmenté, avec une progression importante dans toute les équipes, concernant la qualité des revues et conférences** dans lesquelles les membres du LIFAT publient. Les publications sont très majoritairement internationales dans des revues référencées. Les publications dans des revues non référencées (ACLN) sont désormais très rares.

On note également une **très forte proportion de publications collaboratives**, avec une forte proportion en ouverture à **l'international (plus d'un tiers (39.8%)** des publications en revue ACL et en conférences internationales (C-ACTI)).

Le nombre de publications en revues internationales référencées (ACL) a augmenté ces dernières années (une moyenne de 18 publications par an entre 2011 et 2016 et une **moyenne de plus de 23.6 publications par an** entre 2016 et juin 2021).

Nous communiquons dans des conférences internationales (C-ACTI) et nationales (C-ACTN) réputées dans nos domaines. Une partie des fluctuations des communications s'explique par la périodicité de certaines conférences et par la crise sanitaire. Dans certains domaines, il y a très peu de conférences nationales.

Sur le plan international, les conférences les plus sélectives et réputées ne sont pas nécessairement indexées dans les grandes bases bibliographiques mais elles sont néanmoins bien identifiées, et il n'est pas toujours simple d'y avoir des communications acceptées. Une étude qualitative par domaine et donc par équipe paraît donc plus judicieuses ici.

Sachant qu'une même publication peut être à la fois en collaboration internationale, nationale, inter-équipes, en collaboration avec le LIFO ou avec une autre unité de recherche de l'Université de Tours, on obtient les chiffres suivants :

	Revue/Journaux (ACL)		Conférences/workshops	
% co-auteurs à l'international	BdTLn= 50% – rfai = 34% - root = 64%	<b>49.3%</b>	BdTLn = 34.7% – rfai = 59,5% - root = 41%	<b>30.3%</b>
% co-auteurs nationaux	BdTLn = 47.6% – rfai = 26% - root = 32%	<b>35.2%</b>	BdTLn = 36.5% - rfai = 38% - root = 23%	<b>27%</b>
avec doctorants	75	<b>52.8%</b>	247	<b>64.5%</b>

inter-équipes	3	2.1%	13	3.4%
articles en open-access	108	76%	303	79.1%
Total	142	23.6/an	383	63.8/an

Fig. 3.2 : Statistiques de publications au niveau du LIFAT

Le niveau de publication entre membres du laboratoire est assez diversifié avec des membres à très fort potentiels de publication : 5 EC totalisent chacun plus de 10 articles en ACL et 3 EC ont publiés chacun plus de 30 articles ACTI sur le contrat. Le LIFAT ne compte que 2 membres non publiants sur le contrat. Une analyse plus fine est disponible dans les rubriques liées à chaque équipe.

### 3.1.2. Production d'instruments, de ressources

Les membres du LIFAT contribuent à leur environnement social, économique et culturel par la création de **prototypes et logiciels**, issus de leurs activités de recherche. Notons que ce sont les membres du LIFAT (doctorants mais aussi enseignants-chercheurs) qui contribuent à l'élaboration de logiciels prototypes, permettant de prouver les concepts, avec parfois le soutien de stages de Master ou d'élèves ingénieurs. Le LIFAT **ne dispose pas d'ingénieurs de recherche ou de permanents dédiés à ces fonctions** de développement. La partie 'développement industriel' peut être confiée à un ingénieur recherche, qui sera impliqué dans le projet, ou à une entreprise privée partenaire, qui pourra assurer la maintenance (corrective et évolutive) du logiciel réalisé.

Nous pouvons souligner une **nette évolution des pratiques de mise à disposition du code et des données** des doctorants (et E/C), de leurs protocoles expérimentaux sur ce contrat comme le montrent les données fournies dans le fichier de productions et d'activité LIFAT (plus de **40 logiciels et 25 bases de données ou corpus**).

Les pages web suivantes donnent accès à certaines de ces ressources :

- [Ressources RFAI](#), [Ressources TLN](#) (BdtIn)
- Travaux inclus dans la librairie [NetworkX](#) très utilisée pour la manipulation de graphes en Python (autres Gits en lien avec cette thématique : [Github1](#) [Github2](#) [Github3](#)) .
- Un [Gitlab LIFAT](#) a récemment été mis en place. Pour des raisons de sécurité, il n'est pour l'instant accessible que par VPN mais la visibilité des dépôts publics est à l'étude
- Voir liste des dépôts Git inclus dans le fichier des données de productions et d'activités.

## 3.2. Equipe BDTLN

### Plan

#### Préliminaires

- Membres et organisation
- Objectifs scientifiques et structuration
- Réponses aux recommandations de l'HCERES précédentes
- Axe « Traitement automatique des langues et interactions »
- Axe « Intelligence des données et des services »
- Axe « Exploration et analyse interactive de données »
- Action inter-axe « Extraction de connaissances du Web des données »

#### Référence 1

- Analyse de la qualité / quantité, co-productions inter-équipes de l'unité et avec acteurs externes
- International : originalités, projets et relations académiques pérennes
- National et régional : projets et relations académiques pérennes
- Rayonnement et attractivité

#### Référence 2

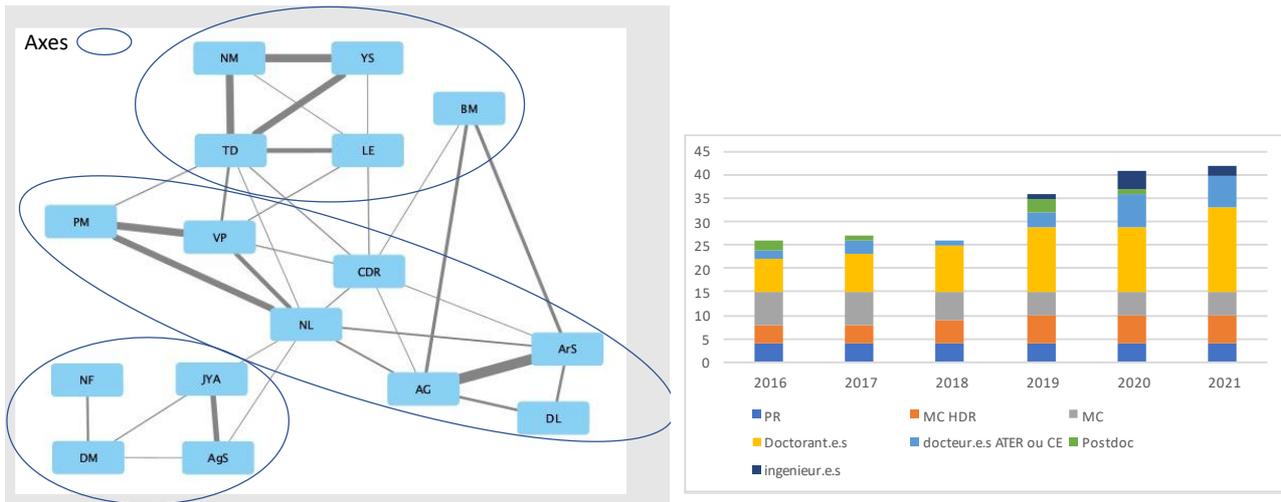
- Répartition, accueil de nouveaux collègues, productions des doctorants et post-doctorants

#### Bilan des forces et des faiblesses - SWOT

### Préliminaires avant Référence 1 : membres, organisation, recommandations précédentes de l'HCERES, bilans des travaux scientifiques par axe

#### Membres et organisation

<b>Responsables d'équipe 2022-2026</b>	Une élection sera organisée à l'automne 2022
<b>Responsables d'équipe 2018-2022</b>	Béatrice MARKHOFF (MC HDR) et collectif
<b>Responsables d'équipe 2016-2018</b>	Arnaud GIACOMETTI (PU) et Laurent ETIENNE (MC)
<b>Membres permanents en mars 2022</b>	Jean-Yves ANTOINE (PU), Thomas DEVOGELE (PU), Cyril DE RUNZ (MC-HDR), Nathalie FRIBURGER (MC), Arnaud GIACOMETTI (PU), Nicolas LABROCHE (MC-HDR), Dominique LI (MC), Patrick MARCEL (MC-HDR), Béatrice MARKHOFF (MC-HDR), Nizar MESSAI (MC), Denis MAUREL (PU), Verónika PERALTA (MC), Yacine SAM (MC), Arnaud SOULET (MC-HDR) [+ Agata Savary (MC-HDR) jusqu'en août 2021]



L'équipe a la particularité d'être majoritairement **située à Blois (12 permanents sur les 15)**. Il y avait à la fin du contrat précédent 4 PU et 11 MC dont 4 HDR. Lors de ce contrat, deux collègues nous ont quittés : un jeune MC souhaitant s'établir en Bretagne (Laurent Etienne) en 2019 et une MCF-HDR ayant obtenu un poste de PU à Saclay en 2021 (Agata Savary). Par ailleurs, un collègue MCF-HDR a été recruté en 2019 (Cyril De Runz), un MC a soutenu son HDR en 2019 (Arnaud Soulet) et nous recruterons pour la rentrée 2022 un.e MC en remplacement d'Agata Savary. Nous serons alors 15 comme en 2016 mais avec depuis 2020 un PU se consacrant exclusivement à présider l'université de Tours (Arnaud Giacometti). Notre capacité à encadrer des thèses reste bonne : 5 MC-HDR au lieu de 4 en 2016, sachant que 4 des MC actuels sont en mesure de soutenir leur HDR pour le prochain contrat (Dominique Li, Nizar Messai, Verónica Peralta et Yacine Sam). Au total ces dernières années l'équipe comprend *plus de 35 personnes* avec les doctorant.e.s et les collègues en CDD (ATER, Postdoc et ingénieur.e.s sur projets). Il y a eu 131 mois de postdoc ou ingénieur payés sur projet sur la période, ce qui est un apport important. Ces collègues et les doctorant.e.s ont également un effet favorable à la **parité femme-homme** : si l'équipe a compris sur la période 2016-2021 **0%** de femme PU et **36%** de femmes MC, soit **26%** de titulaires, il y a eu **27%** de femmes en thèse (29% parmi les thèses soutenues, 28% parmi celles en cours et **16%** parmi celles abandonnées) et il y a cette année **45%** de femmes en contrat d'ATER, postdoc ou ingénieur.

L'équipe a été dirigée deux ans par Arnaud Giacometti qui s'est ensuite investi dans la direction du LIFAT (directeur adjoint) et a été remplacé par Béatrice Markhoff, en collaboration avec un collectif de collègues sur des aspects précis (communication web, animation scientifique, analyse de la production scientifique...). L'équipe se réunit une fois par mois au minimum (l'ordre du jour et le compte-rendu sont publics) : le partage équitable des ressources est régulièrement discuté en veillant à maintenir écoute et respect, toutes les informations liées aux situations personnelles et aux travaux de recherche sont partagées et les décisions engageant l'équipe sont prises collectivement sur la base d'informations rassemblées dans des fichiers partagés que chacun.e met à jour. Tous les membres peuvent suivre les réalisations de l'équipe grâce aux présentations scientifiques au début des réunions mensuelles, au séminaire organisé fin juin, ainsi qu'à travers les « apéros info », ces deux derniers types d'événement étant aussi ouverts aux étudiants (Licence Informatique et Master BDMA) et au public extérieur.

## Objectifs scientifiques et structuration

Les recherches menées dans l'équipe concernent la représentation, la manipulation et l'extraction de connaissances à partir de **données** avec des solutions **centrées utilisateur**. Elles se déclinent selon (1) la structure des données, (2) le type de traitement et (3) l'application visés. Côté **structure de données**, le spectre va des ressources en langues naturelles aux bases de données relationnelles et multidimensionnelles, en passant par les formats spécialisés des ressources spatio-temporelles, semi-structurées du Web et de la représentation des connaissances. Concernant les **traitements**, ils visent d'une part la constitution de ressources et connaissances respectant des critères de qualité et d'autre part des méthodes et algorithmes d'analyse et extraction de connaissances. Les **applications** quant à elles sont tournées vers la société : environnement, emploi, journalisme, patrimoine culturel, santé et handicap. L'équipe s'était organisée en 2016 en 3 axes :

- **Exploration et analyse interactive de données** : l'objectif est de développer des outils d'analyse centrés utilisateurs. D'une part en définissant des langages et algorithmes pour l'extraction de motifs, en particulier des algorithmes anytime pour prendre en compte au plus tôt les préférences d'un utilisateur dans un processus d'extraction, et d'autre part par la modélisation et la découverte, à partir du processus d'analyse exploratoire, des usages, des attentes et des préférences des utilisateurs. Par la qualité de ses publications, cet axe est reconnu tant au niveau national qu'international, que ce soit en fouille de données ou en entrepôts de données.
- **Intelligence des données et des services** : cet axe s'intéresse aux données et aux services d'un point de vue sémantique, pour intégrer des connaissances depuis des sources hétérogènes, composer des services à partir de descriptions sémantiques et plus généralement concevoir des systèmes pour exploiter des données et des services complexes, en particulier des données spatio-temporelles. Ses travaux sont menés dans le cadre de nombreuses collaborations et projets fortement interdisciplinaires, à l'échelle régionale (projet SmartLoire), nationale (deux projets ANR) et internationale (cotutelles, deux projets H2020).

- **Traitement automatique des langues et interaction** : cet axe est dédié aux traitements des langues naturelles avec une forte sensibilité à la spécificité langagière. Il a une grande visibilité nationale et internationale pour sa production de ressources et d'outils pour le TAL, sa coordination de l'action COST PARSEME sur la détection d'entités polylexicales et ses contributions à l'ANR PARSEME\_FR. Il est en première ligne de l'animation de collaborations régionales (direction de l'ICVL depuis sa création) et mène en outre une réflexion éthique sur l'impact des technologies numériques.

Ces axes ont naturellement développé de **multiples collaborations**, cette **complémentarité entre membres de l'équipe** étant illustrée par le graphe précédent (un arc entre deux permanents - représentés par leurs initiales - signifie qu'ils sont co-auteurs d'au moins une publication sur la durée du contrat, l'épaisseur de l'arc étant fonction du nombre de publications communes). En particulier, la fouille de données du Web sémantique a pris une place importante dans l'équipe au cours de ce contrat, avec 4 articles dans les meilleures conférences du Web sémantique (ISWC et ESWC).

Les **travaux scientifiques** dans l'équipe sont détaillés plus loin par axe et action inter-axe, avant les Références 1 et 2.

Le **projet** que portera l'équipe pour le prochain contrat est dans la continuité du bilan fait ici, avec toutefois l'objectif affirmé d'explorer plus de verrous transverses aux axes (plus d'autres, transverses avec les équipes du LIFAT). Ainsi s'il est clair que nous continuerons à proposer des solutions pour faciliter l'analyse exploratoire de données, optimiser la fouille interactive de données (de types et de formats divers) et l'extraction de connaissances, *nous le ferons plus particulièrement encore* pour le Web des données avec sa dimension sémantique. Nous continuerons à contribuer au Linked Open Data / Web sémantique, aux solutions pour en faciliter l'exploitation et à son exploitation par extraction de connaissances, de même que nous intégrerons encore leurs connaissances dans nos approches de conception, découverte et compositions d'objets complexes, *mais aussi* dans nos propositions de mises en relation d'entités d'intérêts dans des textes. Nous souhaitons conserver les apports des méthodes et outils particuliers pour ces données particulières que sont les textes, discours transcrits et autres produits langagiers, pour les développer au sein du TAL *et également* pour les applications qu'ils présentent en termes d'apprentissage. Les applications de nos travaux resteront centrées sur des aspects sociétaux importants dans lesquels nous sommes déjà bien engagés (patrimoine, emploi, journalisme, mobilité, tourisme, santé et handicap), auxquels s'ajoutera l'environnement à travers les projets de l'ARD Junon (Région Centre Val de Loire) dans lesquels plus de la moitié de l'équipe est impliquée. Deux projets ANR, plusieurs projets régionaux, une thèse académique et deux CIFRE démarrent en 2022 et d'autres projets (ANR et régionaux) ont été déposés, témoins du dynamisme du projet scientifique de l'équipe, actuel et à venir.

## Recommandations de l'évaluation HCERES précédente

### Critère 1 : qualité et production scientifiques

Points à améliorer et risques liés au contexte : **charge en enseignement et en responsabilités collectives**

Le problème majeur rencontré par l'équipe BdTin est la charge en enseignement et en responsabilités collectives portée par ses membres.

Recommandations

Un équilibre en termes d'implication dans la formation, la recherche et les tâches collectives permettrait de maintenir une bonne production scientifique.

#### Réponse en fin de contrat 2016-2021

Les charges restent fortes en responsabilités collectives : co-direction du laboratoire et de la fédération ICVL, présidence de l'université, membre de CFVU, de conseil d'UFR ou d'IUT, de conseil de laboratoire, direction d'équipe et de département, responsabilité de quatre formations – Licence, Licences pro, Master.

En enseignement la pression demeure également : fermeture du Master Erasmus Mundus mais duplication de la Licence d'informatique à Tours, en plus de Blois, évolution du DUT en BUT (!) et perte d'un enseignant actif devenu président.

Pour autant, **l'équilibre s'est renforcé dans le sens de la recherche** avec une augmentation significative du nombre de doctorant.e.s et de projets financés. La production scientifique n'a pas faibli, au contraire, mais c'est au prix d'un niveau d'investissement largement supérieur aux termes du contrat d'enseignant-chercheur, ce qui ne peut pas être pérenne. L'équipe soutient ses membres lorsqu'ils ou elles demandent un **CRCT**, car c'est un bon moyen de rééquilibrer les activités dans le sens de la recherche (7 x 6 mois ont été obtenus dans l'équipe au cours de la période).

Cette menace, que l'équipe avait déjà identifiée comme telle en 2016, continue donc à se manifester. Des moyens de recrutement pourraient la faire diminuer, la démission de responsabilités également mais avec le risque d'être marginalisés dans les instances de prises de décision.

### Critère 2 : rayonnement et attractivité académiques

Points à améliorer et risques liés au contexte : **visibilité nationale et internationale des productions**

Un rééquilibrage des financements en faveur des projets nationaux et internationaux par rapport aux projets régionaux serait salubre.

Recommandations

Il faudrait soutenir un effort de communication afin d'attirer des étudiants en provenance des formations locales. La visibilité de l'équipe au niveau international pourrait être améliorée via des participations à des initiatives H2020 et en encourageant la mobilité. Les ressources libres produites par l'équipe doivent être rendues plus visibles et facilement accessibles.

#### Réponse en fin de contrat 2016-2021

Sur les 20 projets et actions de GDR démarrés dans l'équipe entre 2016 et 2021 il y a 3 projets internationaux (**2 projets H2020 et un projet PHC avec l'Irlande**), 12 de portée nationale (dont **5 ANR**) et 5 de portée régionale. **Le niveau régional ne représente donc plus que le quart de nos actions** et de nos financements non récurrents. La mobilité internationale et nationale a été encouragée, particulièrement lors de **CRCT** (Agata Savary 2017-2018, Patrick Marcel 2018-2019, Arnaud Soulet 2019-2020, Verónika Peralta 2020-2021, Yacine Sam 2021-2022 et Nizar Messai 2021-2022).

Le point sur les étudiants est traité au critère 5. Concernant les ressources libres produites par l'équipe, les dépôts git sont devenus la règle, ainsi que des dépôts dans des répertoires nationaux ou internationaux, voir des sites Web pour aider à l'utilisation. Plus de 40 réalisations associées à des publications scientifiques, mises à la disposition du public, sont recensés sur 2016-2021 (cf. fichier de production). L'équipe joue également activement le jeu des dépôts de documents dans HAL et s'investit dans la réponse à l'AMI « Ateliers de la donnée » portée pour l'ensemble des acteurs de la recherche en Région Centre Val de Loire par la MSH Val de Loire.

#### Critère 3 : interaction avec l'environnement économique, social culturel et sanitaire

Points à améliorer et risques liés au contexte : **interactions avec entreprises de portée nationale**

Les interactions présentent une faiblesse toute relative en ce qui concerne des entreprises dépassant le contexte régional.

Recommandations

Il faut viser une augmentation des coopérations avec des entreprises de portée nationale ou internationale (via des projets ANR par exemple).

#### Réponse en fin de contrat 2016-2021

Sur les 8 thèses CIFRE soutenues ou démarrées entre 2016 et 2021, 4 sont avec des entreprises extra-régionales (parisiennes : deux chez Kalidea-groupe Up, une chez SAP, une chez ADWANTED) et pour les locales comme PETIT-PECOT (groupe Gruau) et Neolink, les travaux se font avec des collaborateurs à Paris, Toulouse, etc. De plus, dans l'ANR MOBI'KIDS il y a des collaborations avec 2 entreprises nationales.

#### Critère 4 : organisation et vie de l'équipe

Points à améliorer et risques liés au contexte : **charges d'enseignement et d'administration**

On constate une augmentation globale des charges d'enseignement et d'administration, préjudiciables à l'activité de l'équipe.

#### Réponse en fin de contrat 2016-2021

Réponse faite au critère 1.

#### Critère 5 : implication dans la formation par la recherche

Points à améliorer et risques liés au contexte : **augmenter le nombre de thèses soutenues**

Le nombre de thèses soutenues est en baisse, peut-être seulement de façon conjoncturelle avec un marché de l'emploi attractif détournant les étudiants d'études doctorales.

Recommandations : **attirer des doctorants**

L'équipe devrait trouver des remèdes pour attirer à nouveau des étudiants.

#### Réponse en fin de contrat 2016-2021

**10 étudiants de notre Master local** se sont engagés en thèse dans la période (3 ont soutenu, 1 a abandonné et 6 sont en cours) et en 2022 deux autres commencent également une thèse (CIFRE). Nos efforts de communication, à travers notamment les projets et stages de L3 et Master 2, les apéros info et séminaires de juin, portent donc bien leurs fruits.

Au total il y a eu 40 thèses (dont celles co-dirigées par des membres d'autres équipes du LIFAT ou d'autres universités) soutenues ou démarrées dans l'équipe dans la période 2016-2021, financées sur projet, bourse ministérielle ou régionale, ou par l'ANRT. Sur ces 40 thèses **17 ont été soutenues** contre 5 sur 2011-2015, 6 ont été abandonnées, 17 sont en cours et de nouvelles démarrent. Ainsi la période à venir promet également d'être riche en soutenances.

#### Critère 6 : perspective et stratégie scientifique à cinq ans

#### Points à améliorer et risques liés au contexte : **accès aux données**

Afin de ne pas restreindre l'accès des données produites par l'équipe aux seuls chercheurs du domaine, une licence ouverte garantissant son libre accès et son utilisation sans restriction financière ou juridique serait bienvenue.

#### Recommandations : **points de convergence des axes**

Les trois axes de recherche sont clairement importants et d'actualité. Le contexte est celui de la gestion et l'exploitation de masses de données dans des environnements ouverts et contraints. Il serait intéressant de faire émerger quelques verrous scientifiques et technologiques dont les solutions tireraient profit des avancées réalisées dans les trois axes.

#### Réponse en fin de contrat 2016-2021

Sur l'attention à accorder à la déclaration d'une licence d'utilisation des ressources mises à disposition du public, l'équipe n'a pas encore adopté de politique précise, certaines ressources listées dans le fichier de production disposent d'une déclaration de licence et d'autres pas, cela reste un point d'attention à travailler pour le prochain contrat.

Verrous dont les solutions tireraient profit des avancées des 3 axes : il en existe, comme par exemple l'exploration interactive de ressources du Linked Open Data par des séquences de requêtes SPARQL et de requêtes en langage naturel (Q/A). Cela reste un point d'attention sur lequel travailler ensemble pour le prochain contrat.

La suite de cette section décrit par axe les travaux scientifiques de l'équipe BdTln.

### Axe « Traitement automatique des langues et interactions »

L'axe « Traitement automatique des langues et interactions » met en avant une recherche en Tal consciente de la langue. En effet, les chercheurs de cet axe souhaitent démontrer dans leur démarche scientifique et applicative un souci très marqué pour le fait linguistique. Le Tal actuel développe des recherches dans une démarche centrée données, que l'on parle d'apprentissage automatique statistique ou d'apprentissage profond. Cette évolution, source d'indéniables progrès en termes de performances, questionne également. D'une part, on voit se développer des recherches où l'attention aux faits de langue est très ténue. On se contente le plus souvent d'optimiser une fonction de coût sans retourner (ou peu) aux données. On retrouve ici le questionnement plus général sur l'explicativité du *machine learning*. D'autre part, on assiste de manière contre-intuitive à un éloignement du Tal d'applications réelles : dans de nombreux domaines, on constate que la communauté se focalise sur des tâches qui ne répondent pas nécessairement aux besoins industriels. La raison principale étant sans doute que les tâches étudiées sont guidées par les ressources existantes et non les besoins. Face à ce constat, et compte tenu du nombre relativement modeste de ses membres, l'axe Tal a mis en place au niveau régional des séminaires de recherche qui, excepté pendant la pause covid sur ce genre d'activité, rassemblent régulièrement des chercheurs (de Tours et d'Orléans), en informatique et en linguistique, et des industriels de la région.

Cette démarche s'accompagne souvent d'une dimension multilingue, non pas simplement dans l'idée d'appliquer nos travaux sur différentes langues, mais de les confronter pour une meilleure modélisation linguistique et une meilleure compréhension des phénomènes. De par différentes collaborations avec des universités étrangères, nous avons ainsi travaillé sur des classes de langues variées, comme le français, l'anglais, le polonais, le serbe, l'arabe et l'allemand, ceci afin d'atteindre une plus grande généralité idiomatique. En particulier, cet intérêt pour la langue nous a conduit à développer, parallèlement à nos applications, des ressources linguistiques diverses (corpus annotés et lexiques), libres, et à participer à des travaux de normalisation dans le cadre du groupe de travail ISO TC37/SC4.

Pour entrer plus dans le détail, les travaux de cet axe peuvent être répartis en quatre thématiques.

**La détection en documents d'éléments d'intérêt : entités nommées et entités polylexicales.** Ce niveau, essentiel pour la recherche d'information, permet la détection intelligente des entités dignes d'intérêt dans le document. Nous avons poursuivi sur ce contrat notre participation au projet d'investissement d'avenir Istex, en travaillant sur la détection des entités nommées dans des textes scientifiques et en annotant par notre logiciel CasEN plus de vingt millions de textes de cette bibliothèque numérique. Toujours dans le cadre de la recherche d'entités nommées, nous avons collaboré avec des universités tunisiennes sur la réalisation d'un logiciel traitant la langue arabe, dans le cadre d'une thèse co dirigée. CasEN et sa version arabe, CasANER, sont en libre disposition et ont été utilisés par d'autres équipes. Nous avons poursuivi aussi nos travaux sur les entités polylexicales, avec en particulier, la poursuite de la coordination du réseau européen COST (PARSEME), qui représente vingt-neuf langues et six dialectes issus de dix familles différentes et la création d'un workshop annuel sur ce sujet. Une ANR (PARSEME-FR) obtenue à la veille du contrat a développé des travaux spécifiques au français, avec la soutenance d'une thèse et une autre en cours. Une autre ANR vient d'être acceptée et débutera au contrat suivant : elle porte sur l'induction de lexiques sémantiques pour l'interprétabilité et la diversité en traitement de textes.

**La mise en relations des mots d'intérêt en documents : coréférence, relations temporelles et prédicats.** L'expertise de l'équipe sur le domaine de la coréférence se poursuit à travers une thèse Cifre (société NextIno) sur la pseudonymisation de documents, qui fait suite aux travaux réalisés dans le cadre de l'ANR TALAD. Dans les deux cas, ces recherches mettent l'emphase sur la résolution d'un type de relation trop ignoré par la communauté scientifique, à savoir la coréférence indirecte. Nos travaux ont montré d'une part que le niveau de performance des systèmes de l'état de l'art chutait sur ce type de relation et d'autre part que la mise en place d'un multi-classifieur statistique ciblant séparément chaque type de coréférence pouvait au minimum égaler les performances des solveurs neuronaux génériques sur la coréférence indirecte. Celle sur les relations temporelles a donné lieu à la création d'un corpus annoté en relations temporelles (en suivant la norme ISO TimeML), dans le cadre d'un financement APR-IA de la Région Centre Val de Loire, le projet ODIL, corpus distribué gratuitement. D'autre part, ce contrat a initié des travaux sur la détection de prédicats

d'intérêt dans les articles scientifiques de biologie (en collaboration avec des chercheurs de l'INRAe), tout d'abord par un projet d'investigation (Bio-Systémique), financé par l'appel Chantiers d'usage de la plateforme Istex, puis par un projet ANR (Abliss) en cours. Cette détection sur un grand nombre d'articles scientifiques permet une fouille de la bibliographie (trop abondante !) sur les récepteurs biologiques.

La thématique de **l'aide aux personnes en situation de handicap**, ancienne dans l'équipe, mais peu développée durant le contrat précédent, est repartie avec le projet PREDICT4ALL, financé par la fondation Bennetot, sur le développement d'un module de prédiction de mots qui cible les différentes erreurs orthographiques, en partant de la prédiction de mots Sibylle développée antérieurement par l'équipe. Ce projet a été soumis à une phase d'évaluation clinique longitudinale ayant assez peu d'équivalent par son amplitude et sa durée longitudinale (1 an). L'étude a montré que l'intégration de la correction avec la prédiction demandait à être encore approfondie pour porter ses fruits, mais que la prédiction, en elle-même, a un impact clair sur la qualité de rédaction de patients paralysés cérébraux et/ou dyslexiques. Ces travaux sont complétés par un projet ANR (AAC4All) qui débute en 2022.

Enfin, nos collaborations avec des chercheurs en linguistique ont rendu possible une thématique de **linguistique outillée**, avec deux projets ANR : le projet RAVIOLI qui s'intéresse à la détection automatique des phrases injonctives, en dialogue oral spontané, par l'étude prosodique de l'intonation ; et le projet TALAD, dont le but est la contribution du Tal à l'analyse de discours, principalement dans le cadre de la nomination.

### Axe « Intelligence des données et services »

Cet axe s'intéresse aux problématiques de traitement des données et services d'un point de vue sémantique en se reposant sur les standards du Web. Les contributions développées couvrent un large spectre à l'image des évolutions majeures du Web, allant du Web sémantique, ou Linked Open Data (LOD), au Web social, des services Web classiques aux services de données, APIs REST et Web des Objets (IoT), ainsi qu'aux paradigmes émergents de sécurisation des données et des échanges (SOLID, Blockchain). Il s'organise autour de trois problématiques : (1) l'interopérabilité sémantique (ontologies et utilisateurs), (2) le calcul de similarités sémantiques et (3) la découverte, composition et déploiement de services et d'objets.

**L'interopérabilité sémantique** passe par la conception et l'utilisation d'ontologies de domaine et le support aux utilisateurs lors de leurs interrogations. Elle est au cœur des collaborations avec des disciplines du patrimoine culturel, qui nous ont amenés à créer en 2015 la série de workshop internationaux Semantic Web for Cultural Heritage<sup>1</sup>. Avec les archéologues de Tours qui pilotent le consortium MASA<sup>2</sup> de la TGIR Huma-Num, plusieurs projets communs relèvent de cette problématique : l'ANR SESAMES (2018-2022), le projet européen H2020 ARIADNE+ (2019-2022) et le projet européen H2020 4CH (2020-2023) dont la responsabilité côté université de Tours est prise en charge dans l'axe. Ces deux projets H2020 visent la construction d'une plateforme collaborative s'appuyant sur une base de connaissances, dans un cas (ARIADNE+, 41 partenaires de 23 pays européens et de 4 pays extra-européens) pour le catalogage et si nécessaire l'intégration de million de jeux de données, dans l'autre (4CH, 19 partenaires de 13 pays européens) pour un centre de compétences européen sur la conservation du patrimoine culturel. Dans les deux cas la représentation des connaissances joue un rôle clé, la stratégie générale étant d'utiliser et d'étendre selon les besoins l'ontologie qui fait référence dans ce domaine, CIDOC CRM, avec laquelle sont également réalisés les travaux menés avec le laboratoire CITERES-LAT dans le cadre du consortium MASA, comme par exemple la plateforme OpenArcheo<sup>3</sup> (ODOCH19a). Les terminologies sont aussi un support d'interopérabilité sémantique : un postdoc financé par la Région CVL sur le projet ArchSW a permis de travailler sur l'alignement de vocabulaires utilisés par les chercheurs avec un thésaurus pivot, nommé PACTOLS<sup>4</sup> (ODOCH19b). De même pour l'ANR SESAMES, les contributions portent sur des ontologies d'une part et des terminologies d'autre part pour croiser les connaissances d'architectes et celles d'archéologues sur le patrimoine bâti, avec pour cœur de questionnement l'explicitation des raisonnements scientifiques justifiant la reconstitution de ce qui a pu exister (Heritage2020). Enfin, la thèse de Marwa Boulakbech a proposé des enrichissements de l'ontologie DATATourisme dans le contexte touristique du projet SmartLoire et une thèse CIFRE a été conçue dont l'objectif était l'interopérabilité sémantique de descriptions d'APIs RESTful à l'aide d'une ontologie des interactions HTTP (IC2020), en visant la spécification de ces APIs avec SHACL (la thèse a bien débuté puis le doctorant y a mis fin pour changer d'entreprise et de statut).

Le côté support aux utilisateurs dans leurs interrogations de graphes de connaissances est traité dans le développement du portail OpenArcheo déjà cité et dans la thèse de Thanh Binh Nguyen (co-dirigée avec Mirian Halfeld Ferrari du LIFO) qui a consisté en la modélisation d'un contexte utilisateur pour l'interrogation d'un graphe de connaissances, dans lequel peuvent être spécifiées des contraintes pertinentes pour l'utilisateur, sur les données récoltées par sa requête. Il a pour cela étendu l'ensemble des algorithmes de réécriture bien connus (Chase, Chase&BackChase, PerfectRef, Xrewrite) avec une solution efficace pour ce défi qu'est le filtrage des résultats d'une requête en fonction d'un contexte utilisateur (SOFSEM18).

Le **calcul de similarités sémantiques** est une opération clé pour la découverte de services, de Smart devices (Web des objets) ou d'autres objets. Des solutions ont été définies et implémentées en exploitant les relations conceptuelles contenues dans les ontologies de domaine pour les services Web (thèse de Ahmed Abid, en co-tutelle avec l'Université de Sfax), ou en exploitant les ressources du LOD pour les services Web mobiles dans la thèse de Nasreddine Cheniki (en coopération avec l'Université USTHB d'Alger), enrichissant ainsi automatiquement des descriptions de services mobiles. Le LOD est également la base d'intégration de

<sup>1</sup> SW4CH 2015, 2017 et 2018 <https://sw4ch2018.ensma.fr/> SWJ <https://content.iospress.com/articles/semantic-web/sw210425> et SWODCH 2021 et 2022 <https://swodch2022.inf.unibz.it/>

<sup>2</sup> <https://masa.hypotheses.org/481#more-481>

<sup>3</sup> <http://openarchaeo.huma-num.fr/explorateur/home>

<sup>4</sup> <https://www.frantiq.fr/pactols/le-thesaurus/>

calculs de similarités sémantiques développées dans la thèse de Fouad Komeiha, en cotuelle avec l'Université Libanaise, rassemblées dans un outil au cœur d'un système de découverte d'entités d'intérêt. La thèse de Clément Moreau (projet Smart-Loire) a également porté sur des mesures de similarité sémantique, entre des trajectoires sémantiques tenant compte du comportement humain (FUZZ-IEEE2021, SAC2021).

Enfin, les problématiques de **découverte, composition et déploiement personnalisés de services et d'objets** ont été travaillées dans plusieurs thèses de cet axe, couvrant différentes facettes de l'intégration de données et services dans le Web et des interactions personnalisées avec l'utilisateur. Dans celle de Ahmed ABID déjà citée (soutenue en 2017), des techniques de classification conceptuelles ont été intégrées pour améliorer l'appariement entre services à des fins de recommandation dans le cadre de composition ou de substitution de services Web de type SOAP. Par la suite ce sont les principes des APIs REST qui ont servi à la proposition de deux contributions : pour l'intégration de données par composition de services de données dans le cadre de la thèse de Marwa Boulakbech (soutenue en 2020) et afin d'assurer la migration et synchronisation fluide d'applications dans différents supports d'exécution dans le cadre de la thèse de Clay Palmeira (soutenue en 2019). Dans la première contribution, il a été proposé de développer des mashups spécifiques au domaine pour qu'un utilisateur novice puisse intégrer des données. La solution est fondée sur un algorithme inspiré de la théorie de la configuration qui utilise les raisonnements possibles dans le Web sémantique pour sélectionner et agréger des services de données en fonction des contraintes de l'utilisateur. Le déploiement, dans un environnement fiable et décentralisé, de compositions de services de données découvertes dans le cadre de la plateforme CART développée précédemment fait maintenant l'objet de la thèse d'Amina Brahem (soutenance prévue en 2022) qui traite de la problématique d'exécution en toute confiance, entre partenaires concurrents, de compositions de services Web. Le besoin de préservation de données privées de partenaires concurrents a été résolu en introduisant des algorithmes exécutant dans la Blockchain les parties privées de processus métiers sous la forme de Smart Contracts. La deuxième contribution orientée services REST, par Clay Palmeira, consiste en une architecture permettant la virtualisation de l'environnement de l'utilisateur vis-à-vis du serveur, rendant possible de migrer ses interactions avec un ou plusieurs services d'un terminal à un autre sans dépendre de la synchronisation côté serveur. Cette architecture, qui exploite le paradigme de Liquid Software, a été mise en œuvre dans une plateforme, CUBE, faisant l'objet d'une initiative de création de startup.

Toujours dans cette thématique de découverte et d'intégration de services, les travaux de thèse de Nasreddine Cheniki ont abouti à la proposition d'une plateforme de découverte et sélection de services Web Mobile pour une utilisation contextualisée en fonction des profils et requêtes utilisateurs. Cela est actuellement étendu dans la thèse de Fouad Komeiha pour inclure tout type d'objet connecté dans la recommandation et répondre ainsi au besoin en services et objets connectés dans des contextes domotiques (Web des objets). Les mesures sémantiques proposées ont par ailleurs été couplées à des techniques d'apprentissage et du Web social pour la gestion des interactions dynamiques entre services et mashups dans la thèse de Hemza Labbaci (soutenue en 2022), en co-encadrement avec l'USTH d'Alger. L'approche consiste à modéliser par un graphe le réseau de services déployés et leurs interactions de type composition ou substitution et de prendre en compte les retours d'utilisation pour pondérer ces relations. Les mises-à-jour du graphe par l'intégration de ces retours, de nouveau déploiement et le calcul de similarité à base du LOD, plus des techniques d'apprentissage, permettent la sélection dynamique de nouvelles relations pertinentes dans le graphe.

## Axe « Exploration et analyse interactive de données »

Lors du contrat précédent, l'axe « Exploration et Analyse Interactive de Données » avait défini le projet suivant : développer des outils d'analyse centrés utilisateurs, donnant un statut privilégié à leurs attentes et préférences, proposer des méthodes d'exploration et d'analyse qui soient interactives, selon trois directions :

1. langages et algorithmes pour l'analyse de données,
2. faciliter l'analyse exploratoire de données,
3. évaluation de la qualité des analyses et équité.

Notons que ces directions restent des verrous scientifiques forts, comme cela est décrit dans un article récent de Communications of the ACM ("Automating data science" par De Bie et al., CACM 65(3), mars 2022), où les auteurs arguent que le volet exploration de la data science est celui qui requiert le plus d'attention. Cela est dû au fait que la composante humaine y est la plus forte et que l'effort d'automatisation doit porter sur l'assistance à l'humain, plutôt que viser une automatisation complète. Ceci est illustré notamment par les extraits suivants :

- Important parts of data science are already being automated, especially in the modeling stages, where techniques such as automated machine learning (AutoML) are gaining traction. Other aspects are more difficult to automate, not only because of technological challenges, but because open-ended and context-dependent tasks require human interaction.
- We see great potential for the assistance form of automation, through systems that complement human experts, tracking and analyzing workflows, spotting errors, detecting and exposing bias, and providing high-level advice.

Un projet inter-axe était également défini, à l'interface des axes « Exploration et analyse interactive de données » et « Intelligence des données et services ».

Nous reprenons les trois points ci-dessus pour présenter le bilan des dernières années. Nous citons entre parenthèses les publications les plus significatives associées aux contributions et le cas échéant les collaborations, thèses ou projets dans lesquels elles s'inscrivent.

### 1) Langages et algorithmes

Cette direction de recherche vise à traiter des verrous scientifiques et technologiques relevant de l'automatisation des tâches d'exploration et d'analyse de données, d'une part en abordant ces verrous de manière théorique, et d'autre part en produisant des briques algorithmiques génériques et efficaces. Certaines contributions ont donné lieu à des collaborations internationales, nationales, régionales, ainsi qu'avec les deux autres équipes du laboratoire LIFAT.

Sur le plan théorique, les travaux concernent l'étude du processus d'analyse exploratoire par la définition de cadres formels et de mesures, en insistant sur la prise en compte de l'humain au sein de ce processus.

L'axe exploration a notamment proposé un cadre visant à donner une nouvelle assise théorique et pratique à l'analyse interactive de données structurées (collaboration internationale avec les universités de Bologne et Ioannina, IS2019, EDBT2021), comprenant d'une part la définition et l'implémentation d'un langage de haut niveau pour l'exploration de données et d'autre part la définition formelle du processus d'exploration et son étude théorique (thèse d'Alexandre Chanson, en collaboration avec l'équipe ROOT du LIFAT, EDBT2022). L'axe a également proposé un cadre de fouille de données interactive, basé sur l'échantillonnage de motifs pour traiter des données complexes (thèse de Lamine Diop, soutenue en 2021, PAKDD2016, IDA2017, SDM2018, ICDM2018, PKDD2021). La thèse de Lamine Diop a obtenu le prix de la meilleure thèse de l'association EGC en 2021.

Conjointement, des réflexions ont été menées sur la manière de mesurer, du point de vue utilisateur, l'intérêt ou la proximité lors du processus d'exploration. Diverses mesures d'intérêt pour caractériser les découvertes dans les données ont été proposées (par exemple, ISF2022, EDBT2022), en s'appuyant notamment sur un tour d'horizon des mesures d'intérêt existantes (ADBIS2019). Des mesures de similarité entre trajectoires sémantiques, tenant compte du comportement humain, ont été définies dans la thèse de Clément Moreau (déjà citée dans l'axe précédent).

Les verrous technologiques abordés concernent essentiellement la prise en compte de données massives et en particulier le passage à l'échelle des approches d'extraction de données. L'axe a contribué à la mise au point d'algorithmes d'extraction de motifs sur Hadoop, (collaboration avec le LIFO, DEXA2017), de calcul de Skylines dans des bases de données (EDBT2020, DASFAA2020), et d'extraction de motifs skylines (collaboration avec le GREYC, LIRIS, LORIA, AIJ2017).

Ces contributions théoriques trouvent des applications en fouille de trajectoires sémantiques à partir de données de capteurs (thèse CIFRE de Frédéric Bisone), en mécanique des fluides en collaboration avec l'Université du Zhejiang en Chine (EGC2020, Measurement2021), dans le domaine de la vidéo (thèse de Luong Phat Nguyen soutenue en 2021, co-encadré avec l'équipe RFAT du LIFAT, VISAPP2021).

## 2) Facilitation de l'analyse exploratoire

Ces travaux concernent la prise en compte de l'interaction humaine lors de l'exploration de données, en s'appuyant majoritairement sur l'élaboration de recommandations. L'axe a su capitaliser les acquis du contrat précédent, en développant des techniques de recommandation originales, s'appuyant éventuellement sur du profilage utilisateur, et en diversifiant les cadres applicatifs de ses approches, le plus souvent en partenariat avec des industriels. En l'occurrence, il s'agit de recommander des requêtes décisionnelles (thèse CIFRE de Krista Drushku, IS2019A), du contenu éditorial (thèse CIFRE de Willème Verdeaux, IS2022), des plannings publicitaires (thèse CIFRE de Fodil Benali, ICTAI2021) et des parcours de vie (thèse CIFRE de Nicolas Ringuet, DAWAK2021). Le profilage utilisateur a été abordé par deux thèses autour de l'étude des habitudes d'analyse des utilisateurs : la thèse de Krista Drushku (cf. ci-dessus) sur l'apprentissage des intérêts utilisateurs et la thèse de Clément Moreau (IS2022, DOLAP2020) autour de l'analyse de logs de requêtes pour découvrir des patrons d'analyse.

En plus de ces approches, l'axe développe des techniques d'aide à la compréhension, en proposant des outils d'analyse et de visualisation permettant d'appréhender les déplacements et les activités d'un ensemble d'individus (thèse de Clément Moreau, FUZZ-IEEE2021, SAC 2021, SAC 2020), une méthode de clustering interactif (thèse CIFRE d'Adnan El Moussawi, IDA2017), et en démarrant des travaux sur l'explicabilité et l'interprétabilité. Ces derniers thèmes sont abordés dans deux thèses. L'une porte sur l'explication des recommandations (thèse de Willème Verdeaux, IS2022), l'autre sur l'explication et la confiance dans les modèles de prédiction (thèse de Nicolas Canovas, avec le BRGM).

## 3) Evaluation de la qualité

Ces travaux concernent l'étude de la qualité à la fois en amont et en aval du processus d'exploration de données. En amont, l'axe s'intéresse à l'étude de la qualité des données, avec un travail démarrant sur la prise en compte du contexte dans la gestion de la qualité de données (thèse de Flavia Serra) et un travail sur la qualification de données complexes, notamment géographiques (thèse de Raphaël Bres en collaboration avec le LASTIG de l'IGN, et projet PHC Ulysses avec le University College of Dublin en Irlande). En aval de l'exploration, le travail sur la qualité des explorations interactives de cubes de données s'est poursuivi lors de la thèse de Mahfoud Djedaini (IS2019B) et par la définition d'axiomes permettant d'évaluer la qualité et la complexité des processus d'extraction de motifs locaux (PKDD2018).

## Action inter-axe « Extraction de connaissances du Web des données »

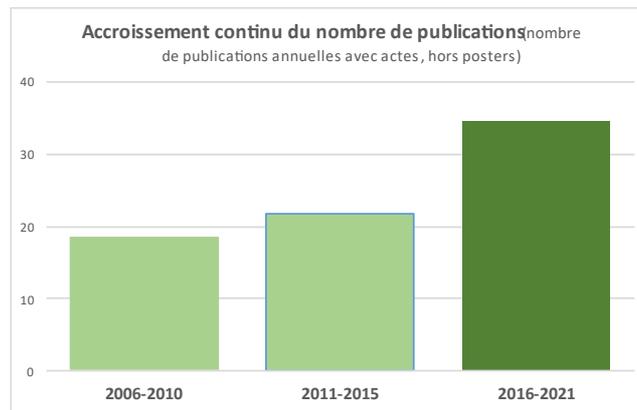
Des travaux transversaux ont débuté en 2016 entre l'axe « Exploration et analyse interactive de données » et l'axe « Intelligence des données et services » sur l'extraction de connaissances dans le Web sémantique. Deux sujets de thèses communs ont été conçus pour des doctorants en co-tutelle, qui ont démissionné au bout d'une année faute de financement suffisant. Les travaux ont continué en s'appuyant sur le projet ANR SESAMES et sur une collaboration avec Télécom Paris. Le contexte du Web sémantique ajoute deux défis scientifiques majeurs à la fouille de données : le biais des données et l'absence de l'hypothèse du monde clos. Sur le premier, notre première contribution a été de proposer une méthode d'estimation de la représentativité des connaissances en s'appuyant sur la loi de Benford (ISWC2018). Sur le deuxième, nous avons montré comment préciser les ontologies en découvrant automatiquement des cardinalités maximales qualifiant certaines propriétés (ISWC2019 cardinalities). L'originalité de notre approche est de s'appuyer sur le graphe de connaissances lui-même pour son enrichissement, sans source extérieure de données. Nous avons également réalisé plusieurs contributions pour bénéficier des graphes de connaissances à des fins d'analyse : requêtes analytiques (ISWC2019 anytime) et génération de tableaux comparatifs (ESWC2021). Dans cette direction, un recrutement est en cours sur un sujet de thèse intitulé « découverte de connaissances stratégiques interprétables dans le Web des données », « connaissances stratégiques » dans le sens de connaissances générées par un processus d'analyse de données pour éclairer des prises de décision, « interprétables » dans le sens où les ontologies du Web permettent de qualifier les données contenues dans les graphes de connaissances analysés (ainsi que

les résultats des analyses). Enfin, nous travaillons sur des moyens de faciliter l'exploration interactive de graphes de connaissances à des utilisateurs non informaticiens, notamment des archéologues. Pour cela et en lien avec l'outil d'interrogation visuelle<sup>5</sup> développé par l'entreprise Sparna implémenté dans OpenArchaeo, nous proposons une méthode de profilage qui expose dans le profil non seulement les classes et propriétés mais aussi les termes utilisés dans le graphe de connaissances (SWODCH2021).

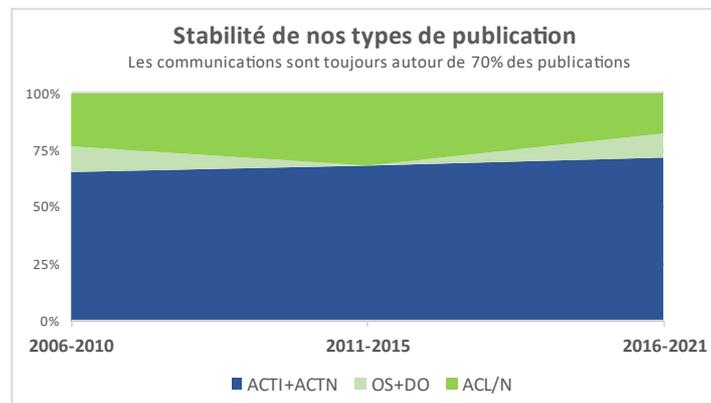
## Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité.

### Analyse de la qualité / quantité, co-productions inter-équipes de l'unité et avec acteurs externes

La production scientifique de l'équipe a continué à augmenter en quantité comme en qualité par rapport aux contrats précédents. Au total sur la période l'équipe a publié 42 articles en revues et 175 articles en conférences (127 en conférences internationales). La figure ci-dessous met en évidence l'accroissement quantitatif.

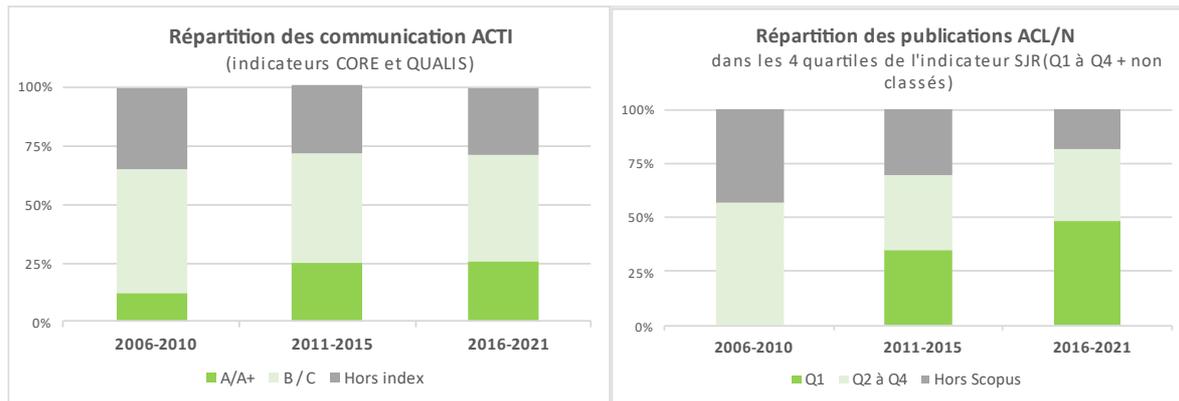


Les publications se répartissent toujours de la même manière entre conférences et revues, conformément aux pratiques dans nos communautés de recherche.



Pour autant, la qualité progresse, aussi bien pour les conférences que pour les revues. Pour les premières et plus particulièrement les conférences internationales (ci-dessous à gauche), la proportion de celles référencées dans CORE dans les catégories A+ et A reste stable, tandis que celle des conférences non référencées diminue. Pour les revues nationales et internationales (ci-dessous à droite), la tendance est encore meilleure avec une augmentation très nette de la proportion des revues classées dans le premier quartile de leur domaine selon l'indicateur SJR.

<sup>5</sup> <http://sparnatural.eu/>



De plus, l'équipe a été nommée pour le meilleur article aux conférences ISWC 2018 et ESWC 2021 et a obtenu le prix du meilleur article à EGC 2021 ainsi que le prix de la meilleure thèse de l'association EGC en 2021.

Les incitations du LIFAT aux collaborations entre ses équipes ont permis à 3 thèses sur financement passant par l'école doctorale d'être co-encadrées avec des membres d'autres équipes du laboratoire, deux n'ont démarré que récemment, en 2020 et 2021, il y a donc encore peu de publications avec un co-auteur d'une autre équipe (ROADEF2021, DAWAK2021, EDBT2022, ROADEF 2022). Deux autres ont démarré très récemment, cette fois avec financements sur projet et CIFRE.

Il est important de noter que les nombreux projets et collaborations amènent les publications de l'équipe BdTIn à être très souvent réalisées avec des partenaires hors LIFAT. Par exemple, 10 thèses sont en co-tutelle ou co-direction avec universités étrangères, 4 en co-direction avec un.e membre du LIFO d'Orléans. Par ailleurs l'équipe a la capacité de collaborer régulièrement avec des chercheurs d'autres disciplines, majoritairement de sciences humaines et sociales mais également de sciences naturelles (voir précisions plus loin, dans l'analyse des projets). Pour toutes ces raisons, au niveau des revues (ACL) 50% des publications comprennent un co-auteur d'une université étrangère et 47,6% un co-auteur d'une autre université française. De même, au niveau des conférences internationales (ACTI) 34,7% des publications comprennent un co-auteur d'une université étrangère et 36,5% un co-auteur d'une université française. Ces pourcentages sont stables par rapport aux précédentes périodes d'évaluation HCERES.

### International : originalités, projets et relations académiques pérennes

L'équipe BdTIn publie très majoritairement dans des revues et des conférences internationales : 80% des articles en revue sont dans des journaux internationaux et, sur l'ensemble des communications avec sélection par comité de lecture, 75% sont en conférences et workshops internationaux. Comme le montrent les figures précédentes, ses publications internationales sont en grande proportion dans des journaux et des conférences internationales parmi les plus réputés de ses domaines d'intérêt, ce qui démontre *une bonne reconnaissance, de la part de ces communautés internationales, de la qualité des travaux menés dans l'équipe.*

Une originalité partagée dans l'équipe et pas seulement au niveau international, est que, si nous avons bien remarqué la domination récente de méthodes de deep learning (DL) à travers les quantités d'articles qui leur sont consacrés, l'attention portée à nos objets d'étude (les langues, les connaissances, les processus de découverte interactive, etc.) prime sur l'utilisation systématique de ces méthodes.

Comme *originalités à l'international*, tout d'abord la période 2016-2021 a vu l'essor du *réseau COST PARSEME* dédié à la prise en compte des expressions poly-lexicales dans le TAL, piloté par une membre de BdTIn (Agata Savary). Ce vaste réseau de collaborateurs a produit des résultats importants comme un corpus énorme, multilingue (16 langages), annoté manuellement, publié en mode ouvert et hébergé dans la grande infrastructure de recherche européenne CLARIN, avec des outils d'identification automatique des expressions poly-lexicales dans les textes. Il a mis en place des actions annuelles d'envergure qui perdurent alors même que son financement finissait en 2017, comme un workshop, Multi Word Expression (MWE), et une campagne d'évaluation dans laquelle chaque année des équipes comparent les résultats de plus de leurs systèmes sur une vingtaine de langues. Cette continuité est assurée par la section MWE du SIGLEX de l'ACL. Ensuite, les collaborations menées par une membre de BdTIn (Béatrice Markhoff) avec nos collègues en SHS autour du patrimoine culturel ont amené l'équipe à participer à *deux projets H2020*, de pluri (regards complémentaires sur un objet commun) à trans (réalisations qui traversent les différents points de vue) disciplinaires. En effet dans ces projets sont menées, par des collègues de différentes disciplines, des réflexions et des réalisations dans le champ des apports de la numérisation du patrimoine culturel, vers des espaces virtuels de recherche et des réseaux de collaboration autour de ces connaissances. Ces projets impliquent eux aussi de nombreux partenaires académiques et du monde non académique (41 dans l'un, 19 dans l'autre, une dizaine en commun). Enfin, les *collaborations menées de longue date dans l'équipe avec l'université Gaston Berger du Sénégal* ont donné lieu à l'élaboration d'un cadre original de fouille de données interactive basé sur l'échantillonnage de motifs pour traiter des données complexes (thèse de Lamine Diop, prix EGC 2021, encadré côté BdTIn par Arnaud Giacometti et Arnaud Soulet).

Tout cela se retrouve dans le tableau ci-dessous, extrait des principaux partenaires internationaux des membres de BdTIn, avec certaines des réalisations communes. Parmi la trentaine de collaborations recensées, ce tableau se limite aux 12 partenaires avec lesquels elles sont les plus actives et devraient le rester, ce pour quoi les collaborations portées par Agata Savary n'y apparaissent

pas. L'équipe collabore sur des directions de thèse avec des universités africaines (Sénégal, Algérie, Tunisie, ainsi que Gabon), une université libanaise et l'Université de la République d'Uruguay.

Partenaires internationaux	Pays	Thèses		Mobilités		Co-publications		Projets et workshops
		co-tutelles	co-directions	entrantes	sortantes	ACL	ACTI	
Université Gaston Berger Saint Louis	Sénégal		1	2	1	2	5	
USTHB Alger	Algérie		2	3			5	
University of Ioannina	Grèce				2	1	7	
University of Duesseldorf	Allemagne				2		6	
University of Bologna	Italie				1	3	3	1
Université de Tunis	Tunisie	1	1				6	
Université de Sfax	Tunisie	2		1		2	1	
University College Dublin	Irlande					2		3
Free University of Bozen-Bolzano	Italie			1		1	2	1
CNR ISTC (Trento)	Italie						3	1
Universit de Tiaret	Algérie					1	2	
PIN Prato, Université de Florence	Italie					1		2

Les projets internationaux financés, listés avec les projets nationaux et régionaux ci-après, concernent le réseau COST PARSEME déjà présenté, les collaborations avec l'University College Dublin montrées dans le tableau des partenaires, ainsi que les deux projets H2020. De plus, de nombreuses collaborations existent hors projets financés et thèses, liées à d'autres responsabilités comme par exemple dans des conférences ou workshops, et il y a très régulièrement des mobilités entrantes et sortantes. C'est le cas notamment avec des universités européennes (Italie, Grèce, Allemagne, Irlande, Pologne) ou nord-américaines (Université du Michigan Dearborn, Université de Québec). Beaucoup de ces collaborations débouchent sur des publications communes (rappel : 50% des articles en revues et 34,7% des publications en conférences comprennent un co-auteur d'une université étrangère).

### National et régional : projets et relations académiques pérennes

L'équipe BdTIn est investie dans de nombreux projets de recherche avec des partenaires académiques, en tant que participante (P) ou coordinatrice (C), que ce soit au niveau international, national et régional, comme le montrent les tableaux récapitulatifs suivants.

Projets nationaux : 771 k€ (656 k€ obtenus sur 2016-2021)					Projets régionaux : 429 k€ (239 k€ obtenus sur 2016-2021)				
PIA(ISTEX)	P	2014	2017	111 000 €	Région CVL (ARVIVA)	P	2014	2016	30 000 €
ANR(DEMOCRAT)	P	2015	2018	4 000 €	Région CVL (DOPAn)	C	2014	2017	160 000 €
ANR(PARSEME - FR)	P	2016	2020	169 676 €	Région CVL (ODIL)	P	2016	2017	22 000 €
ANR(MOBI'KIDS)	P	2016	2022	65 814 €	Région CVL (GIRAFON)	P	2016	2017	35 000 €
ANR(TALAD)	P	2017	2022	87 307 €	Région CVL (RAVIOLI)	P	2017	2020	23 000 €
Fondation(PREDICT4ALL)	P	2017	2020	39 468 €	Région CVL (SmartLoire)	C	2017	2021	114 000 €
PIA Istex (Biosystème)	P	2017	2018	40 000 €	Région CVL (ArchSW)	C	2018	2019	45 000 €
ANR(Abliss)	P	2018	2022	130 604 €					
ANR(SESAMES)	P	2018	2023	106 380 €	Projets internationaux : 115 k€ (100 k€ obtenus sur 2016-2021)				
GDR MAGIS (VGI/Crowdsourcing)	C	2019	2025	3 000 €	COST (PARSEME)	C	2013	2017	11 000 €
GDR MADICS (MADONA)	C	2019	2020	6 000 €	COST (KNOWeSCAPE)	P	2013	2017	2 000 €
Consortium CORLI(FAIREANCORPLUS)	C	2020	2020	4 000 €	Banque Mondiale (CEA-MITIC)	P	2015	2018	2 000 €
GDR MAGIS (mobilité individuelle)	C	2013	2025	3 000 €	H2020 (ARIADNEplus)	P	2018	2022	2 000 €
GDR MADICS (Fender)	C	2021	2022	1 500 €	PHC Utique (TransLvoGelInfo)	C	2020	2021	2 500 €
					H2020 (4CH)	P	2021	2024	96 187 €

Au **niveau national**, l'équipe a des financements dus à son pilotage d'actions des GDR CNRS MADICS (sciences des données, actions Fender et MADONNA) et MAGIS (géomatique, actions mobilité individuelle et VGI/Crowdsourcing). On retrouve aussi dans le tableau des projets nationaux financés les deux projets qui ont contribué à la plateforme ISTE (socle de la bibliothèque scientifique numérique nationale), celui financé par la Fondation Paul Bennebot et les projets ANR DEMOCRAT, PARSEME-FR, TALAD et Abliss, tous décrits dans le bilan de l'axe « Traitement des langues et interactions ». De même, les projets ANR MOBI'KIDS et SESAMES ont été décrits dans les bilans respectifs des deux autres axes. A noter que deux autres projets ANR ont été acceptés pour financement qui démarre en 2022 dans l'axe « Traitement des langues et interactions » : SELEXINI en coopération avec l'axe « Exploration et analyse interactive » et AAC4ALL sur l'aide aux personnes en situation de handicap.

Tous ces projets sont le résultat et le support de relations académiques pérennes, sachant que bien d'autres existent en dehors de financement particulier, par exemple avec Télécom Paris, les universités de Caen, Lyon, Nancy (AIJ) et l'INRAe Clermont-Ferrand (IS2022) ou avec des actions financées par nos partenaires comme par exemple le CNRS MASTODONS QCM\_biochem ou l'IGN (co-encadrement des thèses de Stefan Ivanovic et Raphaël Bres).

Au **niveau régional**, bien que cela n'apparaisse pas dans le tableau des projets régionaux financés, l'équipe participe activement au Réseau Thématique Régionale (RTR) DIAMS, dont elle anime 3 groupes sur les 12<sup>6</sup> : RITUEL (Recherche d'Information, Traitements

<sup>6</sup> <https://www.univ-orleans.fr/lifo/evenements/RTR-DIAMS/groupes/>

Utiles et Eclairés des Langues) avec le LIFO, le LLL et le BRGM, HumanInTheLoop (Human In the Loop for Data Mining and Machine Learning) avec le LIFO et l'INSERM et SeWin (Semantic Web Interoperability) avec la MSH Val de Loire (VdL). De même elle participe au HumaNum Lab sur le Web sémantique porté par la MSH VdL. De plus, elle co-encadre une thèse (de Nicolas Canovas) avec le BRGM. Les projets financés par la Région CVL entre 2016 et 2021 apparaissent dans le tableau précédent, sachant qu'un nouveau projet Région CVL démarre en 2022, SQVALD, budget de 140 000€, sur l'accompagnement des patients hors hôpital dans l'axe « Intelligence des données et des services ». Tous font l'objet de collaborations fortes avec des acteurs régionaux, que ce soient le LIFO, des laboratoires en SHS (historiens du CESR, archéologues de CITERES-LAT, linguistes du LLL) ou en médecine, ou bien des structures comme l'observatoire de l'économie et des territoires du Loir-et-Cher, le Comité Régional du Tourisme ou la direction des sports de la ville de Tours.

### Rayonnement et attractivité

Les bilans par axe, l'analyse des publications, des partenariats et des projets internationaux et nationaux montrent déjà le **rayonnement de l'équipe**, tout comme les thèses en co-tutelles ou co-directions internationales et nationales (commentées par la suite).

Les membres de l'équipe sont également sollicités pour participer à des *comités de pilotage de workshops* (DOLAP depuis 2017 classée B selon CORE, MWE depuis 2014, SW4CH puis SWODCH depuis 2015), des *comités éditoriaux* (revue TAL depuis 2001, Journal of Language Modelling depuis 2012, Revue internationale de géomatique depuis 2016, Data and Knowledge Engineering -Q2- depuis 2018, Infotheca depuis 2017) ou être « senior PC » de conférences (EGC depuis 2016, IJCAI depuis 2020), mener des *expertises scientifiques* (pour le Dutch Research Council en 2021, pour le Linz Institute of Technology en 2018, pour le programme MITACS Accélération au Canada en 2020, régulièrement pour l'ANRT), collaborer à des *ouvrages scientifiques* (Encyclopaedia of Database Systems en 2018, ), et *rapporter sur des thèses de doctorat d'universités étrangères* (Aymn Al-Serafi Université Polytechnique de Catalunya et Université Libre De Bruxelles en 2021, Matteo Francia Université de Bologne en 2021, Daniele Metilli Université de Pise 2021).

De plus l'équipe s'investit dans l'organisation de workshops ou tracks réguliers (DOLAP en 2017 et 2021, SW4CH en 2017, 2018 puis SWODCH en 2021 et 2022, GIA@SAC en 2020 et 2021), de la conférence nationale TALN 2017<sup>7</sup> et a accueilli avec succès et malgré la pandémie la conférence nationale EGC en janvier 2022, ce qui a impliqué un travail continu depuis fin 2020. EGC2022<sup>8</sup> s'est déroulée sur une semaine à la fois en présentiel et en distanciel, avec 215 participants dont les trois quart en présentiel, avec son école en ouverture, 5 ateliers, sa journée Grand Public et ses 5 conférencières invitées de renom. Les conférences plénières ont été retransmises en direct sur Youtube et la plupart des présentations peuvent être réécoutées.

Concernant l'**attractivité** de l'équipe, elle est là encore portée par le dynamisme de ses partenariats sur les projets et les thèses, déjà abordé par ailleurs. Côté recrutements de doctorants les efforts décrits dans le paragraphe « formation par la recherche » et ceux menés avec les entreprises (commentés dans le Domaine 4) garantissent une certaine attractivité. Toutefois sur les financements académiques, aux rémunérations plus réduites et aux sujets plus formels, nous constatons encore des difficultés à attirer de nombreuses bonnes candidatures de l'extérieur.

Nos partenariats et notre formation de docteur.e.s soutiennent régulièrement nos recrutements d'ATER et de postdoc. Lors de recrutements de permanents notre rayonnement amène de nombreuses demandes de candidat.e.s pour réaliser des séminaires, lesquels sont très suivis par les membres de l'équipe étant donnée l'importance que revêt pour nous le recrutement d'un.e collègue. L'attractivité de l'équipe est illustrée par le fait d'attirer sur des postes de MC des chercheurs reconnus avant d'arriver : si l'équipe n'a pas pu remporter la décision pour recruter Karen Fort (poste ouvert à plusieurs équipes du LIFAT), cela a été possible pour Nicolas Labroche et Cyril de Runz, tous deux déjà HDR, qui se sont révélés très pertinents par leur intégration rapide, attestée par leur position centrale dans le graphe de co-publiants.

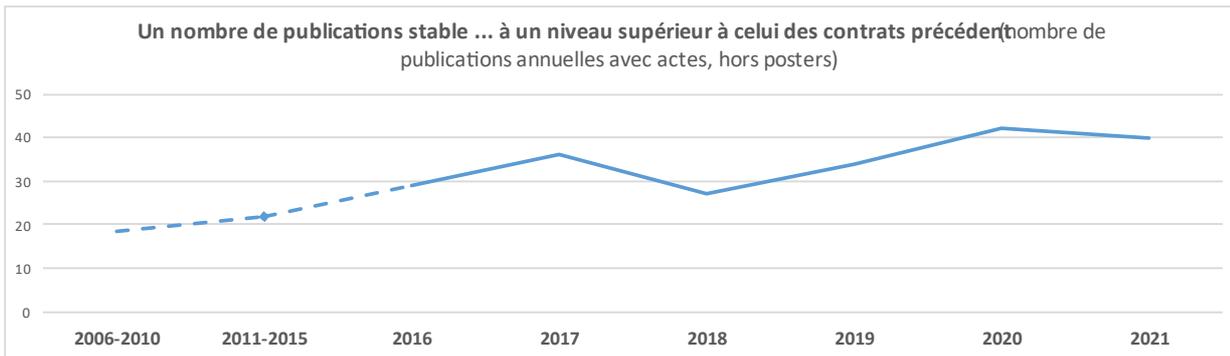
## Référence 2. La production scientifique est proportionnée au potentiel de recherche de l'équipe et répartie entre ses personnels.

### Répartition des publications entre permanents

La figure ci-dessous montre le *nombre moyen de publications par an de l'équipe* sur les 3 derniers contrats, le nombre de permanents étant resté constant entre les 2 derniers contrats (rappelons que la qualité des publications a quant à elle été analysée au début de la Référence 1).

<sup>7</sup> <https://taln2017.cnrs.fr/>

<sup>8</sup> <https://egc2022.univ-tours.fr/>



Dans l'équipe BdTIn, le *nombre moyen de publications par an et par permanent* (dans les catégories OS/DO, ACL/N, ACTI et ACTN) est passé de **1,5** sur 2011-2015 à **2,3** sur 2016-2021. La répartition n'est évidemment pas uniforme mais ce qu'il faut surtout noter est que **tous** les membres de BdTIn ont publié au cours de la période 2016-2021. Les collègues contraints de s'éloigner de la recherche pour un temps sont encouragés lorsqu'ils peuvent y revenir, en les associant aux projets et à l'encadrement des stages de Master et des thèses. Par exemple, l'équipe a dernièrement classé en premier dans la liste de ses prochains sujets à financer sur bourse ministère ou région un sujet de thèse inter-équipe incluant dans ses co-encadrant.e.s une personne dans cette situation. Plus généralement, la répartition des ressources (bourses, budget) se fait après discussion en réunion d'équipe, par décisions prises en commun sur la base de données partagées, notamment d'historiques des attributions. Une dimension facilitatrice est également les présentations régulières en réunion d'équipe ou séminaires de nos réalisations et de nos sujets de futures thèses. De ce fait, un.e collègue qui n'aurait bénéficié d'aucune ressource pendant un temps du fait d'un éloignement de la recherche sera accompagné.e dès qu'il ou elle aura à nouveau la possibilité de s'investir en recherche et ses demandes de ressources seront satisfaites autant que possible.

Par ailleurs l'équipe soutient les demandes de CRCT, certains ayant pu servir à se consacrer à nouveau à une activité de recherche. Ce fut le cas pour Arnaud Giacometti en 2017 à la suite de ses 7 années de Vice-Président aux relations internationales et pour Patrick Marcel en 2019 après son grand engagement dans les 2 Master européens (Erasmus IT4BI et Erasmus+ BDMA) portés par l'équipe au contrat précédent puis de 2016 à 2018.

### Politique d'accueil de nouveaux collègues

L'équipe n'a pas accueilli de jeune chercheur ou chercheuse permanent.e au cours du contrat (pour lequel.e.s l'université applique une politique d'accueil avec réduction de la charge d'enseignement). Pour autant, la politique de répartition des ressources décrite précédemment accompagne aussi les nouveaux arrivants. Par exemple Cyril de Runz a bénéficié un an après son arrivée d'un financement de thèse, sa demande ayant été soutenue par l'équipe au niveau du laboratoire.

### Formation par la recherche et productions des doctorants et post-doctorants

Au début du contrat l'équipe a continué son investissement dans un master Erasmus Mundus (IT4BI jusqu'en 2017-2018 puis BDMA en 2018-2019, qu'elle a contribué à concevoir), comportant un projet dans lequel les étudiants étaient initiés à la recherche. Le master BDMA a perduré après l'arrêt de la participation au consortium international : il comprend encore ce projet d'initiation à la recherche. L'équipe a par ailleurs réalisé un gros investissement dans la direction de thèses au cours de ce contrat, au nombre de 41 au total. Ces thèses sont évoquées dans les travaux scientifiques par axe avant la Référence 1, celles en CIFRE sont également reprises plus loin dans le Domaine 4. Le tableau suivant récapitule les types de financement et de direction (co-tutelle, co-direction avec université étrangère, co-direction avec partenaire français, co-direction avec autre équipe LIFAT) des 17 thèses soutenues, d'une *durée moyenne de 43 mois* et des 18 thèses en cours. Les noms en gras sont ceux des diplômés du master local (10 au total se sont engagés en thèse dans l'équipe à l'issue de leur master). Il y a eu 6 thèses arrêtées sur la période (dont une par un diplômé du master local), en moyenne au bout d'un an, les personnes ayant rejoint des entreprises.

17 thèses soutenues					18 thèses en cours			
Docteur.e	année	financement	type	devenir	Doctorant.e	année	financement	type
	soutenance		direction			début		direction
Ahmed ABID	2017	étranger	co-tut	entreprise France	Willem VERDEAUX	2019	CIFRE	
Mouna ELASHTER	2017	étranger			Raymond ONDZIGUE MBENGA	2019	étranger	co-tut
Nasredine CHENIKI	2017	étranger	co-dirEtr	postdoc Luxembourg	Faten EL OUTA	2019	ED MIPTIS	
Jakub WASZCZUK	2017	ED MIPTIS	co-dirFr	EC Allemagne	Amina BRAHEM	2019	projet	co-tut
Mahfoud DJEDAINI	2017	projet		postdoc France	Fouad KOMEIHA	2019	projet	co-tut
Adnan EL MOUSSAWI	2018	CIFRE		data scientist univ. Tours	Fodil BENALI	2020	CIFRE	
Stefan IVANOVIC	2018	IGN	co-dirFr	postdoc Suisse	Nicolas RINGUET	2020	CIFRE	
Thanh Binh NGUYEN	2018	ED MIPTIS	co-dirFr	EC Vietnam	Flavia SERRA	2020	étranger	co-tut
Krista DRUSHKU	2019	CIFRE		entreprise France	Adam LION-BOUTON	2020	ED MIPTIS	co-dirFrEtr
Clay PALMEIRA da SILVA	2019	étranger		postdoc Belgique	Alexandre CHANSON	2020	ED MIPTIS	co-dirLIFAT
Fatma BEN MESMIA	2019	étranger	co-dirEtr	EC Tunisie	Gregor JOUET	2020	U. LdV	
Caroline PASQUER	2019	projet	co-dirFr	entreprise France	Nicolas CANOVA	2021	BRGM	
Lamine DIOP	2020	étranger	co-dirEtr	ATER univ. Tours	Maëlle BRASSIER	2021	CIFRE	
Marwa BOULAKBECH	2020	ED MIPTIS		EC en CDD univ. Tours	Najet Dhouha HADJ MOHAMED	2021	étranger	co-tut
Frédéric BISONNE	2021	CIFRE		entreprise France	Raphaël BRES	2021	IGN	co-dirFr
Long Phat NGUYEN	2021	ED MIPTIS	co-dirLIFAT		Ben CRULIS	2021	ED MIPTIS	co-dirLIFAT
Clément MOREAU	2021	projet		entreprise France	Praveen SONI	2021	projet	co-dirLIFAT
					Thomas KASTNER	2022	CIFRE	co-dirLIFAT

Les doctorant.e.s et postdoctorant.e.s sont largement associés à la rédaction des articles, principalement en conférences où les doctorants co-signent 54% de la production (il leur faut au minimum trois articles dont un dans une bonne conférence pour envisager une soutenance), cela monte à 62,5% avec les co-auteurs postdoctorants. Pour les articles en revue 27,5% sont co-signés par des doctorants et 40% quand on y ajoute les articles co-signés par des postdoctorants. Pour les autres types de publications (ouvrage, chapitre d'ouvrage, direction d'ouvrage), 19% sont avec des doctorants et 32% en comptant aussi des postdoctorants.

## Synthèse de l'autoévaluation

Comme le montre la matrice SWOT ci-après, les forces de l'équipe résident globalement dans l'investissement important, l'ouverture et le dynamisme de ses membres, capables de mener leur recherche, diriger, coordonner et collaborer dans des projets de toutes envergures. Ce qu'il faut renforcer a trait à une forme de modération pour favoriser la résilience, face aux réalités de l'environnement et, en interne, face aux évolutions des collaborations. L'environnement dans lequel nous évoluons est en effet plein d'injonctions contradictoires, avec des demandes sociétales fortes sur nos thématiques, tant en recherche qu'en enseignement, des demandes des autres disciplines de plus en plus fortes sur nos thématiques également, tout cela avec des moyens supplémentaires non pérennes apportés ponctuellement par des projets à l'administration chronophage, qui se multiplient. Cela ne favorise pas les synergies et le travail commun sur le terme moyen et long. La très forte demande du privé sur nos thématiques de recherche entraîne également des réorientations plus faciles pour celles et ceux de nos doctorant.e.s qui rencontrent des difficultés en début de thèse, personnelles ou financières. L'équipe a ainsi perdu 6 doctorant.e.s, chaque fois dans leur première année de thèse (la COVID 19 ayant aussi joué). Nous en avons développé une vigilance particulière pour évaluer la motivation et l'idée que se font les candidat.e.s d'un travail de thèse lors du recrutement.

Enfin, l'équipe court le risque d'une démobilitation de ses HDR, présents et à venir, face au manque de perspectives d'évolution de carrière à l'université de Tours.

### Production scientifique de l'équipe BdTIn

Origine interne	
Forces	Faiblesses
Capacité à maintenir ou améliorer la qualité et la quantité de notre production scientifique	Difficultés à lâcher des responsabilités (parce que ces responsabilités garantissent plus d'aisance dans l'environnement d'exercice de nos fonctions)
Capacité à encadrer de nombreuses thèses, qualité d'encadrement attestée par les publications des doctorant.e.s et l'avenir des docteur.e.s	Plusieurs démarrages de thèses suivis quelques mois plus tard de prise de poste dans le privé
Richesse et quantité des projets financés, bon ancrage régional, nombreuses collaborations nationales, capacité à développer des recherches avec d'autres disciplines	Structuration par axe prévue en 2016 remise en cause par les dynamiques d'interaction réelles au cours du contrat (sauf axe TAL)

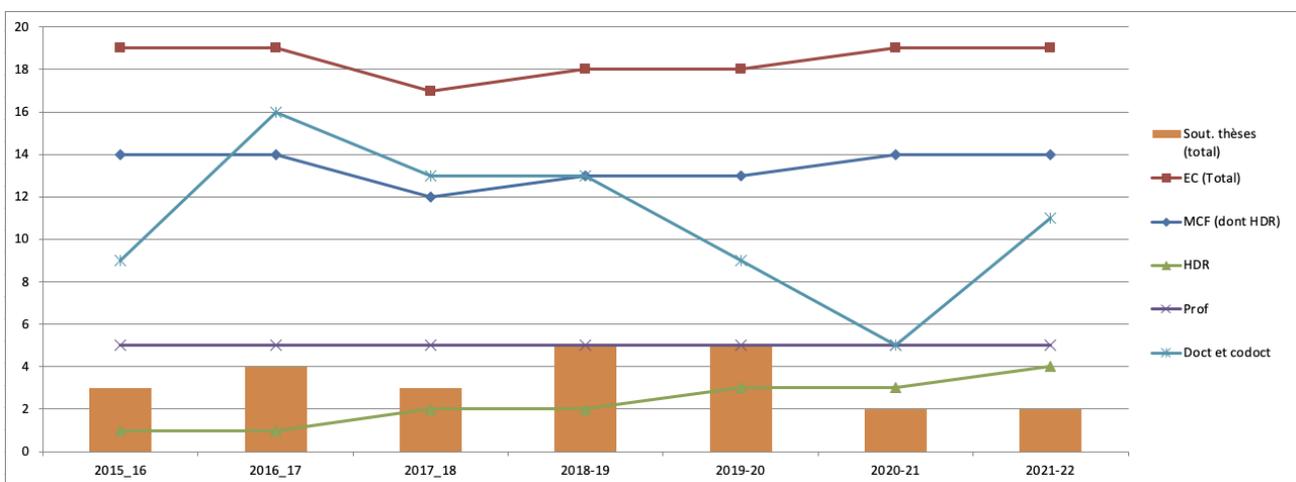
Dynamisme, évolution des interactions intra et extra axes thématiques de l'équipe, augmentation des interactions avec autres équipes du LIFAT	Difficulté à attirer de bonnes candidatures de l'extérieur sur les thèses à financements académiques
Diversification de nos sources de financements, en particulier avec des entreprises via des financements de thèses CIFRE	Manque de supports de PU par rapport au nombre de HDR, aggravé par le changement de fonction de Arnaud Giacometti (président actuel de l'université) et la perte du support de PU occupé par Denis Maurel après son départ à la retraite
Capacité à développer et maintenir des réseaux internationaux	Risque de démobilité des MC HDR par manque de perspective
Capacité à motiver des étudiants de notre Master pour la recherche	
<b>Origine externe</b>	
<b>Opportunités</b>	<b>Menaces</b>
Demande importante des autorités européennes dans le domaine de la science ouverte et des principes FAIR	Marché de l'emploi très concurrentiel : difficulté à recruter des doctorants, postdocs et ingénieurs, des doctorants qui s'engagent puis décident de prendre un poste dans le privé
Nombreuses sollicitations des autres disciplines académiques pour collaborer sur leurs traitements de leurs données	Concurrence entre disciplines : il est nécessaire de prendre des responsabilités à tous les niveaux pour ne pas être "oubliés" dans la répartition des ressources (postes)
Synergies potentielles entre les thématiques de l'équipe	Perte de deux chercheur.e.s clés : Arnaud Giacometti (devenu président d'université) pour l'axe Exploration et Agata Savary (devenue PU à Saclay) pour l'axe TAL, sans perspective de remplacement dans le cas du premier

### 3.3. Equipe RFAI

#### Présentation de l'équipe

L'équipe RFAI (Reconnaissance de Formes et Analyse d'Images) est composée de 19 permanents en poste : 5 PR (dont 1 en poste à Lille sur les mêmes thématiques qu'un PR local), 14 MCF dont 4 HDR (les membres en disponibilité et ne contribuant pas à l'équipe ne sont pas comptabilisés ici). Les effectifs sont donc stables mais on notera un léger *turn-over* puisque 2 membres sont partis (retraite et départ en entreprise), "compensés" en milieu et fin de contrat par 2 recrutements. Sur le contrat on compte désormais 3 HDR en plus (4 au total dont 1 à l'INSA), ce qui permet une meilleure répartition des directions de thèses. En ce qui concerne la répartition géographique, les permanents sont répartis sur 4 sites et composantes (Polytech Tours, UFR Sciences, IUT de Tours, INSA CVL). 1 PR est également en poste à Lille. Par rapport au précédent contrat, le site de l'INSA accueille désormais 4 permanents qui travaillent sur des thématiques proches. 4 permanents restent donc un peu plus isolés géographiquement (2 à l'IUT de Tours, UFR Sciences, Lille). Ce potentiel de 19 permanents (EC soit 9,5 chercheurs eq. temps plein) est à nuancer. En effet, beaucoup de permanents sont engagés dans des charges administratives lourdes (cf. Tableau des charges en annexe B) nécessaires d'une part au bon fonctionnement des composantes/unités et d'autre part à la préservation des intérêts et à la juste considération de la place de l'informatique dans les composantes et établissements. Cet investissement et cette représentation, démultipliés par le nombre de composantes desquelles nous sommes issus, s'est fortement accru au cours du contrat pour diverses raisons (inflation "naturelle" afférentes à ces charges, place croissante de l'informatique dans les politiques des composantes/établissements, situation sanitaire, ...). Ainsi le potentiel recherche réel de l'équipe sur le contrat pourrait être estimé au mieux à 6,5 eq. temps plein (ce calcul, tout à fait questionnable, représente néanmoins un élément d'information important).

Pour ce qui est des non-permanents, on compte une moyenne de 10,8 doctorants par an mais avec des variations importantes selon les années (dépendance vis à vis des projets). Sur la durée du contrat, il y a eu 22 soutenances de thèses (3,7 / an en moyenne) dont 13 de l'Université de Tours ou de l'INSA (2,2 par an), les autres étant des co-encadrements avec d'autres établissements. L'équipe est complétée par 2,2 post doctorants et/ou ATER par an en moyenne. Il s'agit d'une faible proportion (et en baisse) liée d'une part à la nature des contrats et d'autre part à la difficulté à recruter des ATERs ayant à la fois un profil recherche et enseignement compatible avec les besoins (sur les 2 postes ATER auxquels l'équipe peut prétendre, aucun n'est dans les thématiques de l'équipe depuis 2019-2020). Également, l'équipe ne compte toujours pas d'ingénieurs permettant de faciliter la valorisation des développements (le CETU ILIAD3, lorsqu'il est associé aux projets de recherche, permet de compenser en partie ce manque). Ces évolutions au cours de la période sont résumées sur la figure ci-dessous :



La politique d'animation de l'équipe, prend en compte plusieurs spécificités de l'équipe. Outre la dispersion géographique, il faut noter que l'INSA en tant que tutelle possède sa propre politique scientifique et cherche naturellement à accroître sa visibilité au travers de son personnel. Également, pour rappel, l'équipe a subi dans le contrat précédent une restructuration forte en intégrant les membres des petites équipes du LIFAT (ex. FOVEA et HANT). Tout cela, associé au manque de disponibilités lié aux charges administratives, rend le maintien de la cohérence thématique et de la cohésion de l'équipe particulièrement délicat. Dans ce contexte, il a fallu trouver un équilibre dans l'animation pour maintenir à la fois la cohésion et la cohérence sans pour autant rajouter de lourdeurs ou d'obligations au-delà des recommandations essentielles : éviter la dispersion thématique ; partage des informations sur le montage de projets pour associer les membres compétents qui le peuvent et le souhaitent ; objectif d'une publication dans une revue pour chaque thèse (ou dans l'année qui suit). Ainsi, sur les aspects administratifs et fonctionnels, la fréquence des réunions a été réduite avec la mise en place d'une communication asynchrone et le partage de documents en complément. La diffusion d'informations a également été réduite en limitant les redondances et en les ciblant davantage (système d'abonnement à des newsletters). Les décisions importantes restent en revanche prises collégialement et en réunion. En particulier, les propositions de thèses sur bourses académiques sont présentées en équipe de façon à permettre une adaptation du périmètre du sujet pour permettre des co-encadrements. Le choix des sujets retenus s'effectue également après une discussion et un vote, de façon à prendre en compte différents critères (nouveaux collègues, sujets innovants, sous-encadrement, thématique...). Également, l'investissement d'une partie de l'équipe dans des projets conséquents sont toujours discutés. Enfin, nous avons maintenu la fréquence des séminaires internes scientifiques (cf. Liste en annexe B ou sur le [site de l'équipe](#)) reposant sur : les visites de collègues d'autres établissements,

l'initiation ou l'avancement de travaux personnels, les répétitions de soutenances (thèses, HDR), la préparation de conférences pour les doctorants...

### Objectifs scientifiques

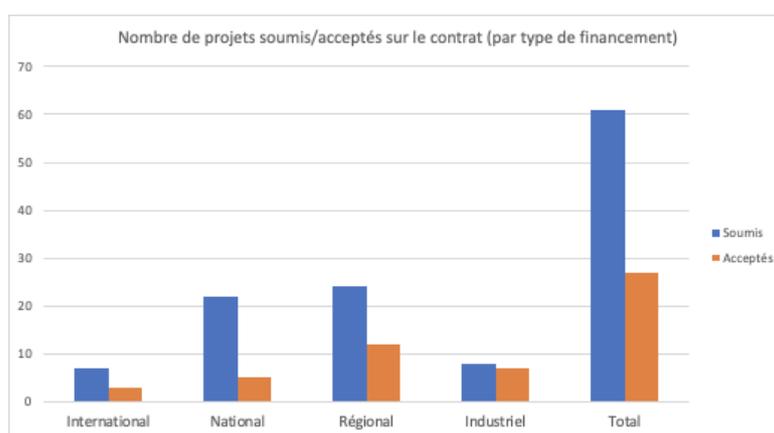
Les activités de l'équipe s'articulent historiquement autour de trois grands thèmes : les outils et méthodes de reconnaissance de formes, l'analyse d'images et de vidéos, la visualisation et l'interaction. Une attention particulière a toujours été portée à la collaboration de ces méthodes au sein des chaînes de traitement dans lesquelles l'utilisateur est susceptible d'intervenir, tant pour l'apprentissage et le paramétrage que pour l'usage final. Sur ce contrat, l'équipe a poursuivi l'intégration des outils et méthodes de *deep learning* dans ses travaux, soit comme outil, soit en tant qu'objet de recherche, sur lequel nous essayons de conserver un regard critique. Les travaux de l'équipe se structurent de la façon suivante :

- 3 axes scientifiques principaux interconnectés :
  - o Les **modèles et algorithmes d'analyse d'images et de vidéos** (environ 30% de l'activité) : il s'agit surtout de travaux sur des opérateurs (par apprentissage profond ou non) et de chaînes d'analyse (reconnaissance de formes, analyse de vidéos) traitant de problématiques spécifiques. Une partie de ces travaux ont été publiés dans revues de qualité (cf. Fichier de production revues [53, 63, 72, 92]). Si une très grande partie de ces travaux sont à présent abordés au travers du deep learning (thèses de T. Birgui Sekou, Q.C. Le, ...), il subsiste des travaux assez singuliers et originaux autour des traitements bas niveaux et des opérateurs temps réels. Ceux-ci ont également permis de mettre en place des relations durables avec l'université de Hong Duc au Vietnam, et de mettre en œuvre une architecture matérielle pour la collecte de flux TV servant à la création de la plus grande base de données portant sur la problématique de détection de segments vidéo (83k vidéos, 10 600 heures). Cette base STVD et le travail de recherche associé feront l'objet d'une publication à la conférence à l'international ICIAP 21 (reportée en 2022) (cf. Portfolio RFAI).
  - o Les **modèles et algorithmes sur les graphes** (environ 10 à 20%) : cet axe traite des problématiques de graph-matching, de reconnaissance de formes structurelles et d'apprentissage sur les graphes. Il s'agit d'un axe historique fort de l'équipe avec au moins 3 personnes travaillant sur des problématiques proches et 4 thèses sur la thématique sur le contrat (cf. Annexe B). L'équipe RFAI est très largement reconnue dans ce domaine au niveau national et international (animation du TC-15 de l'IAPR par exemple). On notera que ces travaux sont étendus aux *Graph Neural Networks* et à leurs dérivés (thèse de M. Martineau – cf. Portfolio RFAI -, ANR CodeGNN en démarrage).
  - o Les **algorithmes et méthodes d'apprentissage ou d'optimisation** (env. 30%) : cet axe est en partie le résultat de travaux collaboratifs initiés avec l'équipe ROOT sur le couplage entre recherche opérationnelle et apprentissage statistique. Même si les projets originaux déposés au cours du contrat n'ont pas été acceptés, cette thématique est en croissance (thèse de G. Lacharme, de M. Darwiche, travaux de R. Raveaux ; thèse de S. Mignon sur le transport optimal, thèse à venir sur l'intégration de post-traitements combinatoires dans l'apprentissage) et commence à donner lieu à des publications intéressantes (cf. Portfolio LIFAT). On trouve également dans cet axe des travaux sur des approches biomimétiques (spécificité de l'ex équipe HANT qui a intégré l'équipe RFAI en 2016).
- 2 axes transversaux :
  - o Les **modèles et algorithmes dédiés aux données temporelles** (séries, flux de données, d'images ou vidéos) (environ 10%) : cet axe se concentre sur l'établissement de modèles ou algorithmes qui cherchent à intégrer directement la relation temporelle qui existe au sein des données. Il peut s'agir de données 3D+t ; de vidéos dans lesquelles les frames successives sont traitées en tant que séquences, pour faire du suivi (thèse de Q. C. Le), de la caractérisation de motifs spatio-temporels (thèse de L. P. Nguyen) ; de séries temporelles (thèse de V. Haykal) ou de séquences (cf. Publication revues [79]) ; ou d'algorithmes d'apprentissage sur flux (cf. Publication revue [80])
  - o Les **méthodes intégrant l'utilisateur** (environ 15%) : plutôt que de concevoir uniquement nos algorithmes et méthodes comme des outils produisant une sortie à partir d'une entrée, une partie de nos travaux s'attachent à intégrer leur utilisation dans un environnement dans lequel l'utilisateur est partie prenante. Dans ce contexte, l'interaction avec l'utilisateur ou l'adéquation de l'outil aux contraintes de l'utilisateur est un élément essentiel. Dans ce contexte, nous avons par exemple travaillé sur : des méthodes interactives de segmentation (cf. Portfolio RFAI Sila-3D) ; des méthodes d'apprentissage interactives (cf. Publication revue [80]) ou contraintes (cf. Publication revue [82]) ; ou encore des méthodes de visualisation ou des dispositifs interactifs basés sur la reconnaissance de formes (thèse D. Boas, S. Shah, P. Carvalho)
- 3 domaines d'application (fortement liés au contexte régional) :
  - o **Arts, Humanités et Patrimoine** (environ 15%) : l'activité sur les humanités numériques et le traitement des documents a sensiblement diminué sur le contrat, essentiellement par manque de "soutien" local, régional et national (peu de projets acceptés malgré plus de 15 soumissions avec des consortiums nationaux et internationaux de qualité - Cf. Daccit, Animus, Genealire 1 et 2) ; en revanche, les travaux autour du patrimoine et du tourisme patrimonial se sont largement développés car plus soutenus localement
  - o **Santé et Handicap** (environ 25%) : domaine toujours très soutenu localement et régionalement avec des demandes en croissances. Il s'agit là du domaine d'application principal de nos travaux
  - o **Environnement** : domaine qui émerge à la fois de la volonté d'implication de certains membres face aux enjeux associés, mais aussi du tissu et soutien régional sur cette thématique

On notera que la plupart des thèses (64%) n'ont pas de domaine d'application privilégié. Cela signifie qu'une proportion importante des thèses portent sur des aspects fondamentaux indépendants de contraintes applicatives liées notamment au type de financement. Il s'agit pour la plupart de thèses financées par des bourses académiques ou par des projets non régionaux (NB : cela ne signifie en aucun cas que les thèses CIFRE ou sur bourses régions ne traitent pas de sujets fondamentaux). Ce phénomène est également observé au niveau des publications dans les revues. La répartition des projets, thèses et revues dans ces axes est fournie en annexe B.

### Éléments de rayonnement, Éléments d'attractivité

L'attractivité et le rayonnement de l'équipe se mesure selon différents critères. En premier lieu, l'équipe a été en mesure de monter ou participer au **montage de projets, de l'échelle locale à internationale, avec des partenaires de renom** (cf Annexe B pour la liste des projets acceptés). Ainsi, sur le plan international, l'équipe a soumis (ou participé au montage de) 7 projets internationaux (dont 3 ont été acceptés) : 3 projets européens qui n'ont pas été acceptés (projet LivingLab avec l'Université de Kyushu au Japon, le PhuseLab de l'Université de Milan en Italie et le département de psychologie de l'Université de Jyväskylä en Finlande ; projet AdamIoT avec Kuwait Petroleum Italie et Espagne, Politecnica de Catalunya en Espagne, Vmware International Limited en Irlande, IDS Europe aux Pays-Bas ; 1 ERC) ; 2 Studiums (dont 1 accepté en partenariat avec l'Université de Californie) ; 2 ANR internationales (DACCIT avec l'EPFL, Deepvision, avec deux universités Canadiennes : Vancouver et Guelph ; accepté ; LIFAT représenté indirectement par l'arrivée d'un collègue précédemment au LIRIS) ; 1 programme de financement de thèse accepté (VIED911). Au plan national, 22 projets ANR ont été soumis (5 acceptés) dont certains avec des partenaires internationaux ou nationaux de renom (LOR et LOR2 avec le Politecnico de Turin et le CIRRELT au Canada, Genealire (J-D. Fekete, INRIA Saclay), SIRCUS (INSERM, CHRU Tours, GIS Autisme) et environ une dizaine de nouveaux partenaires.



L'attractivité se mesure également au travers des **mobilités**. Ainsi l'équipe a accueilli des collaborateurs pour une durée cumulée de 24,2 mois et a effectué des visites pour une durée cumulée de 5 mois (cf. Annexe A et données de production et d'activités). Également la très grande majorité des recrutements (MCF, thèses, post-docs) sont des recrutements externes. L'équipe a aussi participé à l'**organisation de conférences** internationales (Gbr, ICPRS, track de l'ACM Symposium on Applied Computing (SAC)), nationales (EGC) et à l'organisation d'une école d'été (Ecole d'été France Excellence Chine 2017). Également, l'équipe est présente dans le **pilotage d'organisations ou associations scientifiques** internationales : TC15 de l'IAPR (chair et newsletter), Evolution Artificielle (présidence, comité des sages), IDAKS, UniTwin CS-DC de l'UNESCO ; nationales : AFRIF (CA, secrétaire), EGC (secrétaire adjoint), GRCE (CA, secrétaire), IFRATH (présidence et CA), Editions RNTI (trésorier, site web) ; régionales : RTR DIAMS (co-direction), ICVL (comité de direction), SFR FED Neuro-Imagerie (comité de direction), Cycle RES (comité de direction). L'équipe est également impliquée dans l'**expertise de projets** (plusieurs par années) pour l'ANR, l'ANRT et à l'international pour Natural Sciences and Engineering Research Council of Canada, Fonds de Recherche du Québec Nature et Technologie, Fondation Nationale Suisse ; et l'**éditions de revues** : RNTI, IARIA Journal - Journal On Advances in Internet Technology, TSI, Elsevier Internet of Things, MDPI Journal of Imaging, Elsevier Pattern Recognition Letters, IEEE Transactions on Emerging Topics in Computing.

L'ensemble de ces activités ont permis à l'équipe à la fois d'entretenir une visibilité et des relations pérennes tout en en développant de nouvelles sur le contrat. Globalement, ces relations et cette visibilité se mesure au travers des **publications avec des co-auteurs nationaux ou internationaux externes** (30 publications internationales et 30 nationales, ce qui représente plus de 50% de la production en collaboration, alors qu'elle n'était que de 20% pour l'international et 16% pour le national au précédent contrat). Si l'on rentre plus en détails dans ces relations, on peut noter certains changements. En effet, une grosse partie de l'activité internationale découlait auparavant des travaux relatifs aux humanités numériques. Comme cette activité est en décroissance (essentiellement du fait de projets déposés non acceptés), les relations se sont distendues avec les partenaires historiques (Inde, Australie, Espagne, Japon). Cependant, ces relations, qui étaient très fortes et qui restent actives au travers de **réseaux** (TC-15 de l'IAPR, IDAKS, l'organisation de conférences), peuvent se réactiver à tout moment. En contrepartie, de **nouvelles relations** se sont créées, notamment avec l'Algérie-Université de Mascara (2 thèses co-encadrées), avec l'Italie-Milan et Salerne (visites et co-

publications régulières), avec le Vietnam-Université de Hong Duc (liens avec d'anciens doctorants en poste, co-encadrements, publications communes, signature d'une MOU et d'accords d'échanges) et avec l'Espagne-Université de Tarragone (visites et co-publications régulières). Sur le plan national, on note peu de changements, avec la conservation des relations avec les laboratoires partenaires comme le LITIS et le Greyc et la participation à l'animation de réseaux et associations (EGC, EA, AFRIF, GRCE, IFRATH), même si les relations avec les partenaires travaillant essentiellement sur les humanités numériques se sont là encore un peu distendues (LIRIS, L3I, LABRI). Finalement, sur le plan régional, on note à la fois un renforcement des relations avec les partenaires habituels (INSERM, INRAE, CHRU, CESR, MSH) grâce à l'animation de réseaux (SFR FED) et au soutien de la région au travers de projets financés, et l'émergence de nombreuses nouvelles relations, grâce à la mise en place et à la participation au RTR DIAMS et à la participation au montage de l'ARD Junon (le tout résultant de la bonne visibilité des travaux précédents de l'équipe). Ainsi on note de nouvelles collaborations avec l'IDP, le BRGM et l'INRAE, le CHRU, le LIFO, le PRISME... Ces relations devraient très largement devenir pérennes.

### Prise en considération des recommandations de la dernière évaluation

Remarques HCERES	Réponses apportées
Production : Efforts à poursuivre, de manière collective et individuelle	Co-encadrements et association aussi souvent que possible des membres de l'équipe aux activités (=> aucun non publiant ce contrat) mais les charges individuelles et les moyens (nombre de bourses thèses) limitent les possibilités
Rayonnement international : à conforter	<ul style="list-style-type: none"> <li>7 soumissions à des programmes ou projets internationaux (3 acceptés)</li> <li>Nouvelles collaborations pérennes avec l'HDU au Vietnam, l'Algérie et l'Italie, maintien d'invitations (malgré COVID)</li> <li>30 publications avec co-auteurs étrangers</li> </ul>
Formation par la recherche : <ul style="list-style-type: none"> <li>nombre de doctorants à conforter (compte tenu du potentiel d'encadrement) ; augmentation du nombre de thèses (notamment CIFRE)</li> <li>durée des thèses importante</li> </ul>	<ul style="list-style-type: none"> <li>Du fait des charges lourdes et de la politique équipe visant à favoriser les co-encadrements et à améliorer la qualité d'encadrement des thèses, la capacité d'encadrement réelle n'est pas proportionnelle au nombre de permanents. Par ailleurs, on note parfois une difficulté à recruter de bons candidats. Aussi, il n'est pas certain qu'une augmentation significative du nombre de thèses permette d'aboutir à une meilleure production (sauf à avoir une réduction de l'activité projet et plus de bourses académiques en échange). Pour les CIFREs, cf. "Collaborations entreprises" plus bas dans le tableau.</li> <li>42,5 mois soit environ 3 ans et demi contre près de 4 ans sur le précédent contrat</li> <li>Aucun abandon en thèse sur ce contrat</li> </ul>
Cohérence thématique : <ul style="list-style-type: none"> <li>ne pas se disperser thématiquement</li> <li>se positionner de manière originale dans la communauté française et internationale</li> <li>Le CETU ne doit pas provoquer de dispersion</li> </ul>	<ul style="list-style-type: none"> <li>Resserrement autour des axes graphes, apprentissage et interactions</li> <li>Singularités scientifiques <ul style="list-style-type: none"> <li>Couplage RO/ML</li> <li>Graphes et réseaux de neurones sur graphes</li> <li>Modèles temporels</li> <li>Données environnementales (Jumeaux numériques)</li> </ul> </li> <li>Projets CETU dans les thématiques de l'équipe (visu, interaction, traitement de vidéos/images, applications en santé et tourisme) ; ressources propres pour la partie ingénierie</li> </ul>
RH : <ul style="list-style-type: none"> <li>charges administratives ou pédagogiques lourdes</li> <li>veiller à la bonne intégration des jeunes MCF</li> </ul>	<ul style="list-style-type: none"> <li>Malheureusement pas de changements, au contraire (cf. liste des charges en annexe B)</li> <li>Politique d'équipe visant à limiter les réunions administratives sur les points essentiels</li> <li>Les nouveaux collègues ont été associés autant que possible sur les thèses. A ce jour, tous les collègues recrutés récemment ont été ou sont impliqués sur un encadrement et/ou sur des projets</li> </ul>
Collaborations entreprises à renforcer (avec publications)	<ul style="list-style-type: none"> <li>1 CIFRE + 1 bourse co-financée région avec un industriel sur le contrat (+ 1 en 2021-2022 + 2 en cours de montage) ; certaines expériences de CIFRE passées ont abouti à un rapport bénéfice/coût discutable car beaucoup d'opportunités sont liées à des startups parfois fragiles et ou à des entreprises sans support en recherche conduisant à des difficultés à atteindre les objectifs visés (manque de support ou données finalement inutilisables). Nous sommes donc vigilants sur ces points et</li> </ul>

	<p>passons beaucoup de temps à discuter avec les industriels sur les opportunités mais avec peu de sujets retenus in fine. Ce point reste donc à améliorer.</p> <ul style="list-style-type: none"> <li>• Activité de valorisation renforcée en collaboration avec le CETU</li> <li>• Les relations entreprises s'effectuent beaucoup au travers des projets régionaux qui intègrent souvent des partenaires industriels</li> </ul>
--	--

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

Depuis le précédent contrat, l'équipe poursuit ses efforts afin de produire une recherche de qualité. Pour cela, plusieurs stratégies sont mises en place.

#### Formation doctorale :

L'encadrement et le suivi des doctorants dans l'équipe est assez important. Ainsi, certains collègues encadrent volontairement peu de doctorants afin de les accompagner pleinement dans leurs travaux et dans leur formation. Cet élément contribue certainement au fait qu'il n'y ait eu aucun abandon en thèse. Également, l'équipe encourage les co-encadrements. Si cela peut rendre l'orientation du travail plus délicate, cela permet d'aller vers une implication de tous les collègues (notamment les nouveaux collègues mais aussi ceux qui ont des charges lourdes). En termes de qualité, cela permet d'avoir les compétences complémentaires nécessaires sur certains sujets mais aussi d'avoir en permanence plusieurs avis critiques sur la formalisation des problèmes et la justesse des solutions, sur la méthodologie et les protocoles expérimentaux, ainsi que sur la rédaction des documents. Également, chaque doctorant fait au moins une présentation à l'équipe avant sa soutenance et nous souhaitons revenir au principe pré-COVID qui consiste à ce que toute présentation en conférence donne lieu à une présentation en équipe.

#### Politique de publication :

En termes de qualité, l'objectif reste avant tout une reconnaissance par les pairs qui passe par la publication dans des journaux et conférences reconnus (cf. Annexe B). Cette recommandation forte est globalement suivie (publications dans Pattern Recognition, Pattern Recognition Letters, Journal of Mathematical Imaging and Vision, IEEE Transactions on Image Processing / ICCV, ECCV, ICIP, ICPR, CVPR...). Cela n'exclue pas les publications dans des conférences ou workshops plus spécialisés permettant une diffusion plus ciblée (IV, ICDAR, GbRPR, AFIHM...). L'objectif pour chaque doctorant est une revue (ou une conférence de bon niveau) avant la fin de la thèse.

#### Pérennisation des travaux et reproductibilité :

Pour faciliter la reprise par d'autres collègues de nos productions et l'expérimentation sur les mêmes données, l'équipe s'engage de plus en plus dans la publication du code et des bases de données sur des Git ouverts (cf. [Site RFAI](#) et fichier de production). Cela devient une pratique presque systématique même s'il faudrait à minima un ingénieur en soutien pour bien formaliser les dépôts et les rendre pérennes.

#### Politique scientifique :

La recherche de la qualité est aussi permise dans une certaine mesure par la politique d'équipe permettant une certaine souplesse quant à la façon de conduire les travaux et les mécanismes d'affectation des bourses de thèse qui permettent d'attaquer des problèmes fondamentaux et autorisent une certaine prise de risques. Ainsi, la plupart des thèses sur financements ANR ou académiques ne sont pas définis explicitement pour répondre à une problématique applicative parmi les domaines applicatifs de l'équipe (même si ces derniers sont souvent considérés dans les expérimentations). Pour chaque bourse académique, chaque EC, HDR ou non, peut proposer un sujet qui est ensuite discuté en équipe. Le choix définitif est fait en fonction de plusieurs critères (qualité, taux d'activité, intégration de jeunes MCF, ...). Également plusieurs collègues font le choix explicite de réduire leur "dépendance" aux AAP pour conduire des travaux plus exploratoires ou spécifiques, ce que l'équipe soutient également dans la mesure de ses moyens. Ainsi, sur le contrat, on peut noter les travaux suivants qui sont assez singuliers et dont une partie a donné lieu à des publications de bon niveau) :

- a. Traitements d'image bas niveau et temps réel (thèses de T. A. Pham, N. D. Cong et H. Le ; publications revues : TIP 2016 [92], JRTIP 2020 [53])
- b. Transport Optimal et meta-learning (thèse Q. Le : conférences [130, 134], S. Mignon, G. Lacharme)
- c. Travaux sur les liens entre apprentissage et RO/optimisation (cf. Portfolio)
- d. Travaux sur la prédiction de séries temporelles et (thèse V. Haykal), leur couplage avec :
  - a. L'analyse d'images et des informations géolocalisées (à petite échelle dans le projet Desherbrob en cours ; à plus grande échelle dans l'ARD Junon en cours)
  - b. Les graphes spatio-temporels (projet CodeGNN en cours)
- e. Travaux mêlant des techniques de fouille et de reconnaissance de formes pour la caractérisation de vidéos par motifs spatio-temporels (thèse L. P. Nguyen)
- f. Travaux sur les graphes et étendus vers l'apprentissage profond sur graphes (thèses Z. Abu-Aisheh, M. Martineau, M. Darwiche, D. Ho). Outre les publications et corpus produits (cf. données de production et Portfolio), une partie de ces travaux ont été inclus dans la librairie [NetworkX](#) très utilisée pour la manipulation de graphes en Python (Github en lien avec cette thématique : [Github1](#) [Github2](#) [Github3](#) et données de production 5.)

**Bilan :**

Finalement, la production de l'équipe se répartit ainsi :

- 50 revues dont 26 de rang Q1 (52%) et 7 de rang Q2 (soit 66% Q1 ou Q2, cf. liste en annexe B). Cette production représente une moyenne de 0,43 revue/EC/an. Si on essaie d'intégrer dans ce calcul le fait qu'un EC est sensé consacrer 50% de son temps à la recherche, ainsi que le temps pris effectivement par les charges lourdes, on se rapprocherait plutôt de 1,25 revues/EC eq. temps plein/an ;
- 73 conférences internationales et nationales et 9 ouvrages ou chapitres. Parmi elles, 9 sont classées A ou A\* et 34 classées B (rang CORE) soit 60% classées B ou plus. Cette production représente 0,64 conférences/EC/an en moyenne ou encore, selon le mode de calcul alternatif 1,87 conférences/EC eq. Temps plein/an ;
- Parmi ces publications, 30 sont avec des co-auteurs étrangers et 30 avec des co-auteurs nationaux.

Par rapport au précédent contrat (même durée) la production brute en revues augmente de 140%. Relativement au nombre de permanents en moyenne sur la période, cela représente une augmentation qui reste modérée (de 0,39 revues/EC/an sur le précédent contrat à 0,43 sur ce contrat) mais qui peut néanmoins être interprétée de façon positive si l'on considère le contexte évoqué précédemment. En revanche, le nombre de publications en conférences baisse de façon assez significative (91 internationales + 27 nationales sur le précédent contrat) ce qui peut s'expliquer en partie par la crise sanitaire et par la chute du nombre de doctorants entre 2019 et 2021 (seulement 23 conférences sur la deuxième moitié du contrat).

Globalement, le niveau de qualité et la visibilité des publications de l'équipe sur la période 2016-2021 sont légèrement supérieurs au précédent contrat : 52% de revues Q1 (contre 50% sur le précédent contrat) ; plus de 25% des publications (surtout revues) avec des co-auteurs étrangers (contre 20% sur le précédent contrat).

### Référence 2. La production scientifique est proportionnée au potentiel de recherche de l'équipe et répartie entre ses personnels.

La production totale de l'équipe pourrait être augmentée par rapport à l'effectif, notamment au niveau des conférences internationales. Ceci étant plusieurs éléments viennent nuancer cela. Tout d'abord, l'apprentissage profond a provoqué une véritable révolution dans nos domaines, avec énormément de soumissions et une compétition très forte. Il est devenu difficile de soumettre sans se comparer aux meilleures approches "deep" souvent produites par les plus gros laboratoires privés et publics. Également, le système de relecture par les pairs est un peu saturé, engendrant pour les revues des délais parfois très importants. Enfin, certaines conférences sont très sélectives (parfois plus que les revues) et le taux de soumission est toujours très élevé.

Également, on note que la crise sanitaire a eu un impact non négligeable avec un report des publications vers les revues. Enfin, ce contrat a aussi été une période particulière avec un effort soutenu vers la valorisation (production de logiciels, montage de startups, valorisation industrielle avec le CETU..., cf. Domaine 4). Enfin, l'amélioration de la qualité reste notre priorité plutôt que la production quantitative.

Pour ce qui est de la répartition de cette production au sein de l'équipe, celle-ci reste très dépendante des charges individuelles ainsi que des publications faites au travers des encadrements de thèses. Comme expliqué plus haut, ces charges diffèrent d'une personne à l'autre et la production de ce fait ne peut être équilibrée entre tous les membres. Par ailleurs, certains membres peinent ou ne souhaitent pas s'investir à nouveau dans une activité de recherche suite à leur investissement dans d'autres tâches (ce qui peut être renforcé à l'approche d'une fin de carrière). On compte également 2 nouveaux permanents qui viennent de nous rejoindre et dont l'activité scientifique est en cours de reconstruction et dont les publications sont plutôt attendues dans les mois à venir. Finalement, seuls quelques collègues parviennent à publier sur leur temps de recherche propre, ce qui est malgré tout positif car nouveau (dans ces proportions) par rapport au contrat précédent. D'une façon générale et comme déjà expliqué, nous cherchons à réduire ces disparités en impliquant le plus possible tous les membres dans les travaux, notamment au travers des co-encadrements et des participations aux projets mais cela ne résoudra pas tous les problèmes de disparités.

### Respect de l'intégrité scientifique, de l'éthique et de la science ouverte

L'ensemble des enseignants-chercheurs de l'équipe sont conscients des problèmes d'intégrité scientifique et des problématiques éthiques. Ces questions sont normalement abordées au travers des encadrements doctoraux et particularisées en fonction des sujets, car cela relève de la responsabilité des encadrants. Ceci étant l'équipe (au travers des membres plus sensibles à ces sujets) contribue également sur ces points au travers de questions lors des séminaires au cours desquels les EC et doctorants présentent leurs travaux, ainsi qu'au travers des sujets abordés en réunion d'équipe. Certains collègues ont également participé plus activement aux réflexions conduites au niveau du laboratoire (charte éthique, séminaire LIFAT sur la souveraineté numérique). De l'information est également régulièrement postée [sur le site de l'équipe](#), avec transmission automatique par mail aux personnes abonnées. Enfin, des efforts particuliers ont été conduits lors de ce contrat pour la publication des articles en open source (**76% des publications**), la **mise à disposition des codes sources et bases expérimentales** pour aller dans le sens de la reproductibilité (cf. Bases de données et codes sur Git dans le fichier de production et d'activités et sur le site de l'équipe)

### Synthèse de l'autoévaluation

	Forces	Faiblesse
Interne	<ul style="list-style-type: none"> <li>- Visibilité sur certaines thématiques qui se mesure notamment au travers               <ul style="list-style-type: none"> <li>- des associations pour le dépôt de projets (graphes, interactions et visualisation, Recherche Opérationnelle /Machine Learning, vision...)</li> <li>- des publications avec des co-auteurs étrangers ou nationaux</li> <li>- de l'animation (sociétés savantes, associations, groupes de recherche)</li> </ul> </li> <li>- Visibilité régionale bonne</li> <li>- Originalité de certains travaux (fast template matching, couplage Recherche Opérationnelle /Machine Learning, fouille et vision...)</li> <li>- Qualité des publications en revues</li> <li>- Nombreuses collaborations pluridisciplinaires (INSERM, INRAE, CHRU, CESR, CITERES/MSH, BRGM...)</li> <li>- Quelques partenariats internationaux forts (Hong Duc University au Vietnam, Italie)</li> <li>- Production logicielle et valorisation</li> <li>- Aucun membre non publiant sur le contrat</li> <li>- Aucun abandon de thèses sur le contrat</li> </ul>	<ul style="list-style-type: none"> <li>- Volumétrie améliorable pour les publications en conférences</li> <li>- Hétérogénéité de la production dans l'équipe</li> <li>- Publications avec des co-auteurs internationaux à diversifier et renforcer</li> </ul>
	Opportunités	Menaces
Externes	<ul style="list-style-type: none"> <li>- Thématiques (Apprentissage Machine, Reconnaissance des Formes, Analyse d'Images) très porteuses</li> <li>- Collaborations inter-équipes</li> <li>- implication dans les réseaux régionaux ; participation au dépôt de projets structurants (PIA4 et ARD Junon)</li> <li>- Tissu régional (santé, environnement, tourisme)</li> </ul>	<ul style="list-style-type: none"> <li>- Pérennisation des travaux et publications sur des domaines où l'on gagne en compétences et visibilité (cf risque lié à la dépendance aux projets)</li> <li>- domaine lié fortement au machine learning et donc très compétitif</li> <li>- Nombre de doctorants fluctuant et publications relativement liées aux encadrements de thèse (un peu plus de 50%)</li> <li>- Charges de certains collègues ne permettant pas une production homogène entre membres, ni un taux d'encadrement plus important</li> <li>- Manque d'un ingénieur recherche pérenne pour appuyer la valorisation, assurée en partie les EC</li> <li>- Cohérence et cohésion de l'équipe difficile à maintenir, à cause de son histoire, de sa dispersion géographique, de la disponibilité de chacun</li> </ul>

### Projet de recherche

Pour le futur projet de l'équipe, nous ne prévoyons pas de changements majeurs quant au fonctionnement (et à la politique de l'équipe) ni quant aux axes de recherche. Seule l'importance de ceux-ci est susceptible d'évoluer.

Pour les axes majeurs :

- modèles et algorithmes d'analyse d'images et de vidéos : cet axe historique, qui fédère une grande partie des chercheurs de l'équipe va évidemment perdurer avec ses spécificités (analyse d'images temps réel, reconnaissance d'action dans les vidéos, architectures et modèles spécifiques -deep ou non deep). Il pourrait en revanche devenir un axe transversal mettant en les compétences de l'équipe sur ce type de données.
- modèles et algorithmes sur les graphes : ce deuxième axe historique très fort va également se poursuivre avec des travaux qui vont se renforcer autour des Graph Neural Networks et des graphes dynamiques ou spatio-temporels (lien avec l'axe données temporelles en plus du lien déjà existant avec l'axe analyse d'images). Ces travaux sont déjà engagés au travers de notre participation à l'ANR CodeGNN
- algorithmes et méthodes d'apprentissage ou d'optimisation : cet axe qui a émergé au cours du contrat, qui intègre notamment des collaborations avec l'équipe ROOT autour du couplage entre Recherche Opérationnelle et Machine Learning est également susceptible de se renforcer avec environ 4 permanents travaillant sur le sujet. Une thèse sur le sujet est déjà prévue pour septembre 2022.

Pour les axes transversaux (qui singularisent les travaux de l'équipe) :

- Les méthodes intégrant l'utilisateur : cet axe qui vise à produire des outils pour les utilisateurs et qui intègrent ceux-ci dans le processus d'analyse et de décision va également perdurer et se renforcer. C'était déjà historiquement un point fort de

l'équipe et il est devenu encore plus indispensable aujourd'hui avec la pénétration de l'IA (du deep learning notamment) dans la société et en regard de nos domaines applicatifs (santé et handicap notamment). Si cet axe devait intégrer des problématiques d'explicabilité des méthodes et outils lors du contrat qui s'achève, cela n'a pu se faire. Cet aspect devrait enfin pouvoir émerger au travers des projets qui vont être lancés (projet OPTIMEDIAS en santé, projet d'ARD Junon en environnement).

- Les modèles et algorithmes dédiés aux données temporelles : là encore, cet axe qui devait se renforcer lors du contrat qui s'achève, est resté en sommeil, faute de financements. Pour le contrat à venir, il devrait prendre une importance considérable avec les projets qui sont déjà en cours de lancement : Projet région Desherbrob (une partie de thèse sur la prédiction géolocalisée de mauvaises herbes), ANR CodeGNN (1 thèse sur les graphes dynamiques), ARD Junon (1 thèse et 2 postdocs sur la prédiction de séries temporelles), Projet OPTIMEDIAS (1 thèse CIFRE incluant la prédiction pronostique pour la sclérose latérale amyotrophique)

Enfin, en ce qui concerne les domaines d'application (fortement liés au contexte régional) :

- Arts, Humanités et Patrimoine : cet axe va sans doute encore se poursuivre pendant le contrat à venir, notamment autour d'activités de transfert, en collaboration avec le CETU. Cependant, il est possible que les financements (notamment région) se tarissent dans ce domaine (fin de l'ARD Ipat par ex.). Enfin, il ne reste plus qu'une personne travaillant fortement dans le domaine (et quelques autres plus ponctuellement). Il est donc possible que cet axe se réduise considérablement
- Santé et Handicap : cet axe applicatif fort (lié notamment à l'activité régionale CHRU, INSERM, INRAE, livingLab,...) va perdurer et se renforcer notamment suite au Studium et à l'ANR Sircus qui viennent de démarrer mais aussi au travers du projet OPTIMEDIAS. Ce renforcement va se faire en ajoutant aux données images sur lesquelles nous travaillons déjà des travaux sur d'autres données de santé (clinique, biologique) qui sont désormais accessible au travers des grandes plateformes régionales (HUGO pour le grand ouest). Cela devrait permettre d'ajouter à nos méthodes et outils la possibilité de travailler dans ce domaine en intégrant des données multimodales, point crucial pour l'évolution des outils destinés à la santé.
- Environnement : cet axe applicatif est seulement en émergence et porté par 1 personne (avec le renfort de 2 ou 3 collègues). Il sera présent au cours du contrat à venir notamment grâce aux avancées qui seront faites dans le projet Desherbrob et au travers de l'ARD Junon. Cependant, il est difficile de dire quelle sera son évolution, notamment s'il agrégera plus fortement d'autres collègues.

### 3.4. Equipe ROOT

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

##### Préambule

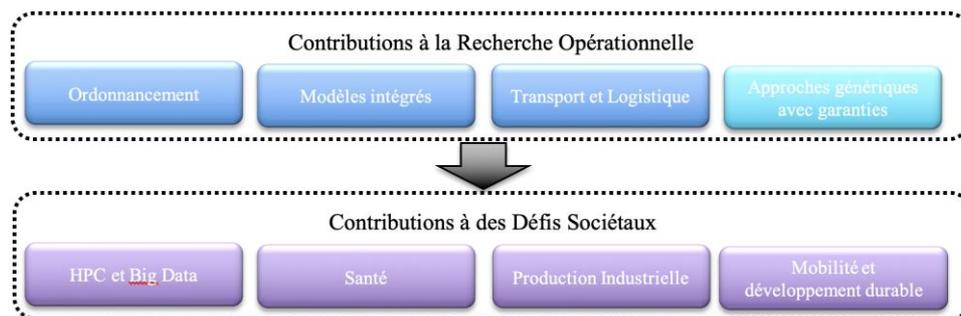
L'équipe Recherche Opérationnelle, Ordonnancement et Transport (EMR-CNRS ROOT), composée aujourd'hui de 13 membres permanents, est issue de l'équipe Ordonnancement et Conduite. Ce changement de nom date de 2016 et reflète une évolution dans les thématiques scientifiques sur lesquelles les membres de l'équipe travaillent. A l'origine, les activités de recherche de l'équipe s'articulaient autour de deux grands domaines de la Recherche Opérationnelle (RO) : l'ordonnancement et la conduite, principalement appliqués aux systèmes de production. Au fil des années, les projets ainsi que les compétences de l'équipe se sont élargis, montrant un ancrage toujours plus important dans le domaine de la Recherche Opérationnelle, mais avec des contributions à la résolution de modèles plus génériques d'ordonnancement et/ou de transport, ainsi que des contributions au champ des méthodes.

L'équipe est aujourd'hui une des équipes les plus importantes en Europe dans les domaines de l'ordonnancement et du transport.

##### Activités scientifiques depuis 2016

Depuis 2016, les activités scientifiques de l'équipe ont continué à s'inscrire dans le domaine de la Recherche Opérationnelle (RO), comme indiqué dans la figure ci-dessous. L'équipe a poursuivi ses contributions au domaine de la RO en participant au développement scientifique du champ des méthodes de la RO ou du champ des modèles. Les principaux modèles abordés relèvent des problèmes d'ordonnancement et/ou des problèmes de transport et de logistique. Les recherches menées au sein de l'équipe se sont inscrites aussi bien dans un cadre académique que dans un cadre de résolution de problèmes plus appliqués. Soulignons que l'étude et la résolution de certains problèmes appliqués ont permis de proposer des éléments de réponse à quatre grands défis sociétaux (en accord avec les défis de la SNR<sup>1</sup> définis en 2014) qui sont :

- **Systèmes de calcul haute performance (HPC)** et de traitement de grands volumes de données (Big Data) avec un focus sur l'optimisation de ressources de calcul dans les systèmes informatiques, comme par exemple les data centers.
- **Santé** : garantir un système de santé performant est un enjeu majeur de notre société. Les contributions à ce défi se positionnent dans le cadre de l'optimisation de l'organisation du système de Santé.
- **Mobilité et développement durable** : favoriser la mobilité et le développement durable sont deux préoccupations au cœur de ce défi avec un focus sur des problématiques relevant de l'optimisation des transports (de personnes et de marchandises), du développement durable (réduction des émissions de GES, utilisation de véhicules électriques, ...) ou bien des deux simultanément.
- **Production Industrielle** : dans le cadre du renouveau industriel, les contributions à ce défi concernent l'optimisation des ressources au sein des systèmes de production dans le but de les rendre plus efficaces et plus flexibles.



Afin d'étayer la production scientifique de l'équipe, le tableau ci-dessous liste les principaux projets scientifiques qui ont été menés sur la période 2016-2021 dont certains sont encore en cours (étiquette **Actif**). Dans ce tableau sont mentionnés les principaux sujets de recherche, le nombre de publications dans une revue internationale et en conférence (notés respectivement R et C), le nombre de thèses, les éventuels organismes de financement et les champs de contribution en lien avec le schéma ci-dessus.

Intitulé du projet	# Publi.	#Thèses	Financement	Champs
Problèmes théoriques d'ordonnancement (méthodes exactes et approchées, caractérisation de solutions optimales) <b>Actif</b>	13R ; 10C	3		Ordo AppGen
Diagnostic de systèmes à l'aide de réseaux de Petri temporisés <b>Actif</b>	3R ; 5C			Ordo ProdIndus

Approches de résolution de problèmes d'ordonnement multi-agents sur des machines parallèles ou dans des systèmes distribués (cf. élément 5 du portfolio ROOT)	1R ; 7C	2		Ordo HPC&Data
Optimisation de la gestion des ressources sur une plateforme de type « Big Data » basée sur le logiciel Hadoop	6C	1	CIFRE	Ordo HPC&Data
Planification de parcours patients <b>Actif</b>	2C	2		Ordo Santé
Étude et résolution de problèmes intégrés d'ordonnement de la production et distribution <b>Actif</b> (cf. élément 1 du portfolio ROOT)	3R ; 9C	2	ANR	ModèlesInt ProdIndus
Résolution de problèmes intégrés dans un contexte de production et livraison de chimiothérapies (cf. élément 1 du portfolio ROOT)	2R ; 10C	1	Région	ModèlesInt Santé
Problèmes de tournées de véhicules sous incertitude	5R ; 3C	1		Trsp&Log
Étude et résolution de problèmes de tournées de véhicules électriques (cf. élément 3 du portfolio ROOT)	6R ; 14C	3	ANR JCJC	Trsp&Log Mobi&DevDu
Aide à la décision pour la conception d'itinéraires touristiques personnalisés	4C	0	Région	Trsp&Log Mobi&DevDu
Calcul de chemins pour la conception de plans d'évacuations de personnes	3R	1		Trsp&Log Mobi&DevDu
Calcul des plus courts chemins multi-objectifs	1R ; 8C	1	CIFRE	Trsp&Log Mobi&DevDu
Modèles mathématiques et algorithmes pour la résolution du problème de tournées dans le milieu de la santé (soins à domicile ou gestion de flotte d'ambulance)	3R ; 6C	1	ANR JCJC	Trsp&Log Santé
Modèles et méthodes pour l'optimisation du processus de picking dans les entrepôts de e-commerce <b>Actif</b>	1R ; 2C	1		Trsp&Log ProdIndus
Recherche Opérationnelle et Reconnaissance des Formes : indexation et appariement de graphes	3R ; 8C	2		

### Production et qualité scientifique

La production scientifique de l'équipe ROOT se concrétise majoritairement par des publications, soit dans des revues internationales de renommée dans le domaine de la RO, soit dans des conférences nationales ou internationales. La politique de publication de l'équipe est 1) de mettre en valeur les contributions scientifiques par des publications d'articles dans des revues internationales et très sélectives du domaine et 2) de communiquer sur les travaux réalisés ou encore en cours dans des conférences afin de se positionner dans la communauté et d'en partager les principaux résultats. Cependant, les conférences visées ne sont pas nécessairement sélectives puisque l'objectif est de toucher la communauté la plus large possible du domaine. De plus, il n'existe pas beaucoup de conférences référencées comme très sélectives dans le domaine de la RO. Concernant les publications en revues, leur nombre oscille entre 5 et 13 par an, principalement dans des revues internationales, avec une moyenne annuelle de 8,8 publications/an (53 publications sur la période), ce qui est similaire au taux de publication sur la précédente période. Ce maintien quantitatif de la production scientifique en revues ne s'est pas fait au détriment de la qualité des revues dans lesquelles l'équipe publie, au contraire. Le tableau suivant indique la répartition des publications en revue de l'équipe selon deux métriques connues : Scimago Journal & country Rank (SJR) et QUALIS.

Quartile SJR									
	Q1	Q2	Q3	Q4	Non réf.				
Nombre d'articles	32	17	2	1	1				
QUALIS									
	A1	A2	B1	B2	B3	B4	B5	C	Non réf.
Nombre d'articles	19	17	3	6	1	0	1	4	2

Ces chiffres montrent que l'équipe ROOT publie dans les revues les plus visibles du domaine. Au-delà de la visibilité, il est important de noter que des articles sont publiés dans des revues très sélectives (notamment 6 dans Computers and Operations Research, 5

dans *European Journal of Operational Research* et 5 dans *Journal of Scheduling*) voire extrêmement sélectives (3 dans *Transportation Science* et 1 dans *Transportation Research Part B: Methodological*). Tous ces éléments indiquent un très bon niveau de publication en revues tant quantitativement que qualitativement, et une bonne reconnaissance des travaux de recherche de l'équipe ROOT.

Le nombre de communications dans des conférences est, en moyenne, de 21,8 communications par an (131 communications sur la période), 60,3% d'entre elles dans des conférences internationales. Ces résultats sont en baisse par rapport à la période précédente qui était de 34,5 communications par an en moyenne. Cette baisse est due en partie à la crise sanitaire du COVID-19. Les conférences dans lesquelles l'équipe communique sont des conférences reconnues dans leur domaine : peu sont référencées dans les grandes bases de données scientifiques, mais elles sont très visibles dans le domaine de la RO. C'est notamment le cas des conférences : Control Decision and Information Technologies (CoDIT), International Symposium on Combinatorial Optimization (ISCO), Models and Algorithms for Planning and Scheduling Problems (MAPSP), Metaheuristics International Conference (MIC), Project Management and Scheduling (PMS) et International Workshop on Freight Transportation and Logistics (ODYSSEUS). Comme le montre la volumétrie et la répartition des communications dans des conférences (voir le fichier des données de production et d'activité attaché à ce dossier), l'équipe ROOT est bien insérée dans ses communautés au niveau national et international. L'équipe a également reçu une récompense : en 2017, un doctorant de l'équipe a obtenu le prix du « Best Student Paper Award » à la 7th International Conference on Industrial Engineering and Systems Management.

Les travaux de recherche théoriques ou appliqués de l'équipe ROOT ont parfois conduit à la réalisation de solutions technologiques, qui ont pour la majorité été développées dans le cadre de collaborations industrielles via le dispositif CIFRE. Ces productions sont décrites au DAE4.R1 de l'équipe ROOT. D'autres productions technologiques ont été conçues pour illustrer et/ou utiliser les méthodes de résolution de problèmes combinatoires développées par l'équipe, sous forme d'outils d'aide à la décision. Par exemple, le projet ANR JCJC PONOSaD s'intéressant à la planification et l'optimisation pour les soins à domicile et porté par un membre de l'équipe, a donné lieu à un prototype logiciel de gestion et de planification de rendez-vous en temps réel pour les coordinateurs de soins mais aussi les intervenants à domicile via leur smartphone. Ce prototype (source et documentation) est disponible sur le site web du projet <http://ponosad.projets.li.univ-tours.fr/> (en open source). L'équipe a aussi participé activement au développement et au maintien d'une plateforme web collaborative VRP-REP (<http://www.vrp-rep.org>) pour la gestion et le partage de données de la littérature pour les problèmes de tournées de véhicules, c.-à-d., le téléchargement de jeux de données et des tableaux de résultats, la recherche de références bibliographiques par jeux de données ou type de problème, la mise au format universel XML pour la représentation des jeux de données et des bibliothèques de code open source pour la lecture/écriture de données et la vérification de solutions. Enfin, dans le cadre d'un projet financé par la région Centre et dans la continuité d'une longue série de travaux de recherche, l'équipe a travaillé sur la mise à jour d'un logiciel d'aide à la décision et à la traçabilité pour la production de chimiothérapies afin d'y intégrer la partie livraison. Ces travaux se sont déroulés en collaboration avec la société Eticsys également partenaire dans le cadre d'un projet financé par la Région Centre-Val de Loire (<https://www.groupeprisme.com/lugap-et-le-groupe-prisme-innovent-pour-les-hopitaux-et-les-patients-atteints-de-cancer/>).

Enfin, la dernière grande catégorie de production scientifique de l'équipe concerne les thèses de doctorat. Sur la période 2016-21, 16 thèses de doctorats ont été soutenues, 10 thèses sont en cours et 3 doctorants ont abandonnés pour raisons personnelles. Ces chiffres sont similaires à la période précédente. Les deux tableaux de l'annexe C listent l'ensemble des thèses soutenues, abandonnées et en cours de l'équipe. A noter qu'en 2021, N. Kullman, a reçu le prix de la meilleure thèse GT2L (GT Transport et Logistique du GdR RO) pour ses travaux sur la prise de décision dynamique en cas d'incertitude dans l'acheminement des véhicules et la logistique.

Pour conclure, l'équipe a maintenu sa production scientifique pratiquement au même niveau que lors du précédent contrat (avec une légère baisse de communications en conférence). Les manuscrits de toutes les thèses soutenues sur la période sont disponibles sur HAL. Toutes les publications (revues, conférences et ouvrages) sont référencées sous HAL, avec 68% des articles publiés en revues disponibles au téléchargement sur HAL et 56% des articles de conférences disponibles au téléchargement sur HAL.

### Collaborations nationales et internationales

Les travaux de recherche de l'équipe sont souvent menés avec des chercheurs d'autres laboratoires en France (national) ou avec des chercheurs étrangers (international) et de plus en plus en collaboration avec des chercheurs des autres équipes du laboratoire ou d'autres laboratoires de l'université de Tours (local).

**Au niveau international**, l'équipe ROOT est également très active pour maintenir ou nouer de nouvelles collaborations avec des laboratoires étrangers. Les principaux indicateurs de cette activité sont les séjours scientifiques entrants ou sortants, les publications communes et les thèses de doctorat en cotutelle. Sur la période 2016-2021, l'équipe a invité et accueilli plusieurs scientifiques étrangers, notamment :

- D'Amérique du Nord :
  - Pr. Angel Ruiz de l'université de Laval au Canada (1 semaine en 2017)
  - Pr. Gilbert Laporte d'HEC Montréal (2 jours en 2017)
  - Pr. Nadia Larichi de Polytechnique Montréal (2 semaines en 2017)
  - Ass. Pr. Valérie Bélanger d'HEC Montréal (1 semaine en 2017)

- Pr. Justin Goodson de l'université de St Louis aux USA (6 semaines en 2016 et 2 semaines en 2017)
- D'Amérique du Sud :
  - Pr. Juan Villegas de l'université de Antioquia en Colombie (1 semaine en 2016)
- D'Italie :
  - Pr. Federico Della Croce de Politecnico di Torino en Italie (1,5 semaines en 2017, 5 semaines en 2021)
  - Pr. Ola Jabali de Politecnico di Milano en Italie (1 semaine en 2018)
  - Ass. Pr. Rosario Scatamacchia de Politecnico di Torino en Italie (1 mois en 2018)
- D'Afrique :
  - Ass Pr. Cheick Dhib de l'Institut Supérieur de Comptabilité et d'administration des Entreprises de Nouakchott (2 mois en 2018)
  - Dr. Hafedh Mohammed Babou de l'école Polytechnique de Nouakchott en Mauritanie (2,5 mois en 2017)
  - Dr. Amir Abdellatif de l'école d'Ingénieurs de Monastir en Tunisie (5 mois en 2019 et 3 mois en 2020)
- Du Vietnam :
  - Ass. Pr. Trang Hong Son de l'université de Technologie de Ho Chi Minh Ville (4 mois en 2017)

L'équipe a également accueilli plusieurs étudiants en doctorat de Politecnico di Torino (Italie) : Riccardo Schiano (2 mois en 2018) et Elena Renner (1,5 mois en 2019). Plusieurs membres de l'équipe ROOT se sont également déplacés à l'étranger pour des séjours scientifiques, le tableau ci-dessous liste les lieux, l'année et la durée totale sur l'année des séjours.

Lieux	Année	Durée totale
CIRRELT (Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport) de Montréal	2016	1 semaine
CRM (Centre de Recherche en Mathématiques) de Montréal (délégation CNRS)	2017	6 mois
Politecnico di Torino (Italie)	2017	1 mois
	2018	7 semaines
	2019	7 semaines
	2020	7 semaines
	2021	14 semaines
Université de Los Andes (Bogota, Colombie)	2017	1 semaine
Université de Technologie de Ho Chi Minh Ville	2017	1 semaine
Université of St Louis (USA)	2018	10 jours

Sur la période, les collaborations internationales se sont traduites par : 64% (34) publications en revue de l'équipe et 41% (54) publications en conférence de l'équipe co-signées par au moins un membre de l'équipe ROOT et un membre d'un autre laboratoire étranger, signe de rayonnement à l'international des activités de l'équipe. L'équipe peut également noter 5 thèses en cotutelles sur la période avec les établissements suivants : HEC Montréal (1), Polytechnique Montréal (1), l'université de Autónoma de Nuevo León au Mexique (1) et l'université de Nouakchott en Mauritanie (2). En plus des chercheurs mentionnés lors des mobilités entrantes, l'équipe peut également mentionner quelques collaborations fortes pour lesquelles il y a eu des publications communes : Alessandro Agnetis (Université de Sienne, Italie), Christos Koulamas (Florida International University, USA), Fabio Salassa (Politecnico di Torino, Italie), Louis-Martin Rousseau (Polytechnique Montréal, Canada), Michel Gendreau (Polytechnique Montréal, Canada), Nguyen Huynh-Tuong (Ho Chi Minh University of Technology, Vietnam), Sonja Rohmer (HEC Montréal, Canada), et Vincent Boyer (Universidad Autónoma de Nuevo León, Mexique).

**Au niveau national**, l'équipe a mené et mène plusieurs travaux de recherche avec des chercheurs affiliés à d'autres laboratoires en France. Ces collaborations peuvent prendre différentes formes, le plus souvent étant des projets financés par un organisme (le tableau ci-dessous liste les projets financés au niveau national ou régional) ou des codirections de thèse (2 thèses ont été en codirection avec le LARIS d'Angers, 1 thèse a été en codirection avec le laboratoire HeuDiasyC de Compiègne et 1 est en codirection avec le LIMOS – antenne de Gardanne).

Acronyme	Intitulé	Partenaires	Financement	Période
<b>ATHENA</b>	<i>Méthodes d'optimisation pour l'étude intégrée de problèmes décisionnels complexes</i>	Porteur : ROOT-LIFAT + Heudiasyc (Compiègne), LAAS-CNRS (Toulouse), et LIMOS (Clermont-Ferrand)	ANR 348k€	2013-2017
<b>PONOSaD</b>	<i>Planification et optimisation au niveau opérationnel pour les soins à domicile</i>	Porteur : ROOT-LIFAT (Collaboration avec le CIRRELT de Montréal)	ANR JCJC 131k€	2013-2017

<b>e-VRO</b>	<i>Electric vehicle routing optimization</i>	Porteur : ROOT-LIFAT (Collaboration avec le LARIS d'Angers, St Louis University aux USA et l'Universidad de Antioquia en Colombie).	ANR JCJC 222k€	2015-2019
<b>DISTRICHIMIO</b>	<i>Ordonnement et distribution des chimiothérapies</i>	Porteur : ROOT-LIFAT + le CHRU de Tours et la société Eticsys	Région CVL 140k€	2017-2020
<b>ETHNOS</b>	<i>Ethique et IA en Santé</i>	Porteur : LIFO (Orléans) + ROOT-LIFAT	Région CVL 30k€	2019-2021
<b>SIOMedic</b>	<i>Sécurité des objets connectés médicaux</i>	Porteur : LIFO (Orléans) + ROOT-LIFAT	Région CVL 100k€ + INSA CLV 15k€	2021-2023

Les collaborations scientifiques ne se résument pas aux projets financés et aux codirections de thèse, mais passent aussi par relations privilégiées avec des chercheurs ou chercheuses expert(e)s dans des domaines scientifiques complémentaires aux membres de l'équipe. Parmi ces relations qui ont donné lieu à des publications communes ou des séjours/visites (entrants ou sortants), nous pouvons en citer quelques-unes majeures : le LIFO (Orléans) avec Mathieu Liedloff, le LAAS (Toulouse) avec Christian Artigues et Lotte Berghman, Heudiasyc (Compiègne) avec Antoine Jouglet, le LARIS (Angers) avec Christelle Guéret et Éric Pinson, le CERMICS (Paris) avec Axel Parmentier et le LIMOS (St Etienne et Gardanne) avec Thierry Garaix et Claude Yugma. Sur la période, les nombreuses collaborations nationales se sont traduites par : 17 publications en revue et 30 publications en conférence co-signées par au moins un membre de l'équipe ROOT et un membre d'un autre laboratoire en France.

**Au niveau local**, l'équipe ROOT a mené et mène plusieurs projets en informatique avec des collègues des autres équipes du LIFAT (équipes BDTLN et RFAI). Les principaux projets portent sur la résolution de problèmes combinatoires de graphes pour le traitement d'images, l'optimisation de requêtes sur des bases de données, l'optimisation de techniques d'apprentissage dans les réseaux de neurones, ou encore l'utilisation de techniques d'apprentissage pour l'optimisation de problèmes combinatoires. Ces collaborations fructueuses se traduisent sur la période par :

- 2 thèses soutenues co-encadrées par au moins un membre de l'équipe ROOT et un membre de l'équipe BDTLN ou RFAI
- 2 thèses en cours co-encadrées par au moins un membre de l'équipe ROOT et un membre de l'équipe BDTLN ou RFAI,
- 3 publications en revue et 10 publications en conférence co-signées par au moins un membre de l'équipe ROOT et un membre de l'équipe BDTLN ou RFAI.

L'équipe a également collaboré avec d'autres laboratoires de l'université, notamment le GREMAN (laboratoire UMR-CNRS multidisciplinaire en matériaux, microélectronique, acoustique et nanotechnologies) pour la modélisation et l'optimisation du canal CPL (courant porteur en ligne) et le calcul de fonction de transfert (2 publications en commun) et le laboratoire IDP (UMR-CNRS Institut Denis Poisson, en mathématiques) pour la résolution exacte de problèmes d'ordonnement sur machines parallèles par des techniques dites d'« Inclusion-Exclusion » (participation d'un chercheur en mathématique à l'encadrement d'une thèse).

Enfin, toujours au niveau local, des membres de l'équipe ont également participé au projet région SMARTLOIRE porté par l'équipe du LIFAT BDTLN et portant sur l'organisation de séjours de tourisme culturel en Région Centre-Val de Loire par composition de services Web (projet financé à hauteur de 196 k€ pour la période 2017-2020). L'équipe a été responsable d'un lot traitant de la conception d'itinéraires touristiques personnalisés et optimisés. Ce projet a permis à l'équipe de recruter un post-doctorant pour 9 mois. L'équipe est impliquée dans un nouveau projet région déposé et accepté en 2021. Il s'agit du projet SQUALD porté par l'équipe BDTLN portant sur le suivi de qualité de vie de patients d'affections de longue durée (financé à hauteur de 200 k€ pour 3 ans).

Comparé à la période du précédent contrat et malgré deux ans de crise sanitaire, l'équipe ROOT a maintenu, voire a accru cette diversité des collaborations, montrant ainsi une bonne implantation de l'équipe au niveau local, national et international.

### Implication dans la communauté et rayonnement national et international

La production des connaissances et l'expertise scientifique des membres de l'équipe ROOT se dessinent également au travers d'autres activités telles que la participation à des réseaux scientifiques ou des sociétés savantes, des activités d'évaluation d'articles, des participations à des jurys de thèse ou HDR, des participations à des comités de programme ou d'organisation de conférence, des activités éditoriales ou encore de direction d'ouvrage.

Tout d'abord, l'équipe s'est fortement investie dans l'animation de structures à vocation scientifique. Sur la période, nous pouvons relever les investissements suivants :

- EURO Working Group PMS : Depuis 2014, l'équipe possède un membre de l'advisory board du groupe de travail Project Management and Scheduling (PMS) d'EURO, l'association des sociétés européennes de Recherche Opérationnelle. Ce groupe de travail se réunit une fois par an lors de l'organisation de la conférence internationale PMS.

- GDR Recherche Opérationnelle : De 2014 à 2018, un membre de l'équipe était le directeur adjoint de ce GDR. Depuis la recontractualisation du GDR RO avec le CNRS en 2020, l'équipe y est impliquée puisque deux membres sont co-responsables d'un axe parmi les sept axes de ce GDR (l'axe OPA, Ordonnancement, Planification et Applications, et l'axe REST, Réseaux, Énergie Services, Transport).
- ROADEF : L'équipe est également impliquée dans la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF) au travers d'un membre ayant le rôle de deuxième vice-président de l'association depuis 2020 (chargé de toute la partie web et de son animation : communication, paiement adhésion, forum, inscription, etc.).
- GT ROSA : En 2014, un membre de l'équipe a créé avec un collègue de l'école des Mines de St Etienne, le groupe de travail ROSA (Recherche Opérationnelle et Santé) du GDR RO qu'il co-anime toujours. Ce groupe se réunit 1 à 2 fois par an et propose des exposés scientifiques autour de la RO et de la Santé. La dernière en date du 11 juin 2021 s'est déroulée sous forme de webinaire sur le thème "l'aide à la décision et la RO face à la Covid-19" (sept intervenants, une quarantaine de participants).
- CNU : L'équipe possède deux membres élus au Conseil National des Universités de la section informatique (CNU 27) depuis 2016 (donc deux mandats pour chacun d'eux). A ce titre, un membre a participé à un comité d'expert HCERES (pour le Centre de Recherche en Informatique de Lens) en 2019.
- CaSciModOT : Un membre de l'équipe est impliqué dans la fédération CaSciModOT qui est une structure d'échange autour de la simulation, de la modélisation, de l'algorithmique, et du traitement de données massives. CaSciModOT est maintenant un cluster de la Région Centre-Val de Loire avec un comité de pilotage constitué de représentants de laboratoire. Le membre de l'équipe y est au titre du LIFAT. Récemment, il a co-organisé la 34e journée du projet CaSciModOT à Tours en 2021.

Les membres de l'équipe s'impliquent aussi dans la communauté scientifique au travers de :

- Comités d'organisation de conférences :
  - Conférences internationales :
    - VEROLOG en 2016,
    - Matheuristic en 2018.
  - Manifestations nationales :
    - 1 journée du GT2L du GDR RO en 2016,
    - 1 journée ANR e-VRO en 2017,
    - 2 journées conjointes des GT GOTHa et Bermudes du GDR RO en 2017,
    - 1 journée thématique HPDA LIFAT-LIFO en 2018,
    - 1 journée de l'axe REST du GDR RO en 202,
    - 7 journées du GT ROSA du GDR RO en 2016, 2017, 2018, 2019, 2020 et 2021.
- Comités scientifiques ou comité de programmes :

Année	Conférences internationales	Conférences nationales
2016	CoDIT, PMS et VECoS	COSI, GISEH, MOSIM, ROADEF et VEROLOG
2017	IESM, ISOR, MIC et MOPGP	ASOCIO, SBPO et ROADEF
2018	CoDIT, ICFDNS, Matheuristic, PMS, SCALCOM, MOPGP et VECoS	GISEH, MOSIM et ROADEF
2019	MOPGP	ROADEF
2020	PMS	GISEH, MOSIM et ROADEF
2021	IEEE SCC, IFIP-APMS et VECoS	ROADEF

- Interventions des membres de l'équipe en tant que conférenciers invités :
  - Lors de l'école d'été Jeunes Chercheurs du GDR RO en 2021 à Paris pour donner un tutoriel et des TP en programmation par contraintes.
  - Lors de l'école d'été Jeunes Chercheurs du GDR RO en 2018 à Troyes (International spring school on Integrated Operational Problems) pour donner des cours et un projet sur "Integrated Scheduling and routing".
  - Lors de l'école d'été Jeunes Chercheurs du GT2L, et en collaboration avec 3 projets ANR (ATHENA, EVERS, et e-VRO), en 2017 à Gardannes pour donner deux exposés liés aux projets ANR ATHENA et e-VRO portés par des membres de l'équipe.
  - Lors de l'école Jeunes Chercheurs du GDR RO en 2016 à Angers pour donner des cours en Ordonnancement.

L'équipe participe à des activités éditoriales dans des revues internationales de renommées dans le domaine puisqu'elle a parmi ses membres des éditeurs associés dans les revues Algorithms (Q3) depuis 2021, Journal of Scheduling (Q1) depuis 2015, et OMEGA (Q1) jusqu'en 2017, ainsi qu'un membre du bureau éditorial de la revue European Journal of Operational Research (Q1) depuis 2015, qui a également participé au jury du « Best Paper Award » de cette revue en 2021.

Enfin les membres de l'équipe possèdent une certaine reconnaissance dans leurs domaines de recherche puisqu'ils sont souvent sollicités pour participer à de nombreux travaux de relectures pour des revues ou conférences dans une volumétrie similaire à celle du précédent contrat (94 relectures pour des articles dans des revues dont au moins les trois quarts sont évalués Q1 sur SJR, et 38

relectures d'articles longs pour des conférences) ainsi qu'à des jurys de thèse (25 en tant qu'examineur et 17 en tant que rapporteur), à des jurys d'HDR (4 en tant que rapporteur) ou encore à des jurys de prix nationaux (prix mémoire de master RO en 2019 et prix de thèse PGMO en 2020).

### Prise en compte des recommandations du précédent rapport

<< Critère 1 : qualité et production scientifiques : L'équipe est encouragée à maintenir ses activités de publication dans les revues répertoriées de haut niveau et viser des congrès sélectifs à fort impact. La collaboration au sein du laboratoire pourrait être un facteur bénéfique pour créer une dynamique nouvelle. Par ailleurs, l'équipe est encouragée à maintenir une page permettant de télécharger ses publications ainsi que ses logiciels. >>

- Comme indiqué dans la section production et qualité scientifique l'équipe a maintenu son activité de publication dans des revues répertoriées de bon niveau (53 publications sur la période dont 60% dans des journaux évalués Q1 dans SJR et 68% dans des journaux évalués A1 ou A2 dans QUALIS). Cependant, le nombre de communications dans des conférences est de 21,2 communications en moyenne par an contre 34,5 communications en moyenne par an sur la précédente période. Cette baisse est en partie due au report de plusieurs conférences à cause de la crise sanitaire. L'équipe a malgré tout publié dans des conférences sélectives (CoDIT, ISCO, MAPSP, PMS, MIC ou encore ODYSSEUS).
- L'équipe ROOT a intensifié ses collaborations au sein du laboratoire avec 2 thèses soutenues et 2 thèses en cours sur des sujets transverses impliquant au moins un encadrant de chaque équipe. 3 publications en revue et 10 publications en conférence sont co-signées par au moins un membre de l'équipe ROOT et un membre de l'équipe BDTLN ou RFAI. Enfin, l'équipe a participé à un projet région conjointement avec l'équipe BDTLN (SMARTLOIRE) et un autre projet démarre avec cette même équipe (SQVALD). Cette possibilité est favorisée par la politique d'attribution des bourses de thèses institutionnelles du LIFAT, qui encourage fortement les thèses inter-équipes.
- Enfin, l'équipe a suivi la recommandation de maintenir une page permettant de télécharger ses publications en créant une collection HAL 'ROOT-TOURS' qui regroupe les manuscrits de toutes les thèses soutenues sur la période, et référence toutes les publications (revues, conférences et ouvrages) avec 68% des articles publiés en revues et 56% des articles de conférences disponibles au téléchargement.

<< Critère 2 : rayonnement et attractivité académiques : Afin d'améliorer encore sa visibilité et son attractivité, l'équipe est encouragée à organiser plus d'événements nationaux et internationaux sur place et à diversifier davantage ses collaborations internationales. >>

- Contrairement à la précédente période, l'équipe a organisé une manifestation internationale à Tours : « International workshop on Heuristics » du 18 au 20 juin 2018 (<https://mh2018.sciencesconf.org>). Cette conférence a réuni une quinzaine de présentations sur 3 jours (sur sélection d'articles de 6 ou 8 pages) et a donné lieu à un numéro spécial dans la revue internationale « Journal of Heuristics » sous la direction de trois membres de l'équipe. L'équipe a également participé à l'organisation d'une autre conférence internationale en 2016, VEROLOG à Nantes (<https://verolog2016.sciencesconf.org>, 142 présentations sur 3 jours). L'équipe a aussi participé à plusieurs manifestations nationales dont 3 organisés à Tours : journée du GT2L en 2016 (6 présentations), journée de l'ANR e-VRO en 2017 (4 présentations) et journées conjointes des GT GOTHa et Bermudes en 2017 (12 présentations sur 2 jours).
- Enfin, sur la période 2016-2021, l'équipe a essayé de diversifier ses collaborations internationales en accueillant plus d'une dizaine de scientifiques étrangers de renommée, en soutenant plusieurs membres de l'équipe dans des séjours scientifiques dans 6 laboratoires étrangers différents, et en signant 5 conventions de cotutelle de doctorats avec 4 établissements différents. L'ensemble des collaborations internationales a abouti à de nombreuses publications co-signés avec au moins un chercheur étranger (34 publications en revue et 54 publications en conférence).

<< Critère 3 : interactions avec l'environnement économique, social culturel et sanitaire : Il est primordial pour l'équipe de maintenir son excellence scientifique. Pour cela, il faut veiller à minimiser les projets industriels à faible valeur académique. >>

- L'équipe ROOT pense avoir maintenu son excellence scientifique, notamment via la publication des contributions scientifiques dans des journaux internationaux de haut niveau. Et, comme le montrera la description de la référence 1 du DAE 4, l'équipe a essayé de maintenir uniquement des projets industriels avec une valeur académique notable pour lesquels des contributions à la recherche sont possibles et valorisables par des publications scientifiques. Avant chaque dépôt de nouvelle CIFRE, l'équipe s'assure que la thématique de la thèse porte bien sur des sujets en phase avec les travaux de recherche de l'équipe et que le contrat prévoit un temps de présence du doctorant au laboratoire de 50% minimum.

<< Critère 4 : organisation et vie de l'équipe (s'il y a lieu) : Il faut continuer à motiver les membres de l'équipe pour qu'ils participent aux travaux scientifiques collectifs. >>

- Ces dernières années, l'équipe a mis en place des réunions d'équipe une fois par mois et des brainstormings ou séminaires une à deux fois par mois. Lors des brainstormings où les doctorants sont présents, l'équipe aborde des sujets prospectifs de recherche laissant l'opportunité à chaque membre de participer à ces éventuels futurs projets de recherche. Ces brainstormings permettent notamment de constituer des projets de recherche collectifs qui sont ensuite déposés à l'ANR

ou à la région. Enfin, pour chaque thèse envisagée, l'équipe incite ses membres à les co-encadrer, pour avoir au moins deux encadrants par thèse, de préférence avec un MCF non HDR.

<< Critère 5 : implication dans la formation par la recherche : Pour minimiser le risque de l'abandon de thèses, il faut veiller à bien évaluer la motivation réelle des candidats. Les doctorants devraient aussi bénéficier d'un encouragement plus systématique pour participer à des écoles d'été. >>

- Malgré un processus de recrutement renforcé pour bien évaluer la motivation des candidats en thèse via un concours sur dossier et des entretiens, il y a eu trois abandons de thèse dans l'équipe, exactement comme dans la précédente période. Les doctorants sont maintenant fortement incités à participer à au moins une école d'été durant leur thèse. A chaque réunion d'équipe, les écoles d'été disponibles du domaine sont présentées afin d'informer les doctorants et encadrants des opportunités. Durant la période 2016-21, l'équipe a financé :
  - 1 doctorant à l'école d'été « Approches Quantiques pour une nouvelle recherche opérationnelle !? » organisée à Montpellier par le GT2L du GDR RO en 2021.
  - 1 doctorant à l'école d'été « Ordonnement, Planification et Applications » organisée à Paris par le GDR RO en 2021.
  - 1 doctorant à l'école de printemps « Ordonnement et Approximation » organisée à Metz par le GT Polyèdres et Optimisation Combinatoire du GDR RO en 2019.
  - 1 doctorant à l'école d'été internationale « Operational Research for Value-based Health Care » organisée à Lisbonne par EURO (association européenne en RO) en 2019.
  - 4 doctorants à l'école d'été internationale « Integrated Operational Problems » organisée à Troyes par le GDR RO en 2018.
  - 4 doctorants à l'école d'été internationale « Interactions between Operations Research and Machine Learning » organisée à Fréjus par l'école des Ponts ParisTech en 2018.
  - 2 doctorants à l'école d'été internationale « Hybridization of Constraint Programming and Operational Research: from practical opportunities to concrete implementation of solutions » organisée à Porquerolles par le GDR RO en 2017.
  - 4 doctorants à l'école d'été « Ordonnement : fondements théoriques et outils de résolution » organisée à Angers par le GDR RO en 2017.
  - 1 doctorant à l'école d'été internationale « Healthcare Operational Research » organisée à Enschede par l'université de Twente en 2017.
  - 3 doctorants à l'école d'été « Les nouveaux enjeux de la logistique et du transport » organisée à Gardanne par le GT2L du GDR RO et les porteurs des ANRs ATHENA, EVERS et e-VRO en 2017.

A noter également que 2 doctorants ont profité d'une mobilité scientifique de 2 semaines et un mois dans des laboratoires français (l'institut de Mathématiques de Bordeaux dans Equipe-projet inria ReAlOpt et au LAAS dans l'équipe ROC)

<< Critère 6 : perspectives et stratégie scientifique à cinq ans : Les compétences de l'équipe sont complémentaires de celles des deux autres équipes du LIFAT et certaines thématiques du LIFO (Orléans). Les actions communes sont un excellent moyen pour renforcer la cohésion de l'unité et faire émerger des thématiques nouvelles. Elles doivent donc être poursuivies.

Étant donnée la taille de l'équipe (elle est aujourd'hui numériquement l'équipe la plus petite du laboratoire), de nouveaux recrutements, notamment des maîtres de conférences, seront nécessaires pour rassembler une force suffisante autour de chaque thématique. >>

- Comme évoqué précédemment, plusieurs travaux de recherche ont été menés en collaboration avec les autres équipes du laboratoire. L'équipe souhaite maintenir ce type de collaboration, d'ailleurs, l'une des perspectives de recherche envisagées dans les prochaines années est l'association des outils de la RO avec les outils de l'apprentissage automatique. Quelques travaux ont aussi été initiés avec des collègues du LIFO (Orléans). Ces travaux ont porté sur les méthodes exponentielles pour la résolution de problèmes d'ordonnement (1 article en revue et 1 communication en conférence) et sur l'éthique de la numérisation et de l'organisation de la santé (1 communication en conférence et le projet région ETHNOS).
- Enfin, depuis la précédente période, l'équipe ROOT est passée de 12 à 13 membres permanents grâce au recrutement d'un nouveau maître de conférences. Elle reste toutefois l'équipe la plus petite du laboratoire.

### Projets de recherche pour les années à venir

Au vu des travaux de recherche en cours, des expertises acquises par les membres de l'équipe et des directions de recherche prises, l'équipe a réfléchi à de futurs axes de recherche prometteurs et à de grandes thématiques de recherche auxquelles elle souhaite contribuer. Tout d'abord, ses travaux continueront à s'inscrire dans le domaine de la Recherche Opérationnelle, et plus précisément dans le développement de nouveaux algorithmes en optimisation combinatoire. Les contributions scientifiques sur lesquelles l'équipe projette de travailler peuvent être regroupées en deux axes non exclusifs :

- **La modélisation et la résolution de problèmes** : l'objectif principal de cet axe est d'étudier de nouvelles problématiques, la plupart issues de défis sociétaux décrits plus loin, et de proposer des méthodes de résolutions innovantes, pertinentes et performantes. Parmi les grands thèmes de recherche que l'équipe souhaite développer ces prochaines années, nous trouvons :
  - Les modèles théoriques pour l'ordonnancement ou le transport, car les nouveaux contextes identifiés apportent de nouvelles contraintes de fonctionnement (techniques, réglementaires) et de nouveaux objectifs à optimiser. De plus, l'ordonnancement et le transport sont les deux domaines de recherche fondateurs de l'équipe. Malgré la richesse de la littérature dans ces deux domaines, de nombreuses contributions sont encore possibles, notamment pour palier au décalage entre les modèles dit académiques et ceux rencontrés sur le terrain et en lien avec les défis mentionnés ci-après.
  - Les modèles intégrés correspondent à des problèmes qui sont composés de plusieurs sous-problèmes combinatoires difficiles, ce qui rend leur résolution encore plus difficile. Pourtant, ces modèles intégrés permettent de considérer des problèmes plus vastes pour une meilleure prise de décision. Ces dernières années, l'équipe a travaillé sur ce type de modèles, en particulier sur le couplage de problèmes de production (ordonnancement) et de livraison (tournée de véhicule). Elle souhaite poursuivre ses recherches dans cette direction encore très prometteuse.
- **La conception et hybridation de techniques d'optimisation** : l'objectif principal est de concevoir et analyser de nouvelles méthodes en développant de nouvelles techniques d'optimisation ou en hybridant des techniques existantes. Parmi cet axe, l'équipe souhaite continuer à explorer les champs suivants :
  - L'association d'outils de la RO et de l'apprentissage automatique pour : (1) améliorer l'efficacité de la résolution de problèmes combinatoires, ou (2) améliorer des méthodes d'apprentissage par une meilleure optimisation de certaines de ses étapes, ou encore (3) accroître la robustesse des solutions dans des environnements dynamiques ou aléatoires en intégrant des phases de prédiction de données dans les méthodes de résolution. D'un point de vue général, les systèmes informatiques récupérant de plus en plus de données de natures variées, et dont il faut tenir compte, cette association de techniques guidera le système à mieux décider en termes de gestion des ressources et planification de tout type d'opérations (production, transport, maintenance et de transfert). Ces perspectives viennent aussi renforcer les collaborations avec les autres équipes de recherche du LIFAT : avec l'équipe BDTLN lorsqu'il s'agit de travailler sur les données ou avec l'équipe RFAI lorsque l'on s'intéresse aux modèles et aux connaissances pour l'apprentissage.
  - Les méthodes basées sur l'algorithmique modérément exponentiel visent à résoudre de façon exacte des problèmes d'optimisation difficiles dans un temps de calcul au pire cas qui soit modérément exponentiel. Lorsque l'on s'intéresse en RO à la résolution exacte de problèmes difficiles, les algorithmes mis au point sont évalués expérimentalement sur la base de leur comportement temporel moyen : leur complexité temporelle au pire cas est donc très souvent du même ordre qu'un algorithme de force brute. Mettre au point un algorithme exact modérément exponentiel, c'est donc proposer un algorithme exact dont la complexité temporelle au pire soit la plus réduite possible, et donc bien meilleure que celle de la force brute. Il y a tout un champ de nouvelles méthodes à explorer pour les problèmes d'ordonnancement et/ou transport. Ces recherches sont en premier théoriques et les résultats obtenus fournissent des informations quantitatives sur la difficulté d'un problème d'ordonnancement. Nous avons déjà beaucoup exploré cette voie et obtenus de nombreux nouveaux résultats. Néanmoins, cet axe de recherche ne s'arrête pas à des études théoriques : créer des algorithmes exponentiels c'est avancer dans la compréhension de mécanismes permettant de réduire la combinatoire d'un problème. Cela permet donc d'en dériver des techniques efficaces en pratique. C'est ce que nous avons fait à partir du mécanisme de la Mémorisation qui nous a conduit à définir des techniques de mémorisation dans les méthodes exactes arborescentes avec d'excellents résultats en pratique sur des problèmes d'ordonnancement (cf. élément 2 du portfolio ROOT). Nous comptons continuer dans cette voie notamment en partant de la technique théorique d'Inclusion-Exclusion. Ce champ de recherche fait également l'objet de collaborations scientifiques naissantes avec des mathématiciens de l'Institut Denis Poisson.
  - Les méthodes de type mathéuristique : l'association de méthodes dites « mathématiques », comme la programmation linéaire, la programmation par contraintes, ou d'autres issues des deux premières (génération de colonnes, décomposition de Benders, etc.), et des méthodes dites « heuristiques/métaheuristiques », restent encore une perspective de recherche d'actualité en RO. Il existe en effet de nombreuses possibilités d'hybridation, qui dépendent de la nature des méthodes choisies et de la manière dont elles sont imbriquées les unes aux autres. L'équipe ayant déjà eu plusieurs succès dans le développement de ce type de méthode, plusieurs membres souhaitent approfondir ce domaine.

L'équipe **se concentrera sur les trois grands défis suivants** : la santé, l'industrie du futur et les mobilités de biens et/ou de personnes. Ces défis s'inscrivent dans les grands défis de la société mis en avant par le CNRS dans son dernier contrat d'objectifs et de performance 2019-2023, signé avec l'Etat (notamment l'intelligence artificielle, la santé et l'environnement, les territoires du futur et la transition énergétique).

- Dans le milieu de la **santé**, l'« hôpital du futur » fait apparaître de nouveaux challenges de réorganisation des services de santé, de numérisation des informations, de personnalisation de soins ou encore d'efficacité à budget contraint. De nouveaux outils d'aide à la décision pour des problèmes de dimensionnement, de coordination, de planification, ou plus généralement d'organisation devront être développés, ce qui nécessitera l'étude, la modélisation et la résolution de nouveaux problèmes. De par son expérience dans le domaine de l'organisation des systèmes de santé, l'équipe a tous les atouts pour aborder les problématiques qui apparaîtront dans les futurs systèmes de santé.
- L'**industrie du futur**, autrement dit la modernisation de l'outil de production et l'accompagnement des entreprises industrielles dans leur transformation numérique, est un enjeu de compétitivité majeur. Cette transition nécessite pour les entreprises un accompagnement, pour moderniser leur système de production, pour accompagner le personnel dans le changement de leur mode de travail, et pour traiter toutes les données récoltées et créer les bons outils d'aide à la gestion. Un des objectifs de l'industrie du futur est de mettre en place une gestion des ressources qui soit plus flexible, plus efficace, grâce à des remontées d'informations et à des systèmes dits "experts" d'aide à la décision. Ces nouveaux environnements de production et les nouvelles informations qui seront disponibles permettront d'aborder les problèmes classiques d'optimisation de la production sous un nouvel angle, avec notamment en perspective une collaboration très nette entre les techniques d'optimisation et les techniques d'apprentissage. L'hybridation de ces deux techniques permettent de développer des solutions innovantes rendant possible au système de percevoir et comprendre son environnement, de mieux décider en fonction de l'état de sa santé, du contexte et des usages. Les collaborations que l'équipe a amorcées avec les collègues d'autres équipes du LIFAT et son expertise des systèmes de production, la place dans une situation très favorable pour aborder ces problèmes.
- Enfin, le dernier défi concerne les **mobilités de biens et de personnes** dans un contexte important de développement durable. Il s'agit de repenser les modes et organisations de transports, optimiser les aspects écologiques et promouvoir et améliorer les mobilités douces. Les outils de la Recherche Opérationnelle peuvent être un atout considérable pour répondre à ces challenges de mobilités, à condition qu'ils soient bien adaptés et efficaces, ce qui ouvre de nombreuses perspectives de recherche.

## Innovation algorithmique en Optimisation Combinatoire

Modélisation et résolution de problèmes

Conception et hybridation de techniques d'optimisation

Santé

*Hôpital du futur*

Système de production et chaîne logistique

*Industrie du futur*

Mobilités des biens et des personnes

*Dév. durable*

**Référence 2. La production scientifique est proportionnée au potentiel de recherche de l'équipe et répartie entre ses personnels**

Afin de décrire la production scientifique vis-à-vis du potentiel des membres permanents et non permanents de l'équipe sur la période 2016-2021, nous présentons d'abord la politique et la vie de l'équipe puis l'évolution et le potentiel de recherche de l'équipe. Nous terminons par la formation et la participation des doctorants et post-doctorants à la production scientifique de l'équipe.

### Politique et vie de l'équipe

L'équipe mène une politique qui favorise les synergies entre ses membres aux compétences complémentaires. Chaque membre de l'équipe peut développer un sujet de recherche en RO de son choix et collabore librement avec qui il veut, en fonction de ses compétences et motivations. Par ailleurs, l'équipe instaure des réunions mensuelles qui permettent de partager les informations sur la vie de l'équipe et du laboratoire. Ces réunions sont également un lieu où sont prises les grandes décisions relatives à la politique de l'équipe (à propos du budget, participations aux conférences, sujets de thèse, etc.). D'autre part, une à deux fois par mois, l'équipe organise des séminaires de recherche ou des brainstormings. Les séminaires traitent généralement d'un sujet de recherche précis et sont réalisés par des enseignants-chercheurs de l'équipe, par des doctorants ou par des chercheurs extérieurs invités. Les brainstormings ont un objectif différent des séminaires, ils visent à aborder des sujets de recherche très ouverts ou réfléchir, voire définir, des potentiels futurs projets de recherche. Ces séminaires et brainstorming sont des événements importants pour l'échange d'information, les conseils et la coordination de la recherche dans l'équipe. Enfin, chaque été, une journée « équipe ROOT » est organisée. Cette journée se décompose en trois parties. La journée commence par un exposé d'un chercheur invité (par exemple, en 2021, Eric bourreau du LIRMM a présenté les liens entre la RO et l'Informatique Quantique). Tous les doctorants de première et deuxième année présentent ensuite une partie de leurs travaux de thèses. Enfin la journée se termine par une ou plusieurs activités conviviales de « team-building ».

### Évolution et potentiel de recherche de l'équipe

À la fin du précédent contrat, l'équipe était composée de 12 permanents (5 Pr, 7 MCF dont 2 HDR). Durant la période, il y a eu un départ à la retraite en 2016 et un départ en mutation pour un poste de Professeur à HEC Montréal en 2018. Ces départs ont eu pour conséquence le recrutement de deux nouveaux MCF en 2017 et en 2020. En parallèle, un nouveau MCF a été recruté en 2018. Ces changements n'ont pas eu un effet positif sur la production scientifique puisque les recrutements n'ont pas pu se faire instantanément (délai de 1 et 2 ans) et l'arrivée d'un nouveau jeune MCF implique un temps d'adaptation et d'investissement pédagogique important lors de ses premières années. L'équipe est maintenant constituée de 13 permanents (5 Pr, 8 MCF dont 1 HDR). Elle est la plus petite équipe du laboratoire en nombre de permanents alors qu'en 2011 elle était la plus grande. En résumé, l'équipe ROOT a produit au total 53 articles publiés en revues, 131 communications en conférences, 5 ouvrages ou chapitres d'ouvrage, et a encadré ou encadre 29 thèses de doctorat sur la période 2016-21. A cette production s'ajoutent toutes les autres activités scientifiques décrites précédemment (portage d'une ANR, portage de 2 ANR JCJC, portage d'un projet Région, implication dans deux autres projets région, nombreuses relectures d'articles, nombreux jurys de thèse et HDR, etc.). Les publications scientifiques de l'équipe ne sont pas uniformément réparties sur tous les membres de l'équipe. Deux membres de l'équipe ont arrêté leurs activités de recherche depuis plusieurs années et n'ont rien produit sur la période 2016-2021. L'équipe a essayé à plusieurs reprises de les impliquer dans des projets de recherche ou de leur donner des opportunités d'encadrement d'étudiants en M2 ou même en doctorat, mais les tentatives n'ont pas abouti. D'autre part, des membres de l'équipe ont assumé ou continuent d'assumer des charges collectives lourdes. Parmi ces charges, nous pouvons mentionner sur la période : la direction de Polytech Tours (6 ans), la Vice-présidence des partenariats et de l'innovations de l'université de Tours (2 ans), la direction du département Formation par Alternance de Polytech Tours (6 ans), la direction adjointe chargée des relations entreprises et des partenariats de Polytech Tours (4+2 ans), la direction de laboratoire LIFAT (2 ans) et enfin des directions d'étude de la spécialité informatique industrielle par apprentissage du département Formations par Alternance (6 ans), du département informatique (4 ans) et du département du premier cycle (parcours des écoles d'ingénieurs Polytech, 2 ans). Malgré ces charges, les personnes concernées maintiennent une activité de recherche conséquente et sont ou ont été impliquées dans d'autres instances nationales et internationales.

### Formation et participation des doctorants

Dans la vie de l'équipe, les doctorants occupent une place centrale car ils sont fortement investis dans leurs activités de recherche. Ils participent aussi librement aux décisions prises en réunion d'équipe. Leur formation implique un encadrement lourd mais riche de la part des permanents de l'équipe : elle pousse ses doctorants à avoir, en fin de thèse, au moins une publication acceptée dans une revue internationale de qualité, des communications dans des conférences internationales et nationales, et d'avoir participé à au moins une école d'été. Ils sont également encouragés à effectuer des séjours scientifiques en France ou à l'étranger pris en charge sur les ressources de l'équipe. Dans le but de favoriser l'émergence de thèses de qualité, l'équipe a mis en place pour chaque thèse, selon la politique du laboratoire, un comité de suivi de thèse qui se réunit une fois par an et qui permet au doctorant d'avoir un retour sur l'avancement de son travail. La venue de chercheurs invités extérieurs à l'équipe est également l'occasion pour eux de présenter leur travail et d'avoir un retour constructif. Au cours de la période, 16 doctorants ont soutenu leur thèse (durée moyenne : 43 mois) et 3 ont abandonné avant la soutenance (taux de réussite : 84%). Au cours de la précédente période 15 doctorants avaient soutenu leur thèse (taux de réussite : 83%, durée moyenne : 41 mois). Le taux de réussite est donc similaire à la précédente période mais on note une légère augmentation de la durée des thèses probablement en raison de la crise sanitaire. Actuellement, 10 thèses sont en cours avec 4 soutenances prévues pour 2022. Parmi les publications scientifiques de l'équipe, 58% (31) des articles en revue de l'équipe et 72% (94) des communications en conférence de l'équipe sont co-signées par au moins un

doctorant, dans plus de 90% de ces publications, le doctorant est le premier auteur. Par ailleurs, une volonté de l'équipe est de favoriser l'encadrement de doctorants par un permanent de l'équipe HDR et un permanent non HDR : il est important que les jeunes collègues MCF puissent s'investir dans l'encadrement doctoral. Seul un post-doctorant a été recruté durant la période pendant 9 mois. Ses travaux ont donné lieu à 2 communications en conférence et un article en revue qui sera publié en 2022. Enfin, l'équipe accueille régulièrement des stagiaires dans son équipe. Sur la période 2016-2021, on peut dénombrer 3 stages de M2 de 5 mois en moyenne, 7 stages M1 de 2,5 mois en moyenne et 1 stage L3 de 3 mois.

### Synthèse de l'autoévaluation

Une vision synthétique de l'autoévaluation de l'équipe ROOT est présentée sous la forme d'une matrice SWOT.

POSITIF	NEGATIF
<p><b><u>Forces :</u></b></p> <ul style="list-style-type: none"> <li>• Qualité de la publication scientifique</li> <li>• Formation doctorale</li> <li>• Visibilité nationale et internationale</li> <li>• Implication dans l'animation de la communauté scientifique au niveau national</li> <li>• Lien avec le monde industriel</li> </ul>	<p><b><u>Faiblesses :</u></b></p> <ul style="list-style-type: none"> <li>• Production scientifique non uniforme au sein des membres de l'équipe</li> <li>• Abandons de thèse</li> <li>• Charges collectives lourdes pour environ la moitié des membres</li> </ul>
<p><b><u>Opportunités :</u></b></p> <ul style="list-style-type: none"> <li>• Collaboration inter-équipes</li> <li>• Réseau solide national et international</li> <li>• Défis sociétaux</li> </ul>	<p><b><u>Menaces :</u></b></p> <ul style="list-style-type: none"> <li>• Manque de projets financés (ANR, Région, etc.)</li> </ul>

Parmi ses **forces**, l'équipe promeut l'excellence dans ses travaux de recherche en mettant en avant ses publications dans des supports visibles, sélectifs et reconnus par la communauté scientifique. Ces travaux de recherche s'appuient sur un nombre important de doctorants que l'équipe forme et incite à publier. Comme autre force, l'équipe possède une très bonne visibilité nationale et internationale au vu des nombreuses sollicitations d'expertises de travaux scientifiques et des collaborations effectives. D'ailleurs, elle est très investie dans l'animation de sa communauté au niveau national, comme le GDR RO ou la ROADEF. Enfin, l'équipe s'intéresse aussi à des applications au monde industriel au travers de CIFRE (STMicronics, SANOFI, SNCF, CYRES et Géovélo), ou de collaborations moins formelles (par exemple le CHRU de Tours et Eticsys via le projet Région Districhimio). Cependant, on peut noter quelques **faiblesses** qui devront être rectifiées. Tout d'abord, on peut noter que la production scientifique n'est pas uniforme au sein des membres de l'équipe, avec notamment deux personnes qui ont "décroché" de la recherche. Il faudra donc essayer de motiver voire inciter ces collègues dans des projets de recherche qu'ils les intéresseraient. Il faudra également veiller à mieux équilibrer la production scientifique en continuant de bien accompagner certains membres, dont les jeunes recrutés, et de leur donner plus d'opportunités (par exemple en les incluant dans des encadrements de thèses ou encore dans des projets de recherche prospectifs qui donneront lieu à des dépôts de projets pour financement). Une autre faiblesse à corriger sera les abandons de thèse en continuant à perfectionner les phases d'évaluation de candidatures et de suivi de thèse. Enfin, un dernier point de vigilance et difficile à rectifier, est l'engagement des membres dans des charges collectives lourdes qui pourrait avoir des conséquences néfastes sur la production scientifique de l'équipe et dans le portage de projets de recherche à long terme. De nombreuses **opportunités** se présentent à l'équipe. Pour commencer, les collaborations inter-équipes ont commencé à porter leurs fruits et ont ouvert de nouvelles perspectives de projets de recherche innovants, la plus prometteuse étant l'association d'outils de la RO et de l'apprentissage automatique. Le réseau national et international de l'équipe peut également créer des opportunités de nouveaux projets de recherche collaboratifs, de nouvelles thèses en cotutelle, ou encore de séjours scientifiques. Enfin, les défis sociétaux exprimés dans les perspectives de recherche de l'équipe sont d'actualité et peuvent déclencher des financements de projet de la part d'agences de la recherche. De plus, ces défis peuvent donner lieu aux développements d'outils d'aide à la décision qui peuvent vraiment profiter au monde socio-économique. Enfin, les **menaces** actuelles portent sur le financement des projets dont les plus importants se sont terminés pendant la période 2016-2021. Il faudra donc relancer cette dynamique dans les prochaines années. Chaque année, l'équipe est impliquée dans le montage et le dépôt de projets ANR ou Région (3 ou 4 en moyenne par an). Actuellement, 2 projets ANR impliquant fortement l'équipe, sont en phase 2 de l'AAPG 2022 de l'ANR (dont 1 en tant que porteurs).

## 4. Domaine d'évaluation 4 : Impacts sociétaux et économiques

### 4.1. Au niveau du laboratoire

Les interactions avec les acteurs socio-économiques sont très nombreuses. Plusieurs collaborations avec des partenaires industriels ont donné lieu à des réalisations logicielles ou à des dépôts de projets.

#### 4.1.1. *Partenariats avec des acteurs socio-économiques, dépôt de marques et brevets*

Plus de **40 logiciels et 25 bases de données ou corpus** ont été développés (cf. document de production pour une liste exhaustive) dans le cadre de travaux de recherche menés au LIFAT en partenariat avec différents acteurs socio-économiques et/ou académiques. Ces derniers sont détaillés par chacune des équipes ayant contribué à leur développement.

De même, le LIFAT a participé à plusieurs dépôts de **déclaration d'invention, de dépôt de marques et à un dépôt de brevet** également décrit dans les sections Equipes.

Les partenariats socio-économiques du LIFAT sont également détaillés par chacune des équipes. Ce qui est notable au niveau global, c'est le très fort nombre de sollicitations d'industriels ou de start-ups qui ont eu lieu durant ce contrat notamment suite au boom de l'Intelligence Artificielle. Une vigilance particulière a donc dû être portée pour rejeter les sollicitations ne comportant pas une part de recherche scientifique suffisante ou hors des domaines d'étude du LIFAT. En tout, **13 conventions CIFRE ont été signées dans la période** : Kalidea, SAP, PetitPicot, Imascaop, Madmix (COF), ST-micro, SNCF, Sanofi, Nereide, Devoteam, Neolink, Nextinno, Adwanted (6 sont encore en cours).

Des collaborations avec les partenaires industriels se font également par l'intermédiaire de projets financés par un organisme ou directement par l'entreprise. Tous les projets visant à valoriser les activités du LIFAT passent par le Service Partenariat Valorisation Contrats de l'Université de Tours et parfois par [CVALO](#).

Au-delà de ces prestations et contrats de collaboration, il faut noter la volonté et la capacité de membres du laboratoire à tenter de **valoriser leurs recherches sous forme de start-up ou spin-off avec** le soutien du SPIV de l'université de Tours ou de l'organisme régional [C-Valo](#) :

1. [Sila-3D](#) : Plateforme de segmentation d'images biomédicales 3D résultant de plusieurs projets pluridisciplinaires régionaux (cf Portfolio)
2. [Todd.TV](#) : projet de start-up s'appuyant sur les travaux en analyse de vidéos temps réel avec applications pour l'analyse de flux TV (cf. Portfolio)
3. [CETU ILIAD3](#) : créé en 2017, le CETU autofinance actuellement 3 emplois temps plein en générant un CA supérieur à 150K€/an (cf portfolio)

#### 4.1.2. *Participation à la diffusion de la culture scientifique*

Le développement croissant de la société numérique et sa mise en réseau ont suscité l'émergence de questionnements concernant avant tout des questions éthiques telles que le respect de la vie privée sur les réseaux sociaux, l'usage et la propriété des données personnelles, le droit à l'oubli numérique, et la question de l'émergence d'une société du contrôle permise par ces techniques. Ces questions parfois éthiques concernent tout le monde et sont souvent sources de débat actuellement. Les membres du laboratoire interviennent ainsi régulièrement auprès du grand public ou des acteurs socio-économiques (cf. document de production pour une liste exhaustive) par exemple au travers de :

- [Centre-Sciences](#) est un organisme régional qui a pour mission de faire partager les sciences et les techniques à tous les habitants de la région Centre. Centre-Sciences est également un centre de ressources pour tous ses partenaires. Chaque année, Centre-Sciences sollicite les membres du LIFAT pour participer à des manifestations tournées vers le grand public :
  - La Fête de la Science, où les membres du LIFAT sont des partenaires habituels, avec des présentations autour de jeux (de recherche opérationnelle), de démos de logiciels d'analyse d'images, de mini-robots, etc.
  - Les « Mardis de la Science » permettent aux chercheurs de Tours de vulgariser leurs travaux
- La « nuit de l'info » se déroule chaque premier jeudi de décembre depuis 2013. C'est une compétition nationale et annuelle qui réunit étudiants, enseignants et entreprises. Durant cette nuit, des entreprises lancent des défis (développements d'applications informatiques innovantes) aux équipes participantes, et proposent des prix pour les équipes ayant le mieux réussi. Les participants disposent d'une nuit pour proposer, implémenter et packager des applications informatiques. Le tout se fait dans une ambiance généralement très conviviale. Sur le site de Tours, jusqu'à 150 étudiants de Polytech Tours et de l'IUT GEII de Tours, extrêmement motivés, ont participé à la nuit de l'info. Des étudiants se sont aussi mobilisés sur le site de Blois.
- La publication d'articles vulgarisés ou d'interviews dans les médias ([un exemple](#)).
- Des doctorants du LIFAT participent également à « Ma thèse en 180 secondes »

Plus ponctuellement, on note durant le contrat :

- Conférence Débat « [Cosmetic Expert : Intelligence Artificielle & Industrie du Futur : Et si on collaborait...](#) » Webinaire Cosmetic valley - décembre 2020
- Table ronde « [Numérisation et mémorisation du patrimoine](#) », à Tours, Semaine de la mémoire - B2V, septembre 2021.
- Conférence « Deep Learning & Graph Representation for Image Analysis ». 10èmes Journées Scientifiques et Techniques du Réseau des Microscopistes INRAE, sept 2021 – Tours
- Diffusion de la culture scientifique par particip de Blois ation, plusieurs fois; à des interventions en collèges comme [parrain scientifique](#) dans le cadre d'actions de science participative initiées par « la main à la pâte » ou par la « Maison de la sciences » (collège Corneille à Tours, collège de Château Renault dans le cadre du projet [Neuro2co](#),
- Le LIFAT a accueilli une artiste, Laurence Dreano, en son sein pendant plusieurs mois en 2019 avec l'aide du CETU ILIAD3 dans le cadre de l'action "Résidence d'artiste" organisé par l'université de Tours. L'objectif était d'introduire, de combiner [numérique et virtuel au sein de sculptures](#).
- « Apéros info » organisés par BDTLN en soirées sur Blois de 2016 à 2020, ouverts aux étudiants et public externe (stoppés momentanément par COVID) : 3 à 4 épisodes par an

Les interventions de diffusion de la culture scientifique se manifestent bien évidemment aussi au travers d'enseignements en intelligence artificielle réalisés à l'Université de Tours : formation en médecine, au CESR sur l'analyse de documents patrimoniaux, à l'INSA et dans le master BDMA de l'Université de Tours, dans le parcours IA de la spécialité informatique de Polytech Tours, formations dans le cadre de la SFR FED ([Voir supports disponibles sur le site LIFAT](#)).

Un investissement s'effectue également au travers de l'encadrement de projets étudiants de ces formations ou encore de vulgarisation auprès des jeunes (stages de 3ème à Polytech Tours, ...)

## 4.2. Equipe BDTLN

### Plan

- Référence 1 – interactions non académiques
- Référence 2 – produits ouverts au monde socio-économique
- Référence 3 – grand public
- Bilan des forces et des faiblesses - SWOT

### Référence 1. L'équipe se distingue par la qualité de ses interactions non-académiques.

#### Interaction avec l'environnement social économique, le monde de la santé et autre

Les projets de l'équipe comportent pour 44% d'entre eux une composante tournée vers la société, souvent en collaboration avec des équipes en SHS et en santé ou handicap. Le tableau ci-après reprend parmi les projets déjà montrés (C=Coordinateur, P=participant) ceux qui ont cette orientation sociétale, que ce soit pour la bio-info, la santé ou le handicap, les dépenses énergétiques, le journalisme, le patrimoine, etc. De plus, bon nombre de ces projets comportent un volet de *science participative*, comme Mob'iKids, les actions prospectives dans le GDR MAGIS, ou le projet international PHC Utique. C'est également le cas dans les projets ANR de l'axe TAL, non rappelés dans le tableau suivant mais qui intègrent généralement un volet annotation, ce qui donne lieu à des travaux dirigés par Jean-Yves Antoine sur le sens des mesures d'accords inter-annotateurs, qui parfois fait aussi appel à la science participative.

Projets avec composante sociétale				
Région CVL (ARVIVA)	P	2014	2016	Patrimoine : Renaissance en Val de Loire
Région CVL (DOPAn)	C	2014	2017	Vulnérabilité énergétique et territoires
CNRS GDR MAGIS (mobilités et trajectoires)	C	2013	2025	Communauté pluridisciplinaire (informatique, géomatique, géographie, capteurs, psychologie, écologie)
ANR (MOBI'KIDS)	P	2016	2022	Suivi et prédiction d'activités de familles en ville
Fondation (PREDICT4ALL)	P	2017	2020	Assistance à personne avec handicap
Biosystémique	P	2017	2018	bio-informatique
Région CVL (SmartLoire )	C	2017	2021	Tourisme en Val de Loire
H2020 (ARIADNEplus)	P	2018	2022	Patrimoine : infrastructure numérique pour l'archéologie
ANR (Abliss)	P	2018	2022	bio-informatique
ANR (SESAMES)	P	2018	2023	Patrimoine : archéologie et architecture patrimoniale
CNRS MASTODONS (QCM-BioChem)	P	2018	2019	bio-informatique
Région CVL (ArchSW)	C	2018	2019	Patrimoine : terminologie et interopérabilité
CNRS GDR MADICS (MADONA)	C	2019	2020	Journalisme, fact-checking
CNRS GDR MAGIS (VGI/Crowdsourcing)	C	2019	2025	Information géographique volontaire et crowdsourcing
ANRT CIFRE (NEOLINK)	C	2020	2023	Parcours de retour à l'emploi
PHC Utique (TransLvoGelInfo)	C	2020	2020	Modes de déplacements citoyens
H2020 (4CH)	P	2021	2024	Patrimoine : centre de compétences numérique
Région CVL (SQVALD)	C	2021	2024	Médecine : accompagnement des patients hors soin hospitalier

## Relations avec des entreprises

Des entreprises participent aux projets de l'équipe BdTIn, par exemple dans Mobi'Kids ou 4CH, ce qui amène parfois à nouer des relations au-delà du projet, par exemple avec la TPE Sparna active dans le domaine du Web sémantique. De plus, les 8 CIFRE menées sur 2016-2021 contribuent à établir plus de contacts. L'un des projets forts du tableau précédent est une CIFRE avec l'entreprise Neolink, « pionnière dans le Web social et résolument attachée aux valeurs de solidarités » selon sa propre présentation. Cette thèse est dédiée à un *système de recommandation interactif* de parcours professionnels et de formation. D'autres contrats CIFRE, montrés dans le tableau ci-après, sont avec des PME (la SCOP Nereides, Petit-Pecot filiale du groupe Gruau, Nextino filiale de Atempo-Wooxo), ou avec de grandes entreprises comme Adwanted, Kalidea du groupe UP (la plus grande SCOP de France) et SAP. D'une façon générale, l'équipe est attentive à ne s'engager dans une thèse CIFRE que lorsque le sujet relève parfaitement de ses thématiques de recherche et peut amener à lever des verrous scientifiques. Dans l'axe « Exploration et analyse interactive de données » la première avec Kalidea, a porté sur le *clustering et l'analyse interactifs de données* clients, capables de s'adapter à l'évolution des caractéristiques des données dans le temps. La thématique de l'élaboration de *recommandations* est au cœur des thèses menées avec SAP (requêtes décisionnelles), Kalidea-Up (contenus éditoriaux), Adwanted (plannings publicitaires) et Neolink (déjà évoquée). Avec Petit-Pecot la thèse a porté sur la *fouille de trajectoires sémantiques*. Dans l'axe « Intelligence des données et services », avec Nereide l'objectif était de réaliser *l'interopérabilité sémantique* de descriptions d'API RESTful à travers une ontologie des interactions HTTP permettant la spécification d'APIs avec SHACL. Enfin dans l'axe « Traitement des langues et interactions » le travail avec NextIno porte sur la *résolution de corréférence indirecte* pour la pseudonymisation de documents.

Projets avec entreprises : 361 k€ (316 k€ obtenus sur 2016-2021)				
CIFRE - KALIDEA	C	2015	2018	45 000 €
CIFRE - SAP	C	2016	2019	35 100 €
CIFRE PECOT Smart Ambulance	C	2017	2021	30 000 €
CIFRE NEREIDES	C	2019	2022	10 000 €
CIFRE KALIDEA-UP	C	2019	2022	78 000 €
CIFRE ADWANTED	C	2020	2023	60 000 €
CIFRE NEOLINK	C	2020	2023	51 720 €
CIFRE NEXTINO	C	2021	2024	52 050 €

### Référence 2 : L'équipe développe des produits à destination du monde socio-économique.

Les outils développés dans le cadre des thèses CIFRE restent en général aux entreprises partenaires. Même si l'équipe ne dispose d'aucun support d'ingénieur, des efforts sont faits (voir annexe X1) pour conserver les outils et autres ressources développés dans les projets afin de pouvoir les conseiller ou les réadapter à d'autres partenaires. Il est important de noter que la collaboration menée avec l'entreprise SAP a débouché non seulement sur la soutenance de la thèse de Krista Drushku et ses excellentes publications (2 articles dans le journal Information Systems (Q1) et 2 en grandes conférences internationales), mais aussi sur le *dépôt d'un brevet* (US Patent App. 15/964,814, 2019) dans lequel apparaissent les noms de ses encadrant.e.s membres de BdTIn (Nicolas Labroche, Patrick Marcel et Verónika Peralta).

Un recensement des prototypes, sites web, benchmarks, corpus et autres ressources déposés dans un Git, un portail, une plateforme ou tout autre registre public a été mené à l'occasion de cette auto-évaluation, dont le résultat est fourni en annexe 1. Leur quantité témoigne de la richesse des développements réalisés comme de la nécessité de mieux s'organiser, malgré le manque total de moyens humains pour cela, afin d'améliorer leur adéquation aux principes FAIR énoncés dans le cadre de la science ouverte, c'est-à-dire les rendre trouvables, effectivement accessibles, suffisamment documentés par des métadonnées favorisant l'interopérabilité, pour faciliter leur réutilisabilité. La migration des nombreux Git personnels vers celui du LIFAT est en cours et une *réflexion sur une homogénéisation dans le choix d'une plateforme de dépôt* est initiée, avec en particulier l'analyse des avantages qu'offrirait la plateforme de dépôt d'archives ouvertes Zenodo<sup>9</sup>.

### Référence 3. L'équipe partage ses connaissances avec le grand public

Sur le dernier contrat l'équipe a mis en place à Blois deux types d'actions ouvertes au grand public, les « apéros info » et un séminaire en juin. Les apéros info interviennent 3 fois par an en début de soirée (18:30 – 20:00) sous la forme de deux présentations grand public de travaux menés dans l'équipe suivies de discussions. Le public est composé essentiellement d'étudiants de licence et de master, plus des personnes amenées par nos partenariats avec des associations d'une part, comme Loir-et-cher Tech<sup>10</sup> (qui organise notamment des « startup weekends » auxquels nous poussons nos étudiants à participer) et d'autre part avec les lycées et collèges (des enseignant.e.s qui découvrent notre présence grâce aux actions en direction des collégiens ou lycéens). Nous invitons le même public, en veillant depuis l'an dernier à inclure nos collègues du LIFAT en poste à Blois à l'INSA, au séminaire de juin où sont présentés

<sup>9</sup> <https://zenodo.org/>

<sup>10</sup> <https://www.loirechertech.fr/page/327551-accueil>

des travaux de l'année qui n'ont pas déjà été exposés au début d'une réunion d'équipe. Les *réunions du groupe RITUEL* du RTR DIAMS, mentionnées dans le bilan de l'axe TAL, sont quant à elles l'occasion de recevoir des entreprises de la région intéressées par les sujets qui y sont abordés.

### Synthèse de l'autoévaluation

Le profil de recherche de l'équipe BdTIn montre clairement que nous nous investissons beaucoup dans notre cœur de métier constitué de l'activité de recherche, l'encadrement de recherches et l'administration liée à la recherche (qui augmente avec le nombre de projets). L'investissement dans le Domaine 4 est de ce fait moins important, ce n'est pas la motivation qui manque mais plutôt des moyens. Dans la limite de ses moyens, l'équipe inscrit ses recherches dans la société à travers (1) ses investissements importants dans de nombreuses collaborations avec d'autres disciplines dans ses projets régionaux, nationaux et internationaux, dont plusieurs intègrent aussi des volets de science participative, (2) ses implications avec des équipes de R&D d'entreprises via 8 thèses CIFRE, (3) ses engagements dans le mouvement de la science ouverte avec une attention aux principes FAIR et l'utilisation de plateformes de dépôts pour les outils et ressources développées et (4) des interactions régulières avec le tissu éducatif et socio-économique local.

Inscription des activités de recherche de l'équipe BdTIn dans la société	
Origine interne	
Forces	Faiblesses
Capacité à travailler avec d'autres disciplines	Manque de temps
Capacité à encadrer des thèses menées avec des entreprises	Manque de moyens : pas de chargé de communication, pas d'ingénieur pour valoriser et pérenniser les réalisations
Capacité à travailler sur des sujets sociétaux avec des partenaires ancrés dans l'environnement régional	
Origine externe	
Opportunités	Menaces
Demandes sociétales fortes	
Mouvement pour la science ouverte, principes FAIR pour les données et bientôt les outils aussi	
Association du LIFAT et plus particulièrement de l'équipe BdTIn à la <b>MSH Val de Loire</b> , qui structure beaucoup d'actions en direction de la société, comme les "boutiques des sciences"	

### 4.3. Equipe RFAI

#### Référence 1. L'équipe se distingue par la qualité de ses interactions non-académiques

Comme mentionné en présentation de l'équipe, nous travaillons en collaboration étroite avec des partenaires académiques de différents domaines. Cette ouverture inscrit également nos travaux dans la société avec des retombées socio-économiques potentiellement fortes. Ainsi, une très grande partie de nos projets (cf. Annexe B) intègre des applications et retombées dans nos domaines cibles que sont la santé et le handicap, les arts, humanités et patrimoine et l'environnement. Ces projets nous permettent également d'établir de nombreuses collaborations avec des partenaires non-académiques. Nous effectuons également des missions de valorisation, en collaboration avec le CETU ILIAD3 qui est ici une aide précieuse. Une partie de ces partenaires sont issus du milieu touristique et en lien avec nos activités sur le patrimoine (ville de Loches, château de Chambord...) ou liés à leurs activités en relation avec l'environnement (Automachine, Frasem, Chambres d'agriculture, Vinôpole, Agreen Tech Valley, Farmviz, Antea group, ATOS...). Parmi ces relations, certaines sont le fruit d'un travail collaboratif parfois ancien, initié notamment au travers de l'encadrement de projets (projets de fin d'étude) d'étudiants de Polytech Tours. D'autres sont nouvelles et issues de réseaux régionaux comme le RTR DIAMS ou le consortium qui s'est formé autour du montage de l'ARD Junon. Sur le contrat, l'équipe a également conclu 2 CIFRE avec deux start-ups :

- MADMIX DIGITAL, entreprise située à Paris et New York qui est un studio de création pour des contenus et interfaces digitales (notamment sur le web)
- IMASCAP : entreprise bretonne spécialisée la chirurgie de l'épaule assistée par ordinateur

Trois CIFRE sont également en montage sur 2022 dont une avec ATOS. Ce nombre de CIFRE peut paraître faible mais il s'explique par les raisons suivantes :

- De nombreuses sollicitations qui rentreraient dans nos thématiques s'apparentent plus à de la prestation ou de la valorisation. Nous avons réorienté plusieurs entreprises vers le CETU ILIAD3 plutôt que vers des collègues de l'équipe déjà trop sollicités
- Un certain nombre de sollicitations viennent de startups en construction qui sont parfois peu solides financièrement et/ou qui ont une vision un peu erronée de ce qu'est la recherche. Certaines expériences passées dans ce contexte nous ont

montré que l'on pouvait perdre beaucoup de temps dans ces interactions et se disperser, s'éloigner de nos objectifs scientifiques et expertises

- Nous avons fait le choix de concentrer notre participation en priorité sur des projets régionaux et nationaux, souvent avec des retombées scientifiques et sociétales estimées plus importantes. De ce fait, nous manquons de ressources humaines (E/C et/ou ingénieurs) disponibles pour ces activités à destination des entreprises.

Finalement, il n'existe donc pas de partenaire non-académique pérenne ou bien identifié sur ce contrat et ce type d'interactions, relativement limité, s'effectue essentiellement au travers des consortiums de projets régionaux qui intègrent ces acteurs. Malgré cela, au travers de l'engagement de l'équipe dans le RTR DIAMS ainsi que dans le projet Junon autour duquel gravitent de nombreux partenaires, la visibilité de l'équipe devrait encore s'accroître au niveau régional, mais aussi national et international. Par exemple, le projet d'ARD régionale JUNON intègre un modèle économique et une stratégie partenariale de moyen terme visant à pérenniser et valoriser les travaux effectués sur toute la chaîne de valeur. Ainsi, des discussions ont déjà été amorcées avec EDF et la ville de Tours sur la création de Jumeaux numériques.

### Référence 2 : L'équipe développe des produits à destination du monde socio-économique

Pendant ce contrat, on note une nette et forte augmentation (par rapport au précédent contrat) de l'activité de valorisation et de transfert. Cette activité s'inscrit en partie dans le cadre de projets de recherche régionaux tels que ceux évoqués précédemment (projets Valmod, ARD Junon, CIFREs,...). Ce volet de valorisation est souvent une condition imposée par la région que nous prenons en compte le plus souvent possible en intégrant le CETU ILIAD3 aux dépôts de projets. En dehors de ces projets, l'équipe s'engage également ponctuellement dans des expertises scientifiques et techniques au travers de prestations réalisées par le CETU ILIAD3, avec l'expertise des membres de l'équipe. Ces produits concernent la valorisation patrimoniale et le tourisme (projets Vivamemori, Vitrail) ou divers sujets : prestation pour le CEA sur l'analyse d'images de matériaux ; prestation GoalLine sur la détection de buts au football avec du Lidar ; projet CLIMENVI pour les chambres d'agriculture sur la visualisation (projections jusqu'en 2100) d'indicateurs climatiques et agro-climatiques pour la viticulture. Ce dernier projet est en cours de valorisation par la mise à disposition, sous licence libre, de l'application et de son code source sur le serveur [Git de l'Université](#), notamment pour permettre de le faire vivre et de le faire évoluer (reprise et adaptation aux spécificités d'autres régions prévues par les chambres d'agriculture).

Au-delà de ces prestations et contrats de collaboration il faut noter l'implication des membres de l'équipe dans le montage de projets de valorisation/start-ups qui ont reçu le soutien de l'organisme régional [C-Valo](#) :

1. [Sila-3D](#) : Plateforme de segmentation d'images biomédicales 3D résultant de plusieurs projets pluridisciplinaires régionaux (voir Portfolio). Ce projet se poursuivra dans l'avenir puisqu'une bourse de thèse co-financée par la Région et l'INRAE va débiter en septembre 2022, qu'un projet visant à exploiter des GNN pour la comparaison de cerveaux vient d'être soumis au dernier appel à projet régional et que ces recherches constituent l'un des axes du projet Loire-Val-Health qui vient d'être déposé dans le cadre du PIA4.
2. [Todd.TV](#) : projet de start-up s'appuyant sur les travaux en analyse de vidéos temps réel avec applications pour l'analyse de flux TV (cf. Portfolio)

Au-delà de ces réalisations, un membre de l'équipe est intervenu dans la rédaction de normes (coordination jusqu'en 2016 pour l'International Electrotech. Committee et/de l'International Standard Organization).

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

Les interventions de l'équipe auprès du grand public se manifestent en premier lieu au travers d'enseignements en intelligence artificielle et reconnaissance de formes réalisés à l'Université de Tours : formation en médecine, au CESR sur l'analyse de documents patrimoniaux, à l'INSA de Blois et dans le master BDMA de l'Université de Tours, dans le parcours IA de la spécialité informatique de Polytech Tours, formations dans le cadre de la SFR FED ([Voir supports disponible sur le site LIFAT](#)). Un investissement s'effectue également au travers de l'encadrement de projets étudiants de ces formations ou encore de vulgarisation auprès des jeunes (stages de 3ème à Polytech Tours,

Également, nous intervenons régulièrement auprès du grand public ou des acteurs socio-économiques par exemple au travers de :

- la fête de la science, des Meet-up IA en région Centre ou de la publication d'articles vulgarisés ou d'interviews dans les médias ([un exemple](#)).
- Conférence Débat « [Cosmetic Expert : Intelligence Artificielle & Industrie du Futur](#) : Et si on collaborait... » Webinaire Cosmetic valley - décembre 2020
- Table ronde « [Numérisation et mémorisation du patrimoine](#) », à Tours, Semaine de la mémoire - B2V, septembre 2021.
- Conférence « Deep Learning & Graph Representation for Image Analysis ». 10emes Journées Scientifiques et Techniques du Réseau des Microscopistes INRAE, sept 2021 – Tours
- Diffusion de la culture scientifique par participation, plusieurs fois; à des interventions en collèges comme parrain scientifique dans le cadre d'actions de science participative initiées par « la main à la pâte » ou par la « Maison de la sciences » ([collège Corneille](#) à Tours, collège de Château Renault dans le cadre du [projet Neuro2co](#),

Finalement, le [site web de l'équipe](#) a été très largement revu au début du contrat avec une diffusion d'informations qui s'effectue assez régulièrement mais dont la mise à jour régulière (en termes de projets, partenaires, ...) laisse à désirer.

### Synthèse de l'autoévaluation

	Forces	Faiblesse
Interne	<ul style="list-style-type: none"> <li>- Soutien du CETU ILIAD3 pour le transfert et la valorisation</li> <li>- Mise à disposition de plus en plus systématique sur des Gits ou sites web des données et codes développés</li> <li>- Croissance de l'activité dans ce domaine</li> </ul>	<ul style="list-style-type: none"> <li>- Pas de partenaires non académiques pérennes</li> <li>- Manque de personnels permanents (ingénieurs) pour diffuser des outils et données de bonne qualité et assurer le maintien des librairies développées</li> <li>- Communication auprès du grand public qui reste assez légère (là encore surtout par manque de temps)</li> <li>- Manque d'un personnel pour l'animation des sites web et/ou la diffusion d'informations sur les réseaux.</li> </ul>
	Opportunités	Menaces
Externe	<ul style="list-style-type: none"> <li>- Demande forte des institutions dans ce domaine</li> <li>- structuration régionale et réseaux</li> </ul>	<ul style="list-style-type: none"> <li>- A moyens constants, une augmentation de cette activité entrainera de facto une baisse du temps consacrée aux autres travaux scientifiques</li> </ul>

## 4.4. Equipe ROOT

### Référence 1. L'équipe se distingue par la qualité de ses interactions non-académiques.

Comme lors du précédent contrat, l'équipe s'est fortement investie dans des partenariats avec des entreprises afin de valoriser les compétences scientifiques des membres au travers de développement de projets en commun. Les principaux partenariats sont détaillés ci-dessous avec leur impact sur la période 2016-2021 :

- Le CHRU de Tours et la société Etycsis (depuis 2007) : L'équipe travaille depuis presque 15 ans sur des problématiques de chaîne logistique autour de la production de chimiothérapies, depuis la prescription jusqu'à l'administration au patient. Tout d'abord deux prototypes logiciels ont été conçus, l'un nommé PLANIF servant à planifier en temps réel et de façon interactive avec les préparateurs, la production des chimiothérapies sur une journée, l'autre nommé PILOTE assurant la traçabilité des informations ainsi que la remontée de données en temps réel. Ces deux prototypes ont ensuite été transférés à la société Eticsys pour qu'elle les reprenne et les commercialise. Depuis, l'équipe continue à travailler avec le CHRU de Tours (en particulier l'unité oncologique en charge des chimiothérapies) et Eticsys pour faire évoluer ces outils et étendre les travaux de recherche en y intégrant les aspects de livraison et gestion des reliquats des molécules cytotoxiques (très coûteuses pour l'hôpital). Ce fut l'objectif du projet DISTRI-CHIMIO financé par la région Centre Val de Loire de 2017 à 2020. Ces travaux ont donné lieu à 2 articles dans des revues internationales et 10 communications en conférence.
- La Compagnie des Mobilités (2008-2018) : cette jeune entreprise, dont le cofondateur est un ancien docteur de l'équipe, vise à promouvoir le développement des modes de déplacement doux tel que le vélo, notamment grâce à leur application Géovélo (<http://www.geovelo.fr>). Sur la période, l'équipe a continué à collaborer au développement d'un calculateur d'itinéraires à vélo multicritères au travers d'une CIFRE. Ces travaux ont contribué indirectement à l'application Géovélo et ont été publiés dans 1 article dans une revue internationale (cf. élément 5 du portfolio ROOT) et 8 communications en conférences.
- L'entreprise Cyres-groupe (2013-2017) : ce partenariat s'est déroulé dans le cadre d'une thèse CIFRE portant sur l'optimisation et la gestion des ressources sur une plate-forme informatique de type Big Data et basée sur le système Hadoop. Les travaux, qui ont donné lieu à 5 publications en conférence, ont porté sur la résolution de problèmes d'ordonnement de tâches sur des serveurs conçus spécifiquement pour stocker et analyser de grandes quantités de données non structurées au sein d'un environnement de calcul distribué.
- La société FICO (2018-2020) : l'équipe a travaillé avec cette société pour réaliser un module d'optimisation multi-objectif intégré au langage MOSEL via des stages de master de recherche (1 publication en conférence internationale).
- STMicroelectronics, depuis 2018 : dans le cadre d'une thèse CIFRE, l'équipe s'intéresse à un système de manutention automatisé pour les lots de plaquettes dans un atelier de photolithographie d'une usine de fabrication de semi-conducteurs. Les travaux visent à développer un outil d'aide à la décision pour la planification des activités de production avec la gestion des compétences et de la qualification de chaque opérateur ainsi qu'un modèle de simulation pour évaluer différents scénarios et proposer une conception efficace de ce système.
- SNCF (depuis 2018) : en collaboration avec le technicentre de la SNCF à St-Pierre-des-Corps, l'équipe travaille sur des problématiques d'ordonnement d'opérations de maintenance sur les rames circulant pour le transport de personnes.

Une CIFRE a débuté en 2020, cependant elle fut interrompue 1 an plus tard par l'abandon de la doctorante (pour raisons personnelles). Une nouvelle CIFRE vient d'être acceptée par l'ANRT, permettant ainsi de poursuivre ce travail dès 2022.

- SANOFI (depuis 2020) : dans le cadre d'une CIFRE démarrée en avril 2020, l'équipe s'intéresse à la logistique d'un centre de distribution de médicaments en charge de la préparation des commandes, de la constitution des palettes et des affectations aux transporteurs. L'objectif est de développer des outils d'aide à la décision pour optimiser la performance des lignes de préparation, ainsi que la performance globale du centre de distribution où les quantités de produits à gérer peuvent être très importantes.

Enfin, depuis 2021, l'équipe est fortement impliquée dans le montage d'une chaire industrielle sur la thématique de l'Usine du Futur avec comme partenaires Polytech Tours et des entreprises locales (ST Microelectronics, SKF, SII, Mecachrome, ATOS, ...). Celle-ci permettrait de fédérer et dynamiser les investissements des entreprises dans leur développement en Région en une dynamique régionale structurante. Le défi de l'industrie du futur 4.0 a été identifié par la Région Centre-Val de Loire dans son Schéma Régional de Développement Économique, d'Innovation et d'Internationalisation (SRDEII). Cette Chaire permettrait de développer davantage des projets multidisciplinaires (entre sciences humaines et innovations technologiques) sur des problématiques concrètes soulevées par les entreprises confrontées aux inévitables transformations de leurs procédés de fabrication. Le projet, porté par un membre de l'équipe, a vu son budget augmenté d'années en années.

Année	Université de Tours	Métropole de Tours	Région Centre Val de Loire	EIT (Europe)	Total
2020	2k€				12k€
2021		30k€		88k€	118k€
2022			108k€	202k€	310k€
2023					

Après des financements de l'université, de la Métropole et de la Région Centre-Val de Loire, le dernier financement obtenu est un financement Européen par l'EIT (European Institute of Innovation & Technologies) SMART'2M (<https://eit-hei.eu/projects/smart-2m/>) appuyant sur Innovation et formations sur l'industrie du futur (partenariat entre Irlande, Serbie, Italie, Grèce, Belgique et France).

L'originalité du projet est de fédérer plusieurs laboratoires en sciences de l'Université de Tours (LIFAT, Greman, Lamé) et des laboratoires de SHS comme Qualipsy, IRJI et Valorem, soit en tout plus d'une trentaine d'enseignants-chercheurs directement impliqués. Cette démarche permet de répondre de manière pertinente à l'amélioration du taux de réussite des projets de recherche portant sur la transformation des process de production. Pour la partie informatique, l'optimisation du système de production intervient dans plusieurs projets intégrés à la Chaire. Deux domaines émergent comme particulièrement pertinents : l'IA pour l'optimisation et la fabrication additive (sans exclure d'autres défis qui pourraient être relevés par le groupe en place).

## Référence 2 : L'équipe développe des produits à destination du monde socio-économique.

Au total, 5 CIFRE ont permis à l'équipe d'interagir directement avec le monde socio-économique (Cyres, La compagnie des mobilités, STMicroelectronics, SNCF et SANOFI). Il y a eu également le projet DISTRI-CHIMIO qui a permis de maintenir une collaboration au long cours avec le CHRU de Tours et la société Eticsys. Ces dispositifs, présentés dans la section précédente, ont été propices à la diffusion de résultats via des documents de travail, des présentations en réunion ou directement par le travail du doctorant lorsqu'il travaillait en entreprise. En 2021, l'équipe a également été impliquée dans l'obtention d'un brevet sur une procédure de virtualisation des fonctions réseau. En plus de ces coopérations, l'équipe intervient auprès d'acteurs du monde socio-économique au travers de projets d'étudiants en école d'ingénieurs (département informatique – Polytech Tours) notamment par des projets de fin d'études en dernière année ou des projets collectifs d'étudiants en quatrième année. Certains sujets de projets ont été donnés et encadrés par des membres de l'équipe et en collaboration avec une entreprise, une association ou un organisme du secteur public. Ces projets à la fois pédagogiques et directement en lien avec les thématiques de recherche de l'équipe ROOT, ont rendu possible l'étude voire le développement de prototypes à destination du monde socio-économique dont voici quelques exemples :

- Planification de tournées de techniciens chez EDF : sectorisation et routage pour assurer un maximum d'interventions (dépannages, relèves de compteur, maintenance).
- CEA – Le Ripault : Outil pour l'ordonnancement et la planification d'ordres de missions de transports spéciaux (marchandises dangereuses).
- GoSystèmes : Optimisation en temps réel de l'ordonnancement des ordres de fabrication dans le milieu de la peinture industrielle.
- CHRU de Tours : Outil d'aide à la conception de planning de personnel de soins à moyen terme.
- Keolis-Tours : Planification d'opérations de maintenance des tramways et aide à l'affectation des rames aux services.
- Radiall : Ordonnancement des tâches à réaliser dans une cellule de production flexible.
- STMicroelectronics Tours : Outil pour la gestion des compétences et planification d'opérateurs.
- DELPHI Technologies Blois : Planning collaboratif.

Enfin, notons qu'un ancien doctorant de l'équipe, Ismaila Ndiaye, a créé une start-up (MakeltMap) proposant une solution web pour l'optimisation et la planification de tournées dans plusieurs secteurs (Optitourplan). Un autre ancien doctorant de l'équipe, Azeddine Cheref est sur le point de créer sa SAS (Société par Actions Simplifiée) autour d'une plateforme web pour la mutualisation des livraisons de produits en circuits courts.

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

L'équipe ROOT participe chaque année à plusieurs actions de vulgarisation. Tout d'abord, elle anime un stand de jeux illustrant des problèmes classiques de RO lors de la fête de la science qui a lieu chaque année pendant deux jours. Elle participe également au programme régional « Ingénieur(e), toi aussi ! » à destination de lycéens au travers d'exposés sur les problèmes de plus court chemin (deux exposés d'une demi-journée par an). D'autres actions de vulgarisation ont été également menées pendant la période dont les principales sont les suivantes :

- Rédaction d'un article de diffusion scientifique dans le bulletin semestriel de la ROADEF Automne portant sur les "Défis pour la RO pour la planification de soins à domicile" en 2016.
- Exposé sur l'« Extraction de connaissances sur des données métaboliques » lors de la 1ère journée du groupe Oncologie 4.0 en 2017 à Angers.
- Participation à une table ronde autour de la conscience et de l'intelligence artificielle, organisée par Pascal NOUVEL dans le cadre de la fête de la science, traducteur de Tom Stoppard « The Hard problem » en octobre 2018.
- Interview par Sophie Bécherel de France-Inter pour sa chronique "Futur proche" diffusée en 2017 (<https://www.franceinter.fr/emissions/futur-proche/futur-proche-12-mai-2017>).
- Exposé sur « L'IA et les financements de la recherche » lors du colloque L'IA au service des entreprises, organisé par la DIRECCTE, à Orléans en octobre 2019.
- Exposé sur « L'intelligence Artificielle » lors de la conférence de l'AMOPA (Association des Membres de l'Ordre des Palmes Académiques 37) en janvier 2020.
- Conférencier invité lors de l'événement UTL (Université du Temps Libre) de l'Université de Tours en mai 2021, l'exposé (en visio) a porté sur « L'intelligence Artificielle ».
- Rédaction de l'article intitulé "la planification des blocs opératoires ?" dans la Bibliothèque Tangente n°75 -2021- numéro spécial portant exclusivement sur la Recherche Opérationnelle.

### Synthèse de l'autoévaluation

Les activités de recherche de l'équipe ROOT en Ordonnancement et en Transport conduisent à une production scientifique importante en volume et en qualité, à une grande visibilité et à des collaborations de recherche au niveau international et national. Les collaborations avec les entreprises sont fréquentes et donnent lieu à la mise en œuvre de travaux de recherche.

Dans son projet scientifique, l'équipe accentuera ses contributions principalement sur les thématiques de la santé et de l'industrie du futur. L'actualité récente montre que ces défis représentent des enjeux capitaux pour notre pays : l'amélioration de l'efficacité des systèmes de santé s'avère indispensable pour les personnes et l'augmentation de la compétitivité s'avère indispensable pour les entreprises. Ces deux grandes thématiques serviront de support au développement de nouvelles problématiques de recherche en ordonnancement et en transport, auxquelles s'ajouteront la question de l'exploitation de grands volumes de données, que ce soit des données relatives à la santé ou aux systèmes de santé, ou des données relatives au fonctionnement des industries.

Pour ancrer ces travaux dans ces deux thématiques, l'équipe ROOT collabore par l'intermédiaire de conventions CIFRE. Ce dispositif ayant de nombreux avantages, l'équipe envisage de poursuivre dans cette voie. Un projet de chaire industrielle sur l'industrie du futur est en cours de montage, porté par un membre de l'équipe. Ce projet, qui présente plusieurs aspects innovants et qui a un caractère fortement pluri-disciplinaire sera un atout pour développer la thématique sur l'industrie du futur.

Pour inscrire ses activités de recherche dans la société, et particulièrement dans les entreprises, l'équipe ROOT collabore principalement avec des contrats de type CIFRE. Ce dispositif ayant de nombreux avantages, l'équipe envisage de poursuivre sur cette voie (actuellement 3 nouveaux contrats CIFRE sont en négociation). Le projet de la chaire industrielle sur la thématique de l'Usine du Futur permettra également de mener de nouvelles collaborations avec des partenaires non-académiques. De plus, le sujet de cette chaire est propice à de nouveaux projets de recherche innovants et en cohérence avec la politique de recherche de l'équipe. De nombreux prototypes sont développés au sein de l'équipe au travers de projets étudiants en école d'ingénieurs en informatique ou en informatique industrielle et à destination du monde socio-économique. Même s'ils sont souvent pertinents et montrent bien la plus-value des résultats des travaux de recherche, le transfert technologique restent encore à améliorer car beaucoup trop de ces prototypes finissent par être "rangés dans des cartons" (la maintenance et l'évolutivité de ces logiciels ne pouvant pas être assurées par des enseignants-chercheurs). Enfin, l'équipe ROOT essaye de s'impliquer dès qu'elle le peut dans le partage des connaissances et la vulgarisation scientifique bien que la recherche opérationnelle ne soit pas parmi les sciences les plus faciles à montrer au grand public.

## ANNEXE A : Données LABRI – Activités à l'international du LIFAT

Les données contenues dans cette annexe proviennent d'une extraction de la base LABRI de l'université de Tours répertoriant les activités majeures des laboratoires ç l'international. Ces données sont exploitées pour le calcul du BQI (Bonus Qualité International) versé au équipes chaque année.

### Cotutelles

Etudiant	Date de première inscription	Pays	Directeur de thèse	Libellé de la thèse
HADJ MOHAMED NAJAT	25/01/2022	Algérie	AGATA SAVARY	Identification des expressions poly lexicales en arabe
VENDE PIERRE	20/04/2021	Canada	EMMANUEL NERON	Outil d'aide à la décision pour la mise en place d'un réseau de transports en commun électriques
SERRA SOSA FLAVIA	13/01/2020	Uruguay	PATRICK MARCEL	Gestion de la qualité des données, basée sur le contexte, dans les systèmes d'information hétérogènes
ONDZIGUE MBENGA RAYMOND	30/04/2019	Gabon	THOMAS DEVOGELE	Etude et mise en place d'un système national d'information géographique décisionnel utilisant les technologies mobiles pour la surveillance intégrée de la tuberculose dans les pays de la zone CEMAC : SYN@TB, cas du Gabon
BRAHEM AMINA	03/04/2019	Tunisie	THOMAS DEVOGELE	Coordination et adaptation dynamique d'exécution de services entre entités autonomes à travers les technologies Blockchain. Application à la réalisation de plans de séjours touristiques dans le cadre du projet Smart-Loire
KOMEIHA FOUAD	25/02/2019	Liban	THOMAS DEVOGELE	Construction of a semantic similarity calculation environment based on linked open data and the semantic web for the recommendation of tourisme services in the context of the smartloire project
HAROUNE MEYA	20/11/2018	Mauritanie	EMMANUEL NERON	Ordonnement multi-projets à contraintes de ressources multicompetences partagées
AHMED SIDI MOHAMED LEMINE	23/11/2018	Mauritanie	EMMANUEL NERON	Nouvelles approches informatiques et mathématiques pour la résolution de problèmes biologiques
HAOUASSI MUSTAPHA	03/10/2016	Algérie	JEAN-CHARLES BILLAUT	Rich vehicle routing problems
SIDI ALY EL ARBY	23/09/2016	Mauritanie	BEATRICE MARKHOFF	Enrichissement d'ontologies à partir du web sémantique

Avec conventions officielle signée entre les 2 pays

### Projets internationaux

Nom du projet	Nom d'appel à projet	Année de	Année de fin	Coordinateur du projet	Statut
COST Action Proposal OC-2021-1-25365 "Machine Learning and Graphs "	H2020 - Défis sociétaux	2021	2022	Université de Rouen	déposé
COST UNIDIVE TAL multilingues	H2020 - Défis sociétaux	2021	2022	Université de Tours	déposé
ANONYMOUS - A unified audio-video analysis framework for surveillance and security	ERC - Consolidator Grants	2018	2023	Laboratoire d'Informatique - Université François-Rabelais de Tours	refusé
Learning for Operations Research (LOR)	ANR International	2019	2022	Université François-Rabelais de Tours	refusé
Learning for Operations Research (LOR)	ANR International	2018	2021	Laboratoire d'Informatique - Université François-Rabelais de Tours	refusé
Mnemon	H2020 - Défis sociétaux	2018	2021	Universität Passau	refusé
ARIADNEplus - Advanced Research Infrastructure for Archaeological Data Networking in Europe - plus	H2020 - Défis sociétaux	2019	2020	PIN	accepté
DACCIT - Digital Analysis of Contrapuntal Combinations and their Intertextual Transformations in Renaissance Music	ANR International	2019	2022	Centre d'Etudes Supérieures de la Renaissance - Université François-Rabelais de Tours	refusé
Digital model for Oncology	Appel à projet par un organisme français Studium	2021	2021	INSERM 1069 - Val de Loire	Accepté
Bourse de these VIED 911	Project 911 - Vietnam International Education Department Fellowships	2019	2022	Université de Hong Duc Vietnam	Accepté
Studium Namub : Alternative animal model to understand the brain	Appel à projet par un organisme français Studium	2019	2020	Institut National de la Recherche Agronomique - Val de Loire	refusé
Adam-IoT - Advanced Data Analytics Models on Fuelling Stations for the Intelligence of Things	H2020 - Défis sociétaux	2019	2022	Università degli Studi di Milano	refusé
H2020 4CH - Competence Centre for the Conservation of Cultural Heritage	H2020 - Défis sociétaux	2021	2023	INFN - Istituto Nazionale di Fisica Nucleare	accepté

Liste éventuellement non exhaustive (seuls certains types de projets peuvent être répertoriés)

### Mobilités sortantes

Nom	Prenom	Statut	Etablissement de destination	Date début	Date de fin	semaine
NGUYEN	DINH CONG	Doctorant	VIET NAM - Hong Duc University	16/12/2018	16/02/2019	9
DELALANDRE	MATHIEU	MCF	VIET NAM - Hong Duc University	29/10/2018	10/11/2018	2
T'KINDT	VINCENT	MCF	ITALIE - Politecnico di Torino	02/10/2018	27/10/2018	4
DELALANDRE	MATHIEU	MCF	VIET NAM - Hong Duc University	31/10/2019	12/11/2019	2
T'KINDT	VINCENT	PR	ITALIE - Politecnico di Torino	09/09/2019	04/11/2019	8
MARCEL	PATRICK	MCF	ITALIE - Università degli Studi di Bologna 'Alma Mater Studiorum'	04/02/2018	03/03/2018	4
MARCEL	PATRICK	MCF	GRECE - Panepistimio Ioanninon	13/05/2018	27/05/2018	2
T'KINDT	VINCENT	PR	ITALIE - Politecnico di Torino	31/08/2020	19/10/2020	7
PERALTA	VERONIKA	MCF	ESPAGNE - Universitat Internacional de Catalunya	09/02/2020	22/02/2020	2
MENDOZA	JORGE	MCF	CANADA - Centre interuniversitaire de recherche sur les reseaux d'entreprise, la logistique et le transport	01/08/2016	28/02/2017	31
DELALANDRE	MATHIEU	MCF	VIET NAM - Hong Duc University	29/10/2017	13/11/2017	3
SAVARY	AGATA	MCF	ALLEMAGNE - Heinrich-Heine-Universität Düsseldorf	01/09/2017	28/02/2018	26
BILLAUT	JEAN-CHARLES	PR	ITALIE - Politecnico di Torino	13/05/2017	27/05/2017	2
T'KINDT	VINCENT	PR	ITALIE - Politecnico di Torino	31/08/2021	19/10/2021	7
TAIBU BIRGUI	SEKOU	Doctorant	University of Liverpool	22/04/2019	11/05/2019	3
						101

Mobilités supérieures à 2 semaines uniquement

### Mobilités entrantes

	Nom personne arrivante	Prenom personne arrivante	Etablissement de provenance	Date de début	Date de fin	semaines
Postdoc	ABDELLATIF	AMIR	TUNISIE - Ecole Nationale d'Ingénieurs de Monastir (ENIM)	01/04/2020	30/06/2020	13
Doctorant	KENNICHE	AHLEM	ALGERIE - Université de Mostaganem	26/02/2020	15/03/2020	3
Prof	AL-HAJJ	MOUSTAFA	LIBAN - Université Libanaise	01/12/2019	21/12/2019	3
PostDoc	RENER	ELENA	ITALIE - Politecnico di Torino	04/11/2019	19/12/2019	7
PostDoc	SERRA	FLAVIA	URUGUAY - Universidad de la República	01/11/2019	31/12/2019	9
Doctorant	DIOP	CHEIKH TALIBOUYA	SENEGAL - Université Gaston Berger de Saint-Louis	01/10/2019	10/11/2019	6
Prof	GROSSI	GIULIANO	ITALIE - Università degli Studi di Milano - Bicocca	18/06/2019	10/07/2019	4
Doctorant	ABDELLATIF	AMIR	TUNISIE - Ecole Nationale d'Ingénieurs de Monastir (ENIM)	01/04/2019	30/06/2019	13
Prof	SCATAMACCHIA	ROSARIO	ITALIE - Politecnico di Torino	25/11/2018	21/12/2018	4
Doctorant	HABIB	MOHAMED	MAURITANIE - Ecole Normale Supérieure de Nouakchott	01/09/2018	28/02/2019	26
Doctorant	BOUH	MEYA	MAURITANIE - Ecole Normale Supérieure de Nouakchott	01/09/2018	28/02/2019	26
Doctorant	LEMINE	MOHAMED	MAURITANIE - Ecole Normale Supérieure de Nouakchott	01/09/2018	28/02/2019	26
Doctorant	DIOP	LAMINE	SENEGAL - Université Gaston Berger de Saint-Louis	01/09/2018	31/12/2018	18
Doctorant	DHB	CHEICK M	MAURITANIE - Université de Nouakchott	15/07/2018	10/09/2018	9
Prof	LANZAROTTI	RAFFAELLA	ITALIE - Università degli Studi di Milano	01/07/2018	15/07/2018	2
Prof	SERRATOSA	FRANCESC	ESPAGNE - Universitat Rovira i Virgili	25/06/2018	22/07/2018	4
Postdoc	ABDELLATIF	AMIR	TUNISIE - Ecole Nationale d'Ingénieurs de Monastir (ENIM)	01/04/2018	30/06/2018	13
Doctorant	DIOP	CHEIKH TALIBOUYA	SENEGAL - Université Gaston Berger de Saint-Louis	25/09/2017	23/10/2017	4
Doctorant	BABOU	HAFEDH MOHAMMED	MAURITANIE - Ecole Normale Supérieure de Nouakchott	17/07/2017	30/09/2017	11
Prof	BOCCIGNONE	GIUSEPPE	ITALIE - Università degli Studi di Milano - Bicocca	25/06/2017	25/07/2017	5
Prof	LARICHI	NADIA	CANADA - Ecole Polytechnique de Montréal	14/05/2017	28/05/2017	2
Prof	TRANG	Hong Son	VIETNAM - université de Technologie de Ho Chi Minh Ville	01/01/2017	03/03/2017	9
Prof	GOODSON	JUSTIN	ETATS-UNIS - Saint Louis University	07/05/2017	28/05/2017	3
Prof	BLUMENSTEIN	MICHAEL	AUSTRALIE - University of Technology Sydney	11/11/2016	23/11/2016	2
Prof	DELLA CROCE	FEDERICO	ITALIE - Politecnico di Torino	01/11/2016	30/11/2016	5
Doctorant	CHENIK	NASREDDINE	ALGERIE - Université des sciences et de la technologie Houari Boumediène	16/10/2016	20/05/2017	31
Prof	AL-HAJJ	MOUSTAFA	LIBAN - Université Libanaise	19/09/2016	13/10/2016	4
Prof	BLUMENSTEIN	MICHAEL	AUSTRALIE - University of Technology Sydney	02/07/2016	16/07/2016	2
Doctorant	FURTADO	PEDRO	PORTUGAL - Universidade de Coimbra	01/03/2016	31/03/2016	5
Doctorant	LHIOUI	MALEK	TUNISIE - Université de Sfax	01/03/2016	31/03/2016	5
Prof	GOODSON	JUSTIN	ETATS-UNIS - Saint Louis University	01/02/2016	15/03/2016	7
Doctorant	KULLMAN	NICHOLAS	ETATS-UNIS - George Washington University	01/01/2016	31/03/2016	13
					Total	294
					moyenne	9,1875

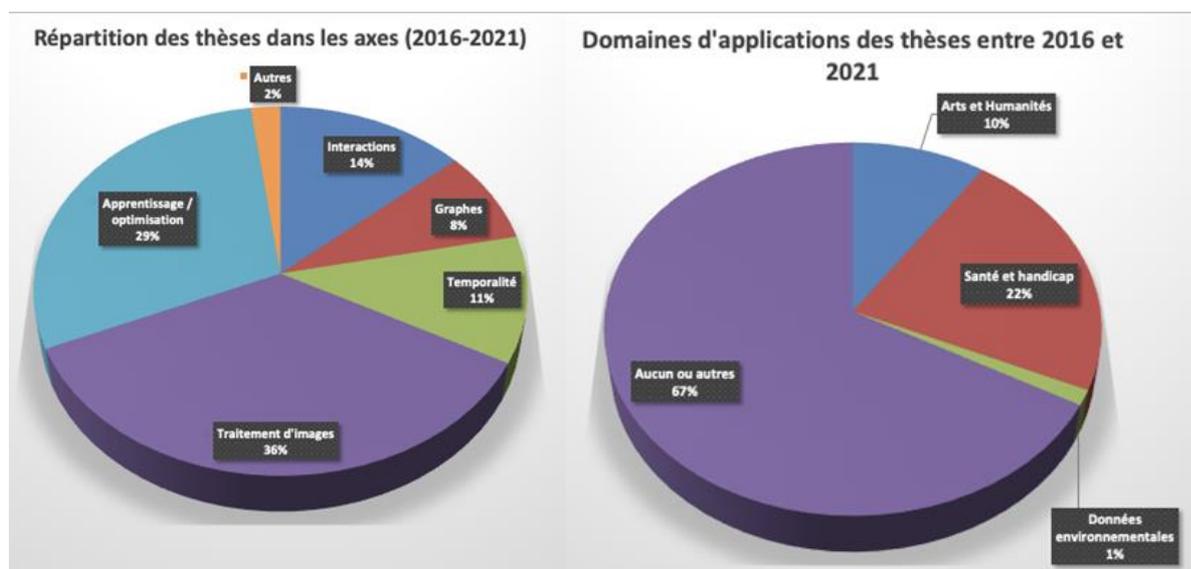
Mobilités supérieures à 2 semaines uniquement

## ANNEXE B : ANNEXE RFAI DU DAE

### Liste des charges administratives afférentes aux membres de l'équipe sur le contrat :

- Direction de laboratoire LIFAT (7 ans)
- Direction de l'école doctorale MIPTIS (6 ans)
- Vice-présidence en charge du numérique
  - À l'Université de Tours (4 ans)
  - À l'INSA-CVL (directeur fonctionnel – 2 ans)
- Direction de départements :
  - Informatique de Polytech – 6 ans,
  - Informatique de l'UFR Sciences Tours – 3 ans
- Direction du département du premier cycle (parcours des écoles d'ingénieurs Polytech)
- Direction des études de la spécialité informatique (6 ans) et direction adjointe du département informatique (2 ans)
- Co-responsable du DUT Information Numérique dans les Organisations (2 ans)
- Responsable de la licence professionnelle technico commerciale à l'IUT (1 an)
- Autres :
  - Responsable des relations avec les entreprises pour le département informatique de Polytech Tours (4 ans)
  - Participations aux conseils de composantes/laboratoires (permanent)
  - Chargé de mission : relations avec les lycées (6 ans) et collégiens
  - Responsable VAE (Spécialité informatique)
  - Responsable relations internationales (formation par apprentissage) ; Responsable des mobilités entrantes (spécialité informatique) ; correspondant relations internationales (laboratoire)
  - Commission Numérique et Pédagogie Innovante à l'IUT
  - Comité numérique de l'Université de Tours (1 an)
  - Responsabilités d'années : spécialité informatique par apprentissage (3A, 5A), spécialité informatique (4A, 5A), L2 et L3 UFR Sciences ;
  - Responsabilités et coordination de parcours (Architecture, Systèmes et Réseaux, IA)
  - Coordination LIFAT-INSA, (2 ans)

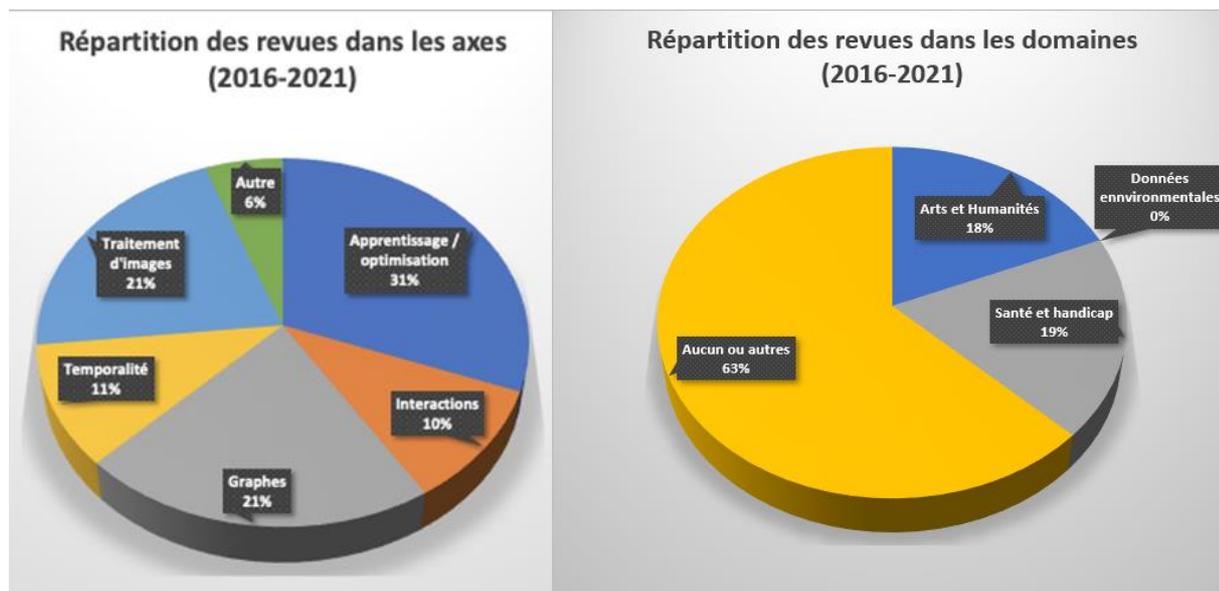
### Liste des thèses et répartition dans les axes :



Doctorant	Soutenance	LIFAT (1) / Externe (0)	Titre	Financement				Partenaires		Retombées socio-économiques
				Académique	Projet	CIFRE	Autre bourses	Académiques	Non académiques	
??	oct-26	1	En définition : Détection et reconnaissance Jockey dans les vidéos			1			Alezan.AI	
??	oct-26	1	En définition : diagnostique et pronostic à partir de données de soin et de recherche. Application à la sclérose latérale amyotrophique			1		CHRU, INSERM, LIFO	ATOS	Santé : diagnostic et pronostic (sclérose, pneumopathie, troubles neurodéveloppementaux)
??	oct-25	0	Graph Neural Networks for morpho-functional analysis and comparison of brain structures	1				INRAE, INSERM		
??	oct-25	1	Learning Spatio-temporal data by graph representations		1			LITIS, Greyc		
??	oct-25	0	New convolutional operators for GNN		1			LITIS, Greyc		
??	oct-25	1	integration of ad-hoc combinatorial post processing in learning architectures	1						
Thomas Kastner	mars-25	1	Détection de fraudes par prédiction de scores d'attrition			1			MACIF	
Simon Mignon	oct-24	0	Modèles génératifs et transport optimal pour la restauration d'images	1				IDP		
Farouq Benchalall	oct-24	0	Desherbrob : reconnaissance/localisation et prédiction de la pousse de mauvaises herbes		1			PRISME	Agreen Tech Valley, Frasem, Automachine	Environnement (limitation des intrants en agriculture)
Hellen Walle	oct-24	1	Intelligence artificielle et fouille de données pour la détection des troubles du spectre de l'autisme	1				CHRU, GIS Autisme		Santé et handicap : troubles autistiques
Praveen Soni	oct-24	0	Génération automatique des tableaux de bord des projets BI4PEOPLE		1			ERIC, ELICO, IRIT		
Ben Crulis	oct-24	1	Assistant intelligent pour l'aide aux déplacements des aveugles	1						Handicap : malvoyants
Hao Le	oct-23	1	Appariement de template temps-réel et robuste pour la détection de séquences TV / Real-time and robust template matching for the detection of partial-copy videos	1				HDU (Vietnam)		
Guillaume Lacharme	avr-23	1	Optimisation d'hyper-paramètres de modèles d'apprentissage profond	1						
Xiao-Lin Li	déc-21	0	Étude Numérique Visant à Améliorer l'Efficacité et la Précision des Algorithmes d'Optimisation par Essais de Particules Appliqués à la Détection des Dommages Structuraux d'une Poutre en Porte-à-faux				1	Lamé		
Boukhari Nouredine	déc-20	0	Algorithmes évolutionnaires pour la résolution de problèmes multi-objectifs				1	Université Mascara		
Nemich Mohamed Amine	déc-19	0	Modèles de fourrage bio-inspirés pour la résolution des problèmes d'optimisation à contraintes de ressources				1	Université Mascara		
Diem HO	mars-22	1	Approach for solving the graph-matching problem based on ant colony algorithm	1						Humanités
Luong Phat NGUYEN	déc-21	1	Extraction and Characterization of Spatial-Temporal Patterns in Videos	1						
Romain CARLETTO	mars-22	1	Méthodes d'apprentissage automatique pour améliorer et automatiser la création de formats publicitaires et prédire leur impact	1		1		Madmix digital		Création digitale
David Boas	juil-20	1	Realité augmentée pour la chirurgie de l'épaule			1		Imascap		Santé : chirurgie assistée par ordinateur
Quoc Cuong Le	déc-20	1	Couplage de cameras intelligentes	1						Vidéo surveillance
Taibou BIRGUI SEKOU	oct-19	1	Intégration d'approches par apprentissage de dictionnaires dans un modèle variationnel pour la segmentation	1						
Fabien Baradel	juin-20	0	Structured deep Learning for Video analysis		1					
Nguyễn Đình Công	juin-20	1	Template matching dans les captures vidéos				1	HDU (Vietnam)		
Haykal Vanessa	déc-19	1	Prévision de séries et classification avec apprentissage profond	1						
Maxime Martineau	nov-19	1	Reconnaissance d'insectes		1			IRBI		Environnement : suivi des espèces
Shah SHIVANI	mars-19	0	Machine Learning and visualization applied on Big data				1	CEA Saclay		
Mostafa Darwiche	déc-18	1	Méthodes de recherche opérationnelle pour la comparaison de graphes	1						
P. Da Silva De Carvalho	déc-17	0	Visual Platform for Integrating Data with Uncertain and Weak Structure				1	CRP (Luxembourg)		
Marouen Sta	sept-17	0	Etude de la structure de la substance blanche cérébrale par IRM pondérée en diffusion : comparaison tractographie IRM - tissu cérébral ; conception d'un outil d'aide à la décision pour la reconstruction tractographique		1			INSERM		Santé : imagerie médicale
Gaetan Galisot	mars-18	1	Segmentation interactive d'images 3d multimodales	1				INRAE		Santé : imagerie médicale
Oussama Hadjeri	avr-17	0	Segmentation de nerfs dans les images échographiques		1			PRISME		Santé : imagerie médicale
Abdessalem Bouzaïeni	mai-18	0	Extension automatique de l'annotation d'images pour la recherche/classification d'images			1		LORIA		
Arundhati Tarafdar	07/2017	0	Wordspotting from multilingual and stylistic documents				1	ISI Kolkata		Humanités
Frederic Rayar	11/2016	1	Clustering interactif multi-échelle pour l'indexation de masse d'images	1						Humanités
Julien Berhouet	2016	0	Chirurgie de l'épaule en RA				1	CHRU		Santé : chirurgie assistée par ordinateur
Zeina Abu Aisheh	2016	1	Indexation de masses de graphes : Application à la recherche d'images par le contenu	1						

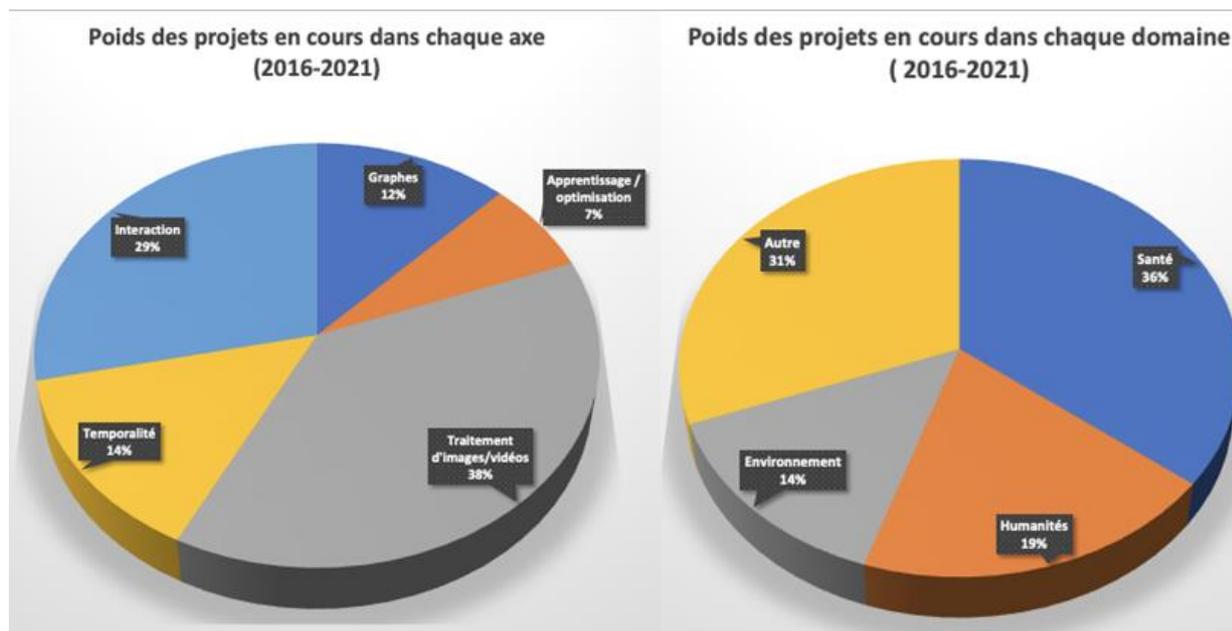
## Liste des revues principales (Q1 et Q2) et répartition dans les axes :

Auteurs	Année	Titre	Titre journal	co-auteurs étrangers	co-auteurs nationaux	Rang
Conte D., Grossi G., Lanzarotti R., Lin J., Petrini A.	2021	Analysis of a parallel MCMC algorithm for graph coloring with nearly uniform balancing	Pattern Recognition Letters	1		Q1
Huang Y., Li D., Niu H., Conte D.	2021	Visual identification of oscillatory two-phase flow with complex flow patterns	Measurement - Journal of the International Measurement Confederation (IMEKO)	1		Q1
Li X.-L., Serra R., Olivier J.	2021	An Investigation of Particle Swarm Optimization Topologies in Structural Damage Detection	Applied Sciences		1	Q2
Li X.-L., Serra R., Olivier J.	2021	Performance of Fitness Functions Based on Natural Frequencies in Defect Detection Using the Standard PSO-FEM Approach	Shock and Vibration		1	Q2
Martineau M., Raveaux R., Conte D., Venturini G.	2021	Graph matching as a graph convolution operator for graph neural networks	Pattern Recognition Letters			Q1
Raveaux R.	2021	On the unification of the graph edit distance and graph matching problems	Pattern Recognition Letters			Q1
Rosas-Gonzalez S., Birgui-Sekou T., Hidane M., Zemmoura I., Tauber C.	2021	Asymmetric Ensemble of Asymmetric U-Net Models for Brain Tumor Segmentation With Uncertainty Estimation	Frontiers in Neurology		1	Q2
Boccignone G., Conte D., Cuculo V., D'Amelio A., Grossi G., Lanzarotti R.	2020	An Open Framework for Remote-PPG Methods and their Assessment	IEEE Access	1		Q1
Conte D., Serratos F.	2020	Interactive online learning for graph matching using active strategies	Knowledge-Based Systems	1		Q1
Darwiche M., Conte D., Raveaux R., T'Kindt V.	2020	Graph Edit Distance: Accuracy of Local Branching from an application point of view	Pattern Recognition Letters			Q1
Silleresi S., Prevost P., Zebib R., Bonnet-Brihault F., Conte D., Tuller L.	2020	Identifying Language and Cognitive Profiles in Children With ASD via a Cluster Analysis Exploration: Implications for the New ICD-11	Autism Research			Q1
Berhouet J.*, Favard L., Boas D., Voisin T., Slimane M.	2019	Thoughts on a new surgical assistance method for implanting the glenoid component during total shoulder arthroplasty. Part 1: Statistical modeling of the native pre-morbid glenoid	Orthopaedics and Traumatology - Surgery and Research			Q1
Cortés X., Conte D., Cardot H.	2019	Learning Edit Cost Estimation Models for Graph Edit Distance	Pattern Recognition Letters			Q1
Darwiche M., Conte D., Raveaux R., T'Kindt V.	2019	A local branching heuristic for solving a Graph Edit Distance Problem	Computers and Operations Research			Q1
Mille J., Leborgne A., Tougne L.	2019	Euclidean Distance-Based Skeletons: A Few Notes on Average Outward Flux and Ridgeness	Journal of Mathematical Imaging and Vision		1	Q2
Werneck R. D. O., Raveaux R., Tabbone S., Torres R. D. S.	2019	Learning cost function for graph classification with open-set methods	Pattern Recognition Letters		1	Q1
Zhuang X., Li L., Payer C., Štern D., Urschler M., Heinrich M., Oster J., Wang C., Smedby Ö., Bian C., Yang X., Heng P.-A., Mortazi A., Bagci U., Yang	2019	Evaluation of algorithms for Multi-Modality Whole Heart Segmentation: An open-access grand challenge	Medical Image Analysis	1		Q1
Abu-Aisheh Z., Raveaux R., Ramel J.-Y., Martineau P.	2018	A parallel graph edit distance algorithm	Expert Systems with Applications			Q1
Abu-Aisheh Z., Raveaux R., Ramel J.-Y.	2018	Efficient k-nearest neighbors search in graph space	Pattern Recognition Letters			Q1
Alaei A., Conte D., Martineau M., Raveaux R.	2018	Blind Document Image Quality Prediction Based on Modification of Quality Aware Clustering Method Integrating a Patch Selection Strategy	Expert Systems with Applications			Q1
Boccignone G., Conte D., Cuculo V., Damelio A., Grossi G., Lanzarotti R.	2018	Deep construction of an affective latent space via multimodal enactment	IEEE Transactions on Cognitive and Developmental Systems	1		Q1
Martineau M., Raveaux R., Conte D., Venturini G.	2018	Learning error-correcting graph matching with a multiclass neural network	Pattern Recognition Letters			Q1
Slimane M., Berhouet J., Rol M., Spiry C., Chevalier C., Favard L.	2018	Shoulder patient-specific guide: First experience in 10 patients indicates room for improvement	Orthopaedics and Traumatology - Surgery and Research		1	Q1
Abu-Aisheh Z., Gaüzère B., Bougleux S., Ramel J.-Y., Brun L., Raveaux R., Héroux P., Adam S.	2017	Graph edit distance contest: Results and future challenges	Pattern Recognition Letters		1	Q1
Girard N., Bertet K., Visani M.	2017	Dichotomic lattices and local discretization for Galois lattices	Advances in Data Analysis and Classification		1	Q2
Martineau M., Conte D., Raveaux R., Arnault I., Munier D., Venturini G.	2017	A survey on image-based insect classification	Pattern Recognition		1	Q1
Mondal T., Ragot N., Ramel J.-Y., Pal U.	2017	Comparative study of conventional time series matching techniques for word spotting	Pattern Recognition	1		Q1
Abu-Aisheh Z., Raveaux R., Ramel J.-Y.	2016	Anytime graph matching	Pattern Recognition Letters			Q1
Alaei A., Raveaux R., Conte D.	2016	Image quality assessment based on regions of interest	Signal, Image and Video Processing			Q2
Hamdi A., Monmarché N., Slimane M., Alimi A.	2016	Fuzzy Rules for Ant Based Clustering Algorithm	Advances in Fuzzy Systems	1		Q2
Mondal T., Ragot N., Ramel J.-Y., Pal U.	2016	Flexible Sequence Matching Technique: An Effective Learning-free Approach For word-spotting	Pattern Recognition	1		Q1
Pham T. A., Delalandre M.	2016	Effective Decompression of JPEG Document Images	Transactions on Image Processing (TIP)			Q1
Moncef Hidane, Mireille El Gheche, Jf Aujol, Yann	2016	Image Zoom Completion	Transactions on Image Processing		1	Q1



### Liste des projets acceptés et répartition par axes :

Acronyme	Origine financement	type Financement	Type projet	Année début	Année fin	Porteur (0/1)	Partenaires académiques (dont CETU ILIAD3)	Autres partenaires	Retombées socio-économiques
Junon	Régional	ARD	recherche	2021	2024	0	IDP LIFO BRGM INRAE	ATOS, ANTEA, Farmviz, OVLT	Environnement : jumeaux numériques pour le suivi et la prédiction du niveau et de la qualité (eau/air)
ValoPat	Régional	CPER	investissement	2020	2024	0	CESR IPAT		Valorisation du patrimoine (immersion)
REGETO	National	GDR ISIS	networking	2021	2023	1	IDP		
SIRCUS	National	ANR	recherche	2022	2025	0	CETU, INSERM	GIS Autisme	Santé et handicap : autisme
CodeGNN	National	ANR	recherche	2022	2025	0	GREYC LITIS		
DOING	Régional	APR IA	recherche	2022	2025	0	LIFO, LLL		Santé
CEA Monts	Industriel	Prestation	Valo/transfert	2021	2022	0	CETU, CEA		Analyse de d'images de matériaux
Studium/Calcium	International	Studium	networking	2021	2022	0	UMR 1069 NCC, Université de Californie San Francisco		Santé : prédiction et décision en oncologie
DESHERBROB	Régional	APR-IR	recherche	2020	2023	0	PRISME	Automachine, Frasem, Agreen tech Valley	Environnement : réduction des intrants en agriculture
VECCAR	National	Projet MSH	networking	2020	2021	0	CETU, MSH		Humanités/patrimoine
Super IA	Régional	FEDER+Région	investissement	2019	2020	1	CETU, CASCIMODOT, BRGM		
BI4PEOPLE	National	ANR	recherche	2019	2024	0			
GoalLine	Industriel	Prestation	Valo/transfert	2019	2020	0	CETU	entreprise	Culture/loisirs
SILA3D	Industriel	C-Valo	Valo/transfert	2019	2020	1	CETU, INSERM, INRAE		Santé : imagerie médicale
Todd.TV	Industriel	C-Valo	Valo/transfert	2019	2020	1			Usage des médias
MadMix	Industriel	co-financée Région	recherche	2018	2021	1		MadMix Digital	Création digitale
Visit	Régional	APR IR	recherche	2018	2022	1	CETU, CESR	Loches (CG37)	Patrimoine
PEI CLIMENVI	Régional	FEADER	Valo/transfert	2018	2022	0		Chambres d'agriculture CVL, Vinopôle, IFV, Meteo France,	Environnement : viticulture et changement climatique
FIBRAVASC	Régional	APR IA	recherche	2017	2021	0	CETU, INSERM, CEA		Santé : imagerie médicale
Neuro2Co	Régional	APR IR	recherche	2017	2022	0	CETU, INRA, INSERM	Centre Science, Beauval, Neurospin	Santé : imagerie médicale
IMASCAP	Industriel	CIFRE	recherche	2017	2020	1	CHRU	IMASCAP	Santé : chirurgie assistée par ordinateur
DANIEAL2	Régional	APR	recherche	2016	2020	0	PRISME	Société Adechotech, Clinique Medipôle Garonne	Santé : imagerie médicale
DEEPEVISION	International	ANR / NSERC Canada	recherche	2016	2019	0	LIRIS, LIP6, University of Guelph (Canada), Simon Fraser University (Canada)		
FIBRATLAS II-III	National	ANR	recherche	2014	2021	0	CETU, INSERM – CHRU Tours		Santé : imagerie médicale
LUMINEUX	Régional	APR	recherche	2015	2019	0	PRISME		Environnement : éclairages automatisés
VALMOD	Régional	APR	recherche	2013	2019	0	CESR	Domaine de Chambord	Patrimoine
VIED911	International	VIED scholarship - program 911	recherche	2016	2019	1	Hong Duc University (HDU)		



### Liste des séminaires et soutenances de l'équipe

(cf. <http://www.rfai.lifat.univ-tours.fr/calendrier/>) :

- Présentation Fatima Debbat on 10 February 2022 13 h 00 min
- Apprentissage et Optimisation discrète on 12 November 2021 13 h 00 min
- Convolutional network slimming: convex quadratic programming for filter selection on 2 December 2021 10 h 30 min
- Differentiable Graph Matching (R. Raveaux) on 30 September 2021 10 h 30 min
- PhD defense of Luong Phat NGUYEN on 10 December 2021 13 h 45 min
- Soutenance HDR J. Mille on 1 July 2021 14 h 00 min
- Répétition soutenance Quoc on 3 September 2020 14 h 00 min
- PhD defense: "Tracking with smart cameras" (Quoc Cuong Le) on 7 September 2020 14 h 00 min
- Séminaire Projets transversaux LIFAT on 1 October 2020
- PhD defense: "Réalité Augmentée pour la chirurgie de l'épaule" (David BOAS) on 6 July 2020 14 h 00 min
- PhD defense: "Real-time LoG-based operator for scene text detection" (Nguyễn Đình Công) on 25 June 2020 9 h 30 min
- journée du RTR DIAMS on 14 November 2019 9 h 30 min
- PhD defense: Time series prediction with deep learning (V. Haykal) on 2 December 2019 14 h 00 min
- Répétition soutenance Maxime on 7 November 2019 10 h 30 min
- PhD defense – Deep learning onto graph space: application to image-based insect recognition (M. Martineau) on 13 November 2019 10 h 00 min
- Répétition soutenance de Taibou on 10 October 2019 10 h 00 min
- PhD defense Taibou BIRGUI SEKOU on 14 October 2019 14 h 00 min
- HDR Romain Raveaux on 26 June 2019 10 h 00 min
- Protected: Romain's HDR rehearsal on 13 June 2019 10 h 30 min
- Séminaire RFAI (Yann Soullard, LITIS Rouen) on 19 March 2019 14 h 00 min
- RFAI Seminar (O. Augereau) on 13 March 2019 10 h 30 min
- RFAI Seminar (Xuan Son Nguyen) on 14 March 2019 14 h 00 min
- Defense Mostafa Darwiche (répétition) on 22 November 2018 10 h 00 min
- séminaire de Francesc Serratos on 16 May 2018 16 h 15 min
- Visite Vincent Poulain d'Andecy (Itesoft Yooz) on 17 April 2018 8 h 30 min
- Séminaire Hugo Raguet on 18 April 2018 10 h 30 min
- Séminaire Alasdair Newson on 11 April 2018 10 h 30 min
- Séminaire Xuan Son Nguyen on 10 April 2018 14 h 00 min
- Séminaire RFAI : Maxime DEVANNE (IMT Atlantique) on 27 March 2018 10 h 30 min
- Séminaire Olfa Ben Ahmed on 23 March 2018 10 h 30 min
- Séminaire Sébastien Bougleux on 15 March 2018 10 h 30 min
- Soutenance de Thèse Gaetan Galisot on 21 March 2018 14 h 00 min
- Gilles Verley : "De la perception à l'énonciation : application à la dégustation" on 11 January 2018 11 h 30 min
- Séminaire Sébastien Bougleux (GREYC) on 7 December 2017 15 h 00 min
- Séminaire David Boas on 14 December 2017 10 h 30 min
- Séminaire Maxime Martineau & Gaëtan Galisot on 16 November 2017 10 h 00 min
- Présentation de Mostafa Darwiche on 19 October 2017 10 h 30 min

- Presentations (Vanessa Haykal & Taibou Birgui Sekou) on 5 October 2017 10 h 00 min
- Soutenance annuelle Cong Nguyen on 25 October 2017 14 h 00 min
- RFAI Seminar: Julien Mille & Quoc Le & Mathieu Delalandre on 20 July 2017 10 h 00 min
- RFAI Seminar: Nicolas Ragot on 10 July 2017 14 h 00 min
- Séminaire G. Boccignone on 6 July 2017 14 h 00 min
- Journée du RTR Image on 27 June 2017 9 h 00 min
- Protected: RFAI Seminar (R. Sabourin, ETS Montreal) on 16 June 2017 10 h 30 min
- Soutenance de thèse de Arundhati TARAFDAR on 12 July 2017 14 h 00 min
- HDR Defense (N. Ragot) on 15 June 2017 14 h 00 min
- RFAI Seminar Yanna CRUZ CAVALCANTI (IRIT Toulouse) on 6 April 2017 10 h 30 min
- RFAI Seminar: Summary of AFIHM'17 on 30 March 2017 10 h 30 min
- RFAI seminar: Learning Strategies for Error-Tolerant Graph-Matching on 9 March 2017 10 h 30 min
- RFAI team meeting on 19 January 2017 10 h 30 min
- Séminaire RFAI D Conté on 13 December 2016 10 h 00 min
- Séminaire CADS D Conté on 8 December 2016 16 h 00 min
- Séminaire CADS M Blumenstein on 18 November 2016 14 h 00 min
- Séminaire CADS M Blumenstein on 17 November 2016 10 h 15 min
- Soutenance de thèse Frédéric Rayar on 22 November 2016 10 h 30 min
- Team meeting and Training for PhD (Julien Berhouet) on 29 September 2016 10 h 30 min
- PhD Defense Julien Berhouet on 3 October 2016 14 h 30 min
- RFAI Seminar Seiichi Uchida (Kyushu University / Japan) on 16 September 2016 14 h 00 min
- RFAI Seminar: Dr Rémi Giraud (LABRI) on 13 October 2016 10 h 30 min

## ANNEXE C : ANNEXE ROOT DU DAE

**Thèses soutenues de l'équipe ROOT :**

Docteur	Titre de la thèse	Financ.	Période	
<b>Z. ABU-AISHEH</b>	Approches distribuées pour la modélisation et l'indexation de graphes : application à la Reconnaissance des Formes	Allocation de recherche	Oct-13	Mai-16
<b>A. FROGER*</b>	<i>Planification des opérations d'entretien des équipements de production d'électricité sous incertitude</i>	Allocation de recherche	Oct-13	Déc-16
<b>A. MONTOYA*</b>	<i>Optimization of freight transportation with electric vehicles</i>	Allocation de recherche	Oct-13	Déc-16
<b>I. NDIAYE</b>	<i>Calcul de chemins pour la conception de plan d'évacuation de personnes</i>	ANR	Oct-12	Mars-16
<b>A. CHEREF</b>	<i>Méthodes de Recherche Opérationnelle pour la résolution de problèmes intégrés d'ordonnement de la production et de distribution</i>	ANR	Nov-13	Mars-17
<b>M. CISSE</b>	<i>Modèles mathématiques et algorithmes pour la résolution du problème de tournées du personnel de soins à domicile</i>	ANR	Oct-13	Juin-17
<b>A. Jlassi</b>	<i>Optimisation de la gestion des ressources sur une plate-forme de type « Big Data » basée sur le logiciel Hadoop</i>	CIFRE - Cyres	Fév-13	Déc-17
<b>L. SHANG</b>	<i>Exact algorithms with worst-case guarantee for scheduling: from theory to practice</i>	Allocation de recherche	Oct-14	Nov-17
<b>M. DARWICHE</b>	<i>When Operations Research meets Structural Pattern Recognition: on the solution of error-tolerant graph matching problems</i>	Allocation de recherche	Oct-15	Déc-18
<b>T.T.T. TA</b>	<i>New single machine scheduling problems with deadlines for the characterization of optimal solutions</i>	Bourse étrangère	Nov-14	Juil-18
<b>T. VAN UT</b>	<i>Approches de résolution de problèmes d'ordonnement multi-agent à machines parallèles identiques</i>	Bourse étrangère	Fév-14	Déc-18
<b>A. GIRET</b>	<i>Calcul des plus courts chemins multi-objectifs</i>	CIFRE - Géovélo	Jan-15	Abandon en 2018
<b>E. LOPEZ-AGUILAR</b>	<i>Outpatient Appointment Scheduling</i>	Bourse étrangère (Cotutelle)	Déc-15	Abandon en 2019
<b>H. CHEVROTON</b>	<i>Résolution par des algorithmes exacts et approchés de problèmes intégrés d'ordonnement de la production et de tournées de véhicules</i>	Allocation de recherche	Oct-17	Déc-20
<b>L. ECHEVERRI</b>	<i>The multi-period electric vehicle routing problem: model and solution approaches</i>	Projet ANR	Déc-16	Déc-20
<b>N. KULLMAN</b>	<i>Dynamic decision making under uncertainty in vehicle routing and logistics</i>	Projet ANR	Jan-17	Mai-20
<b>B. ZAHOUT</b>	<i>Algorithmes exacts et approchés pour l'ordonnement des travaux multiressources à intervalles fixes dans des systèmes distribués : approche monocritère et multiagent</i>	Allocation de recherche	Oct-16	Juin-21
<b>A. ROBBES</b>	<i>Résolution de problèmes intégrés dans un contexte pharmaceutique : ordonnancement et tournées de véhicules</i>	Projet Région	Avr-18	Oct-21

<b>V. JOURDAN</b>	<i>Ordonnancement de la maintenance corrective de rames de trains</i>	CIFRE - SNCF	Mars-20	Abandon en 2021
-------------------	---	--------------	---------	-----------------

\* Doctorants inscrits à l'école doctoral d'Angers.

### **Thèses en cours de l'équipe ROOT :**

Doctorant	Titre de la thèse	Financement	Date début	Soutenance prévue en
<b>M. HAOUASSI</b>	<i>Modèles et méthodes pour l'optimisation du processus de picking dans les entrepôts de e-commerce</i>	Région + Bourse étrangère (cotutelle)	Oct-16	2022
<b>M.L. AHMED SIDI</b>	<i>Nouvelles approches informatiques et mathématiques pour la résolution de problèmes biologiques</i>	Bourse étrangère (cotutelle)	Nov-18	2022
<b>M. BOUH HAROUNE</b>	<i>Ordonnancement multi-projets à contraintes de ressources multi-compétences et partagées</i>	Bourse étrangère (cotutelle)	Nov-18	2022
<b>J. WERLING*</b>	<i>Modélisation et simulation robuste des flux d'une unité de fabrication de semi-conducteurs avec prise en compte des ressources humaines</i>	CIFRE STMicroelectronics	Oct-17	2022
<b>O. PLOTON</b>	<i>Apports de l'algorithmique exponentielle à la résolution exacte ou approchée de problèmes d'ordonnancement</i>	Auto-financement (PRAG en Math.)	Oct-19	2023
<b>G. LACHARME</b>	Optimisation d'hyperparamètres de modèles d'apprentissage profond	Allocation de recherche	Oct-19	2023
<b>P. VENDE</b>	Outils d'aide à la décision pour la mise en place d'un réseau de transports en commun électriques	Région + Bourse étrangère (cotutelle)	Avr-20	2023
<b>A. CHANSON</b>	Optimisation et l'analyse interactive de données : le Problème du Voyageur de Données	Allocation de recherche	Oct-20	2023
<b>L. BEN ALI</b>	<i>Approches par recherche Opérationnelle et apprentissage pour l'optimisation des activités d'un site de distribution</i>	CIFRE SANOFI	Avr-20	2023
<b>S. MOULARD</b>	<i>Planification des parcours patients en ambulatoire</i>	Allocation de recherche	Sept-21	2024

\* Doctorant inscrit à l'école des Mines de Saint Etienne

## ANNEXE D : ANNEXE BDTLN DU DAE

Titres des thèses et HDR 2016-2022				
<b>Habilitation à Diriger de Recherches</b>				
NOM	Titre			date
SOULET Arnaud	Découverte de motifs centrée sur l'utilisateur			22/11/2019
<b>Thèses soutenues</b>				
NOM	Titre	financement	durée	devenir
ABID Ahmed	Techniques du Web sémantique et de la Fouille de données pour les architectures orientées services	étranger co-tut	4	entreprise
BEN MESMIA Fatma	Utilisation de cascades de transducteurs pour la recherche des entités nommées dans des textes arabes.	étranger co-dirF	5	EC Tunisie
BISONE Frédéric	Smart Ambulance : fouille de données hétérogènes issues de trajectoires de véhicules de secours	CIFRE	4,5	entreprise
BOULAKBECH Marwa	Mashup de services par configuration : application au domaine touristique	MIPTIS	4,5	EC en CDD
CHENIKI Nasredine	Découverte de Services Mobiles	étranger co-encadrE	4	postdoc
DIOP Lamine	Échantillonnage de motifs structurés sous contraintes	étranger co-dirE	2,5	ATER
DJEDAINI Mahfoud	Modèle, langage et benchmark pour l'analyse exploratoire	projet	3	postdoc
DRUSHKU Krista	Assistant personnel intelligent pour l'analyse visuelle de données d'entreprise	CIFRE	3	entreprise
EL MOUSSAWI Adnan	La segmentation et l'analyse interactive de données clients sensible à l'évolution de leurs caractéristiques dans le temps	CIFRE	4	data scientist Univ Tours
ELASHTER Mouna	Une plateforme d'extension du dictionnaire relationnel multilingue Prolexbase vers la langue arabe	étranger	4	
IVANOVICStefan	Mise à jour de référentiel géographique par fusion multi-source de données collaboratives	IGN co-dirF	3	postdoc
LABBACI Hemza	A Social Approach for Web Service Discovery and Community Detection	étranger co-encadrE	4	EC en CDD
MOREAU Clément	Fouille de trajectoires sémantiques	projet	3	entreprise
NGUYEN Long Phat	Extraction et caractérisation de motifs spatio-temporels pour l'analyse de séquences dans les vidéos	MIPTIS co-dirLIFAT	3	
NGUYEN Thanh Binh	Vers une exploitation intelligente des données du web : interrogation avec contraintes de l'utilisateur	MIPTIS co-dirF	4	EC Vietnam
PALMEIRA da SILVA Clay	Semantic-aware mashup of RESTful Web Services	étranger	3	postdoc
PASQUER Caroline	Garder la trace, mettre de l'ordre et relier les points : modéliser la variation et l'ambiguïté des expressions polylexicales	projet co-dirF	3	entreprise
WASZCZUK Jakub	Analyse syntaxique des unités polylexicales	MIPTIS co-dirF	3,5	EC Allemagne
<b>Thèses arrêtées par doctorant.e.s</b>				
NOM	Titre	financement	durée	devenir
BESSROUR Rayene	Découverte interactive de contrastes dans des données non-étiquetées	MIPTIS	1	entreprise
DJEDOUBOUM Assidé Christian	Interface entre interrogation du web sémantique et fouille de données volumineuses	étranger co-tut	1	entreprise
FOURRIER Léo	Génération automatique d'explications adaptées aux utilisateurs : évaluer et améliorer la confiance dans les modèles prédictifs et la connaissance du phénomène prédit	BRGM	1	entreprise
GUPTA Anubhav	Apports et exploitation de la recherche d'entités nommées dans des corpus scientifiques multidomains. Application au corpus Istex	projet	1	entreprise
LIRZIN Mathieu	Interopérabilité sémantique dans les systèmes multi-tiers	CIFRE	1	entreprise
SIDI ALY El Arby	Enrichissement d'ontologies à partir du web sémantique	étranger co-tut	1	entreprise

thèses en cours			
NOM	Titre	financement	durée (mars 2022)
BENALI Fodil	Fouille de données pour la construction et la recommandation de plan de placements publicitaires	CIFRE	1,5
BRAHEM Amina	Blockchain's fame reaches the execution of personalized touristic itineraries	projet co-tut	2,5
BRASSIER Maëlle	Résolution des coréférences pour la pseudonymisation de documents	CIFRE	0,5
BRES Raphaël	Evaluation intrinsèque de la qualité et de la fraîcheur des données géographiques collaboratives	IGN co-dirF	0,5
CANOVAS Nicolas	Génération automatique d'explications adaptées aux utilisateurs : évaluer et améliorer la confiance dans les modèles prédictifs et la connaissance du phénomène prédit	BRGM	0,5
CHANSON Alexandre	TAP : problème du voyageur de données	MIPTIS co-dirLIFAT	1,5
CRULIS Ben	Assistant Intelligent pour l'aide aux déplacements des aveugles	MIPTIS co-dirLIFAT	0,5
EL OUTA Faten	data stories for interactive intentional analytics	MIPTIS	2,5
HADJ MOHAMED Najet Dhouha	Identification d'expressions polylexicales en arabe	étranger co-tut	0,5
JOUET Gregor	Apprentissage profond: Interactions contextualisées	U. LdV	1,5
KASTNER Thomas	Détection de motifs atypiques dans des masses de données par des méthodes de fouille de données et d'apprentissage profond	CIFRE co-dirLIFAT	0
KOMEIHA Fouad	Java library for Linked Open Data Based Similarity Measures	projet co-tut	2,5
LION-BOUTON Adam	Conception et induction automatique d'un lexique syntaxico-sémantique d'expressions polylexicales au service de la diversité linguistique	MIPTIS co-dirE	1,5
ONDZIGUE MBENGA Raymond	Etude et mise en place d'un système national d'information géographique décisionnel utilisant les technologie mobiles pour la surveillance intégrée de la tuberculose dans les pays de la zone CEMAC	étranger co-tut	2,5
RINGUET Nicolas	Système de recommandation interactif de parcours de vie	CIFRE	1,5
SERRA Flavia	Context-based Data Quality Management for Heterogeneous Information Systems	étranger co-tut	1,5
SONI Praveen	Génération automatique des tableaux de bord des projets BI4PEOPLE	projet co-dirLIFAT	0,5
VERDEAUX Willeme	modèles explicables et analyse d'hypothèse pour la recommandation multi-parties prenantes	CIFRE	2,5

## Portfolio LIFAT

Plusieurs faits marquants ont jalonné la vie du LIFAT depuis 2016. Nous les présentons en distinguant ceux qui concernent tout le laboratoire et ceux qui sont plus spécifiques aux équipes de recherche.

### Fait marquant au niveau LIFAT : le CETU ILIAD3

Le LIFAT est fier d'avoir largement contribué à la création du [CETU ILIAD3](#) en 2017. Un Centre d'Expertise et de Transfert Universitaire est une composante de l'université de Tours (comparable à une spin-off intégrée) dédiée à la valorisation de la Recherche dans un domaine particulier et **devant s'autofinancer** en offrant des services de prestation et conseil R&D. Le CETU ILIAD3 est spécialisé dans le domaine de l'ingénierie logicielle en traitement d'images et numérisation 3D.

Composé d'ingénieurs et ingénieurs-chercheurs, il propose des services pour le compte de laboratoires de recherche, d'organismes privés ou publics, de PME ou de grands groupes dans le cadre de collaborations public/public, privé/public ou de prestations de services incluant des développements technologiques spécifiques et innovants.

Les réalisations suivantes sont, par exemple, à mettre au crédit du CETU ILIAD3 depuis sa création :

- Projet VisiT : développement d'une application mobile accessible au plus grand nombre permettant la découverte interactive d'œuvres muséales grâce à la réalité augmentée. (Partenaires : Région CVL, Département Indre et Loire, Château de Loches, Château d'Azay-Le-Rideau/CMN) – 2019-2021
- Prestations pour « Chambord 2019, 500 ans de la Renaissance » (exposition saison été 2019) concernant des dispositifs de médiation culturelle, conception et réalisation puis mise en service au cours de la saison touristique (Partenaires : Région Centre Val de Loire, Domaine National de Chambord) – 2018-2019
- Projet « Sculpture 3D », Numérisation 3D de Sculptures de la Région CVL, conception et réalisation de dispositifs d'interaction pour la médiation de contenu historique pour les visiteurs du musée de Beaux-Arts de Tours, Exposition Sculpturoscope – 2016-2018. (Partenaires : Région Centre Val de Loire, DRAC Centre Val de Loire, Centre des Musées Nationaux, Ville de Tours)
- Prestations aux côtés de chercheurs de l'INSERM 1253, pour la mise en œuvre de méthodes de traitement et d'analyses d'images, mise en œuvre de techniques d'IA pour la segmentation automatique de structures cérébrales. 2018-2019-2020 (Partenaires : INSERM U1253, INRAE PRC, CHU Tours)
- Prestation pour le compte du CEA Monts, sur la caractérisation de matériaux grâce à des algorithmes d'IA. – 2021 (Partenaire CEA)

La création du CETU est une réussite ayant produit un fort impact socio-économique local démontré par les figures ci-dessous. On note par exemple que le CETU ILIAD3 est générateur de plus de **2,5 emplois temps plein depuis 2019**.

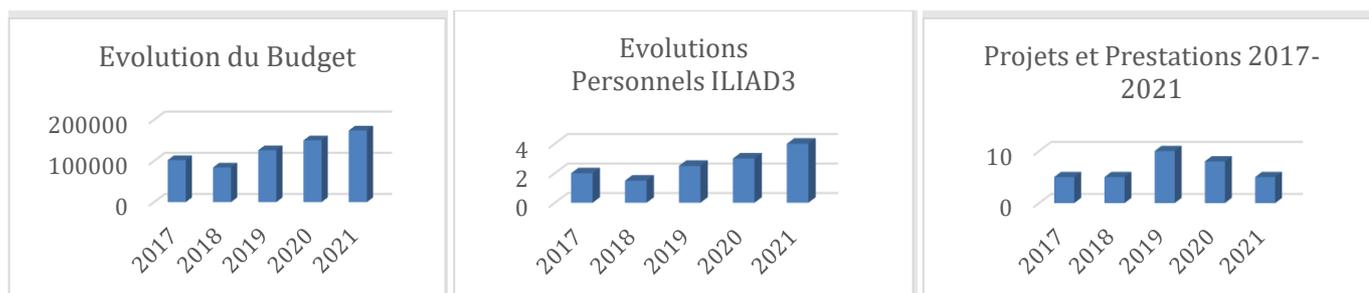


Figure 1 : Impacts socio-économiques de ILIAD3

Matériels disponibles au CETU ILIAD3 :

- Stations de calcul graphique GPU
- Scanner Laser 3D Faro
- Scanners Optiques 3D GoScan/Faro
- Dispositifs d'interactions (Kinect, LeapMotion, etc...) et de visualisation 3D (hologramme)

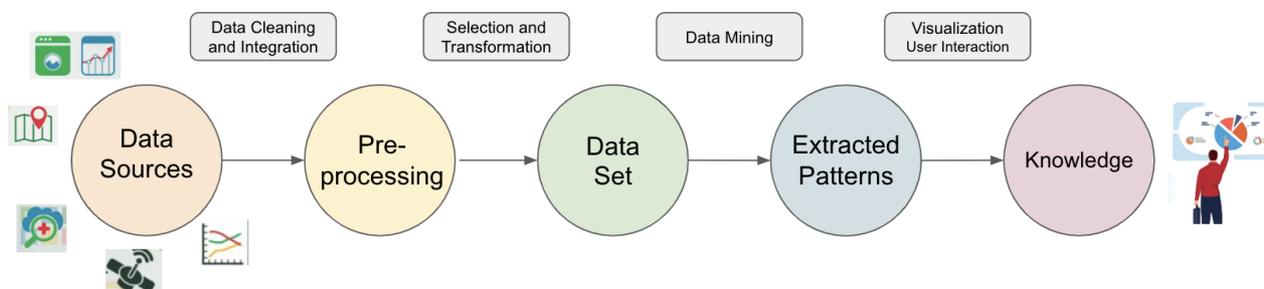
### Fiche LIFAT : Interaction utilisateur et explicabilité

Le LIFAT possède historiquement un positionnement fort sur les méthodes orientées utilisateurs selon deux axes principaux :

- Un axe qui vise à la définition de **méthodes de fouilles de données ou d'apprentissage automatique interactives** dans lesquelles l'utilisateur peut guider la convergence de la méthode vers une solution plus en phase avec ses préférences ou ses besoins. Des travaux ont notamment été conduits sur des méthodes de clustering semi-supervisée [1], ou dans le développement de méthodes d'extraction de motifs [2]. Des méthodes d'élicitation des intentions utilisateurs ont été proposées dans le cadre de la thèse CIFRE de Krista Drushku avec la société SAP pour personnaliser la recommandation d'objets composites dans le contexte de l'intelligence décisionnelle [3]. Enfin, d'autres travaux, plus récents, initiés dans le cadre de la thèse CIFRE de Nicolas Ringuet avec la société Neolink, s'intéressent à la modélisation multidimensionnelle et multi-granulaire du mécanisme de co-construction de trajectoires de vie pour l'assistance au retour à l'emploi ou le maintien à domicile des personnes âgées [4].
- Un axe dans lequel le focus est mis sur le développement d'outils interactifs adaptés à des utilisateurs exploitant des techniques d'intelligence artificielle. En particulier, le LIFAT a développé, dans le cadre d'une collaboration au long cours avec des centres de rééducation (Kerpape, PFNT Hôpital de Garches), un système d'aide à la communication destiné à des personnes souffrant de déficiences motrices sévères (voir la fiche du portfolio "**Ressources TAL et aide à la communication pour personnes handicapées**"). Le LIFAT a également travaillé à la mise en place d'un outil de réalité augmentée pour aider à la chirurgie de l'épaule : aide au positionnement d'une prothèse dans les cas d'atteinte de la glaine (CIFRE IMASCAP) [6]. Également dans le cadre de plusieurs projets successifs avec l'INRAE et l'INSERM (Neurogeo, Neuro2co, ...) le LIFAT a mis au point des outils de **segmentation interactive pour l'analyse d'images médicales 3D** [7]. Une plateforme logicielle disponible en ligne ([SILA3D](#)) résultant de ces travaux est exploitée par des neuroanatomistes pour l'analyse de cerveaux humains et de divers espèces animales (cf. portfolio RFAI).

Si les méthodes précédentes permettent à un système intelligent de mieux interagir ou de mieux cerner les besoins d'un utilisateur, elles ne permettent pas à un utilisateur de nécessairement mieux comprendre pourquoi une analyse fournit les résultats qui sont les siens. Cette thématique, liée à l'interprétabilité et l'explicabilité, aujourd'hui de premier plan dans la communauté nationale et internationale, a été étudiée au LIFAT depuis plusieurs années par le biais de thèses, d'animation de groupes de recherche CNRS ou de projets :

- Le LIFAT est impliqué dans **deux thèses** actuellement : la thèse CIFRE de Willème Verdeaux avec la société Kalidea-Up qui s'intéresse à l'explicabilité des systèmes de recommandation. L'approche proposée étend la méthode classique de l'état de l'art LIME, en la portant au domaine de l'explication de recommandations et en redéfinissant la notion de localité autour d'une instance à expliquer. La méthode innove aussi en proposant d'apprendre une explication sur la base de l'ordre des items d'un utilisateur plutôt que sur leurs évaluations [5]. Enfin, la thèse de Nicolas Canovas, débutée récemment en septembre 2021, est conduite en collaboration avec le BRGM (Bureau de Recherche et de Géologie Minière) à Orléans et s'intéresse aux aspects liés à la confiance dans les modèles prédictifs.
- Le LIFAT est impliqué dans l'animation de deux groupes de travail sur l'explicabilité. Notamment, le LIFAT a initié et animé l'atelier **Fender du GdR CNRS Madics** en collaboration avec un collègue du laboratoire I3 Telecom ParisTech, apportant un regard pluridisciplinaire entre informatique et SHS sur la thématique de l'explicabilité. L'action HELP du GdR Madics fait suite à l'atelier Fender, co-portée par le LIFAT, TelecomParisTech et l'IRIT, s'intéresse à la question de l'explicabilité de bout-en-bout d'un système d'analyse de données.



Contrairement à la majeure partie des travaux dans le domaine qui ne se focalisent que sur l'explication des modèles de prédiction, l'idée ici est de pouvoir expliquer l'impact de chaque choix dans la chaîne de traitement des données et de s'intéresser à la relation entre l'explication et l'utilisateur qui la reçoit. Le LIFAT est également impliqué dans l'organisation et la direction de l'**atelier EGC Explain AI**, dont la première édition a eu lieu à la conférence EGC'2022. Cet atelier est devenu un groupe de travail EGC et est désormais associé à l'action HELP du GdR Madics.

- Le LIFAT participe également à l'ARD **Junon** sur l'établissement de jumeaux numériques, pour lesquels les questions d'explicabilité vont être mobilisées pour valider les propositions de jumeaux et s'assurer que l'on peut avoir confiance dans leur utilisation. Enfin, l'explicabilité sera également au cœur des méthodes de prédiction qui seront mises en œuvre dans

le projet régional OPTIMEDIAS déposé à la région en collaboration avec l'INSERM et visant à diagnostiquer et pronostiquer au plus tôt certaines pathologies (comme les scléroses latérales amyotrophiques par exemple).

### Références

- [1] Adnan El Moussawi. Clustering exploratoire pour la segmentation de données clients. Thèse de doctorat du LIFAT, soutenue le 25/09/2018
- [2] Arnaud Soulet. Découverte de motifs centrée sur l'utilisateur. (User-Centric Pattern Discovery). Université de Tours, France, 2019
- [3] Krista Drushku, Julien Aligon, Nicolas Labroche, Patrick Marcel, Verónica Peralta. Interest-based recommendations for business intelligence users. *Inf. Syst.* 86: 79-93, 2019
- [4] Nicolas Ringuet. Challenges in Lifelong Pathways Recommendation. *ADBIS*: 310-316, 2021
- [5] Léo Brunot, Nicolas Canovas, Alexandre Chanson, Nicolas Labroche, Willème Verdeaux. Preference-based and local post-hoc explanations for recommender systems. *Information Systems* 102021, 2022
- [6] David Boas. Réalité augmentée pour la chirurgie de l'épaule, thèse de doctorat du LIFAT soutenue le 6 juillet 2020
- [7] Gaëtan Galisot, Jean-Yves Ramel, Thierry Brouard, Elodie Chaillou, Barthélémy Serres, Visual and structural feature combination in an interactive machine learning system for medical image segmentation. *Machine Learning with Applications*. Volume 8, 15 June 2022

## Portfolio BdTIn

Pour illustrer le positionnement scientifique de l'équipe BdTIn, ses nombreuses collaborations et sa reconnaissance au niveau national et international, nous présentons trois articles et quatre fiches thématiques.

Le premier article, dans l'inter-axe *Extraction de connaissances du Web des données*, a fait partie de la sélection des trois meilleurs articles à la conférence ISWC 2018, l'une des plus reconnues internationalement dans le domaine du Web sémantique. Il illustre donc la qualité des productions dans ce thème et la reconnaissance de son intérêt par la communauté internationale. De plus il est un fruit d'une collaboration avec Télécom Paris établie par Arnaud Soulet lors de son CRCT en préparation de son HDR. Le deuxième article, orienté *recommandation dans l'analyse interactive de données*, publié dans une revue reconnue également dans la communauté, montre le dynamisme de l'équipe en matière de collaborations avec des entreprises via des thèses CIFRE, ici avec la grande entreprise SAP. Cet article a été accompagné d'un dépôt conjoint de brevet. Le troisième article, publié dans la conférence de référence COLING en 2020, représente le thème *détection d'entités d'intérêt* et en l'occurrence détection d'entités polylexicales, objets de l'action COST PARSEME et du projet ANR PARSEME-FR, qui furent d'une importance prépondérante pour l'équipe au cours de la période 2016-2021 dans son rayonnement tant national qu'international. Dans les auteurs nous soulignons les noms des doctorantes.

Les quatre fiches thématiques quant à elles mettent en valeur la cohérence des démarches menées dans l'équipe BdTIn, que ce soit en termes scientifiques ou de production de ressources dans le cadre de collaborations. Ces démarches cohérentes sur le long terme sont nombreuses à s'incarner au sein des axes, entre eux, entre équipes du LIFAT et au-delà. Celles que nous avons sélectionnées ici concernent des travaux autour (1) de la *narration de données*, (2) des *ressources pour le TAL et l'aide aux personnes porteuses de handicaps*, (3) de l'*extraction interactive de motifs* depuis des données, qui maintient l'expert humain dans la boucle, et enfin (4) de la notion de *similarité sémantique, pour la découverte, la fouille et la composition d'objets complexes*. L'équipe les a choisis parmi d'autres parce qu'ils illustrent des points forts complémentaires à ceux des trois articles, également pour la qualité de leurs résultats et pour leur « couverture interne », chaque fiche rassemblant des contributions de plusieurs membres de l'équipe.

### Élément 1 : article nominé à ISWC 2018

**Arnaud Soulet, Arnaud Giacometti, Béatrice Markhoff, Fabian M. Suchanek : *Representativeness of Knowledge Bases with the Generalized Benford's Law*. ISWC (1) 2018: 374-390 - hal-01824490**

Le verrou visé dans cet article est celui des mesures de la qualité des données dans les grands graphes de connaissances (GC) du Web, comme par exemple DBpedia. Il est connu que ces connaissances sont incomplètes par rapport à la réalité, mais c'est plus leur manque de représentativité qui peut limiter sérieusement la portée des analyses qui s'appuieraient sur elles. Cet article propose une méthode pour mesurer la représentativité d'une propriété (par ex. la taille de la population d'un lieu) dans un GC. Elle consiste à comparer la distribution de la propriété dans le GC avec sa *distribution réelle approximée par la loi de Benford généralisée*. Elle permet d'approximer le nombre de faits minimum à ajouter dans le GC pour qu'il soit représentatif de la réalité. Les résultats expérimentaux montrent que la méthode passe à l'échelle de DBpedia (à qui il manquait en 2018 de l'ordre de 45 M de faits pour être représentatif).

### Élément 2 : article dans Information Systems 86 (2019)

**Krista Drushku, Julien Aligon, Nicolas Labroche, Patrick Marcel, Verónica Peralta : *Interest-based recommendations for business intelligence users*. Information Systems 86: 79-93 (2019) - hal-02976606**

La proposition de cet article est de détecter et caractériser les intérêts d'un analyste de données au cours d'une session d'analyse afin de lui recommander des requêtes, de façon collaborative. Nous définissons et caractérisons l'intention des interactions de l'utilisateur avec le système, ainsi qu'une mesure de similarité pour grouper ces intentions en clusters. Le système de recommandation qui en découle s'appuie sur un modèle de Markov représentant la probabilité pour un utilisateur de passer d'un intérêt à un autre. Il a été validé par une étude utilisateur approfondie, montrant d'une part une meilleure précision de la mesure de similarité vis-à-vis des intérêts des utilisateurs, par rapport à des mesures de similarité de requêtes de l'état de l'art et, d'autre part, un plus grand intérêt des utilisateurs pour ce système par rapport à d'autres systèmes de recommandation de l'état de l'art.

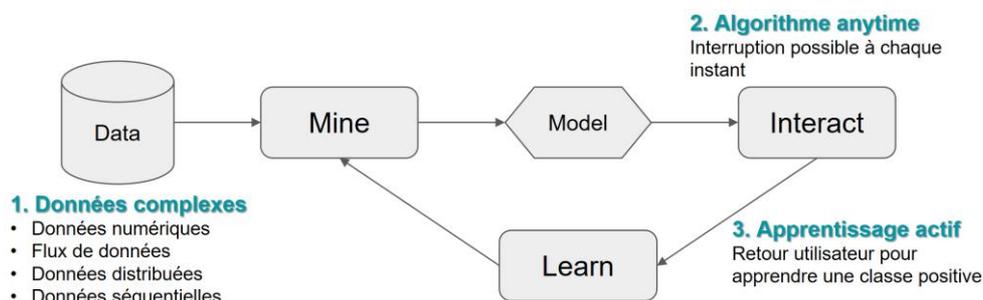
### Élément 3 : article à COLING 2020

**Caroline Pasquer, Agata Savary, Carlos Ramisch, Jean-Yves Antoine : *Verbal Multiword Expression Identification: Do We Need a Sledgehammer to Crack a Nut?* COLING 2020: 3333-3345) - hal-03013636**

Ce travail est représentatif de la démarche de tous nos travaux menés en TAL à savoir conduire des recherches conscientes du fait linguistique de manière agnostique, i.e. sans être liés à une approche (symbolique, apprentissage statistique, apprentissage profond) a priori, mais en adoptant au contraire celle qui est la plus appropriée à la tâche et aux données considérées. Dans le cas présent, nous avons ainsi développé une approche originale composée de la combinaison de 8 filtres linguistiques utilisés de manière variable suivant la langue. Cette démarche "old school" et très peu coûteuse d'un point de vue computationnel a obtenu les meilleurs résultats de l'édition 2020 de la Share Task Parseme sur plusieurs langues, et est restée au niveau de l'état de l'art sur les autres.

## Élément 4 : Fiche “Découverte interactive de motifs”

La fouille de motifs au sein de BDTLN vise des systèmes interactifs où le retour de l'utilisateur sur les motifs est continuellement intégré au processus de fouille pour améliorer l'analyse (étape « Interact » dans la figure). Ces travaux ont été initiés par des réflexions menées dans le cadre de projets PEPS PREFUTE (PEPS 2015, 2016 CNRS) pour Fouille interactive fondée sur les PRÉférences UTILISatEUR (partenaires : GREYC (Caen), LIRIS (Lyon) et LORIA (Nancy)).



*Processus de fouille interactif (inspiré de van Leeuwen 2014) avec nos contributions*

L'originalité de nos travaux est d'exploiter les méthodes d'échantillonnage de motifs introduites par Boley et al., 2011. Au lieu d'extraire tous les motifs, il s'agit de tirer aléatoirement un échantillon de motifs (parmi ceux qui auraient été extraits avec une méthode classique) afin de fournir instantanément des motifs à l'utilisateur. Ce changement de paradigme relâche la complétude de l'extraction de motifs au profit d'une plus grande rapidité et de garanties statistiques. Ce thème, objet d'une partie de l'**HDR d'Arnaud Soulet** soutenue en 2019, a donné lieu à plusieurs articles dans les **meilleures conférences** du domaine. L'ensemble des développements est **disponible librement** (voir ci-après).

- 1. Données complexes** : Nous avons montré comment traiter des données complexes où les méthodes traditionnelles étaient inopérantes. Il est, par exemple, possible de traiter des données continues [SDM18] ou des flux de données sans oubli [PKDD21]. La **thèse de Lamine Diop**, soutenue en 2021 et **primée par l'association EGC**, a permis de considérer des données séquentielles [ICDM18] et des données distribuées [ADBIS20].
- 2. Algorithme anytime** : Nous avons montré comment utiliser l'échantillonnage de motifs pour concevoir des algorithmes anytime de détection de données aberrantes [PAKDD16]. Ce travail a été **distingué** par les conférences EGC 2016 et PAKDD 2016.
- 3. Apprentissage actif** : Nous avons montré comment intégrer les retours de l'utilisateur sur les motifs extraits pour caractériser un jeu de données non-étiqueté [IDA2016].

Ces travaux sont discutés au sein du Réseau Thématique Régional DIAMS dans le Groupe de Travail (co-animé par BDTLN avec des collègues du Laboratoire d'Informatique Fondamentale d'Orléans) « Human in the loop for machine learning and data mining ».

### Références

[PKDD21] Giacometti, A., & Soulet, A. (2021, September). *Reservoir Pattern Sampling in Data Streams*. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases (pp. 337-352). Springer, Cham.

[ADBIS20] Diop, L., Diop, C. T., Giacometti, A., & Soulet, A. (2020, August). *Pattern sampling in distributed databases*. In European Conference on Advances in Databases and Information Systems (pp. 60-74). Springer, Cham.

[ICDM18] Diop, L., Diop, C. T., Giacometti, A., Li, D., & Soulet, A. (2018, November). *Sequential pattern sampling with norm constraints*. In 2018 IEEE International Conference on Data Mining (ICDM) (pp. 89-98). IEEE.

[SDM18] Giacometti, A., & Soulet, A. (2018, May). *Dense neighborhood pattern sampling in numerical data*. In Proceedings of the 2018 SIAM International Conference on Data Mining (pp. 756-764). Society for Industrial and Applied Mathematics.

[IDA17] Giacometti, A., & Soulet, A. (2017, October). *Interactive pattern sampling for characterizing unlabeled data*. In International Symposium on Intelligent Data Analysis (pp. 99-111). Springer, Cham.

[PAKDD16] Giacometti, A., & Soulet, A. (2016, April). *Frequent pattern outlier detection without exhaustive mining*. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 196-207). Springer, Cham.

### GitHub

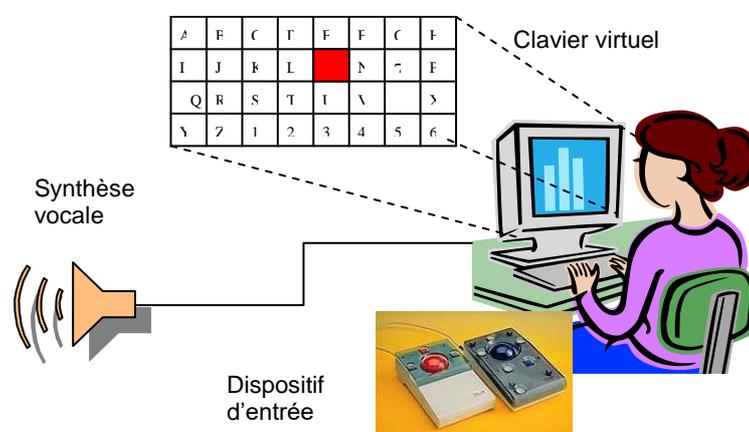
- <https://github.com/LDIOPBSF>
- <https://github.com/asoulet>

## Élément 5 : Fiche “Ressources TAL et aide à la communication pour personnes handicapées”

L'équipe BdTln contribue à la production de nombreuses ressources pour le TAL, des outils et des corpus. Une partie de ces ressources est accessible via le site web <https://tln.lifat.univ-tours.fr/>, une autre, produite dans le cadre de PARSEME, est ici : <http://hdl.handle.net/11234/1-3367> et là : <https://parsemefr.lis-lab.fr/doku.php?id=outcomes>.

L'équipe s'investit également dans la production de ressources pour l'aide à la communication (AAC en anglais, pour Alternative and Augmentative Communication). L'objectif est de restaurer les capacités de communication de personnes qui souffrent d'un handicap moteur sévère. Leurs possibilités de communication à l'écrit (écriture manuscrite ou saisie de texte sur clavier d'ordinateur) sont excessivement réduites. Par ailleurs, l'insuffisance du contrôle moteur peut concerner l'appareil phonatoire. Dans ce cas, la communication est également privée de son support oral habituel. Les systèmes AAC doivent proposer une aide à la saisie de message (et éventuellement leur vocalisation) adaptée à leurs capacités très réduites d'action sur le monde physique. Cette saisie va s'effectuer sur un tableau virtuel de symboles (mots, lettres, phonèmes voire icônes) qui est affiché sur l'écran d'un ordinateur. Le message est construit en sélectionnant successivement sur le clavier virtuel les symboles qui le composent. Ces systèmes reposent sur trois composants principaux :

1. Un dispositif physique (joystick microgravité, commande oculaire, licorne, interface cerveau-machine... ou simple bouton poussoir) qui joue le rôle de périphérique d'entrée et est adapté au tableau clinique de l'utilisateur.



2. Le clavier virtuel, qui peut être associé à un éditeur intégré ou permettre le pilotage d'une application externe.
3. Enfin, une synthèse de parole à partir du texte.

Pour pallier la lenteur et la pénibilité de la composition des messages, il est nécessaire d'optimiser la saisie et c'est là qu'interviennent nos recherches. Deux approches complémentaires sont envisageables :

- la première consiste à optimiser le temps d'accès à la touche recherchée, donc à réduire le temps de défilement pour y accéder. Cette optimisation dépend de l'organisation générale du clavier (IHM),

- la seconde approche vise à minimiser le nombre de saisies, en complétant automatiquement certaines parties du message en fonction du contexte. C'est ici qu'intervient nos recherches (TAL) en prédiction lexicale.

L'équipe BDTLN est reconnue depuis plus de 20 ans pour ses travaux dans ce domaine, qui se caractérisent par leur pluridisciplinarité puisqu'ils associent le TALN, l'IHM mais aussi l'orthophonie et l'ergothérapie. Nos recherches sont menées dans le cadre d'une collaboration au long court avec le **centre de rééducation de Kerpape** et l'**hôpital de Garches**. Nous avons développé les systèmes d'aide à la communication Sibylle et désormais LifeCompanion qui ont été adoptés par plusieurs centaines de personnes handicapées. Ces systèmes sont hautement adaptables à l'utilisateur et au thème courant du discours, ce qui requiert une modélisation du contexte dans un espace sémantique latent (LSA), représentation distribuée équivalente aux techniques de plongements neuronaux. L'équipe s'est intéressée (**projet Predict4All**) à la combinaison de la prédiction avec une correction orthographique en cours de saisie, ce qui est essentiel pour les patients souffrant de troubles langagiers associés (troubles DYS) et pourtant négligé à l'heure actuelle.

**L'équipe pilote un projet ANR (AAC4All)** commençant en 2022, dont l'ambition est de réunir les acteurs francophones en AAC (LIFAT, LIG, IRIT, MoDyCo, CRMFF Kerpape, Garches) pour réaliser une plateforme ouverte qui permettra de mutualiser les efforts et de fournir une aide à la recommandation, aux patients mais aussi aux aidants et soignants.

**Vidéo de présentation de nos systèmes en utilisation dans les centres de rééducation :**

[https://www.youtube.com/watch?v=Jb3hHolgTkE&feature=emb\\_title](https://www.youtube.com/watch?v=Jb3hHolgTkE&feature=emb_title)

## Élément 6 : Fiche “TAP (Traveling Analyst Problem)”

Une narration de données peut être définie comme la composition structurée de messages transmettant des “trouvailles” faites dans des données, communiquées le plus souvent visuellement pour faciliter leur réception. Ce type d’activité est omniprésente lorsqu’il s’agit de restituer des analyses exploratoires de données, comme en datajournalisme. L’étude de la narration de données est relativement récente et a attiré l’attention de différentes communautés de recherche en informatique (gestion de la donnée, visualisation, recherche d’information). Cette étude est permise notamment par l’exploitation de corpus riches (articles de presse, notebooks). L’équipe BDTLN aborde cette étude en ayant défini le problème dit du « voyageur de données » (TAP, pour Traveling Analyst Problem en anglais) [1]. Ce problème multi-objectifs, fortement NP-difficile, formalise la génération automatique sous contrainte de séquences d’actions d’analyse.

*Definition 4.1 (Traveling Analyst Problem (TAP)).* Let  $Q$  be a set of  $N$  queries, each associated with a positive time cost  $cost(q_i)$  and a positive interestingness score  $interest(q_i)$ . Each pair of queries is associated with a metric  $dist(q_i, q_j)$  representing the cognitive distance of browsing from one query result to the next. Given a time budget  $\epsilon_t$ , the optimization problem consists in finding a sequence  $\langle q_1, \dots, q_M \rangle$  of queries,  $q_i \in Q$ , without repetition, with  $M \leq N$ , such that:

- (1)  $\max \sum_{i=1}^M interest(q_i)$
- (2)  $\sum_{i=1}^M cost(q_i) \leq \epsilon_t$
- (3)  $\min \sum_{i=1}^{M-1} dist(q_i, q_{i+1})$ .

### Définition formelle du problème du voyageur de données [1]

L’étude théorique et des applications de ce problème sont au cœur de **trois thèses et d’une action du GDR Madics** :

1. **Étude théorique** : l’étude théorique du problème TAP se fait dans le cadre de la **thèse d’Alexandre Chanson**, en collaboration avec l’équipe ROOT du LIFAT. Cette thèse est l’occasion de collaboration avec l’Italie (université de Bologne) et a débouché notamment sur une approche de génération de notebooks de comparaisons [3].
2. **Application à la recommandation de parcours de vie** : la **thèse CIFRE de Nicolas Ringuet** porte sur la recommandation de parcours de vie et a une forte dimension sociétale. Le TAP a été utilisé comme première approche pour l’implémentation d’un système de recommandations de parcours de vie [2].
3. **Application à la construction de narrations de données** : l’équipe BdTln porte l’**action Madona du GDR Madics** (Maîtriser l’Analyse de Données pour la Narration journalistique). Cette action vise à rassembler chercheurs en informatique, chercheurs en sciences de l’information et de la communication et datajournalistes autour de la documentation et de la modélisation des processus d’analyse des datajournalistes. Les travaux de **thèse de Faten El Outa** s’inscrivent dans cette action. Ils sont menés en collaboration avec l’université de Ioannina (Grèce) et ont donné lieu à une approche de génération automatique de narrations personnalisées à partir de notebooks [4], également modélisée comme une instance du TAP.

### Références

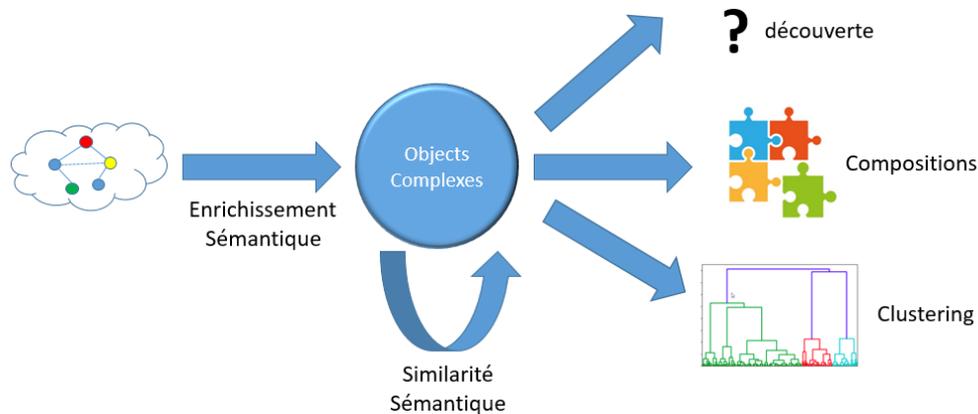
- [1] Alexandre Chanson, Ben Crulis, Nicolas Labroche, Patrick Marcel, Verónica Peralta, Stefano Rizzi, and Panos Vassiliadis. The traveling analyst problem: Definition and preliminary study. *Proceedings of DOLAP@EDBT/ICDT 2020*, pages 94–98. CEUR-WS.org, 2020.
- [2] Alexandre Chanson, Thomas Devogele, Nicolas Labroche, Patrick Marcel, Nicolas Ringuet, and Vincent T’kindt. A chain composite item recommender for lifelong pathways. *DaWaK 2021*, volume 12925 of *Lecture Notes in Computer Science*, pages 55–66. Springer, 2021.
- [3] Alexandre Chanson, Nicolas Labroche, Patrick Marcel, Stefano Rizzi, and Vincent T’kindt. Automatic generation of comparison notebooks for interactive data exploration. *Proceedings of EDBT 2022*, pages 2:274–2:284. OpenProceedings.org, 2022.
- [4] Alexandre Chanson, Faten El Outa, Nicolas Labroche, Patrick Marcel, Verónica Peralta, Willeme Verdeaux, Lucile Jacquemart: Generating Personalized Data Narrations from EDA Notebooks. *DOLAP 2022*: 91-95

### GitHub

- <https://github.com/patrickmarcel/sqlEDAqueryGenerator>
- <https://github.com/Blobfish-LIFAT/NotebookCrowdsourcing>

## Élément 7 : Fiche “Similarité sémantique pour la découverte, la fouille et la composition d’objets complexes”

Le nombre d’**objets complexes** (documents, activités, séquence d’activités, services, Things, etc.) disponibles et partagés est en constante augmentation. Afin de faciliter et automatiser le traitement (découverte, fouille, composition, etc.) de ces volumes importants d’informations, les connaissances (ontologies, données liées, etc.) et les techniques qui leur sont associées (raisonnement sur les concepts et relations, mesures de similarité, algorithmes de fouille) sont indispensables pour le **calcul de similarité sémantique** entre objets sémantiquement enrichis/annotés.



L’équipe BdTin a contribué ces dernières années à ce domaine en faisant soutenir plusieurs thèses en lien avec des projets concrets. Ainsi, des solutions ont été définies et implémentées en exploitant les relations conceptuelles contenues dans les ontologies de domaine pour la découverte de services Web dans la **thèse de Ahmed Abid** (en cotutelle avec l’Université de Sfax), pour les séquences d’activités journalières dans la **thèse de Clément Moreau** (projet ANR Mobi’Kids), ou en exploitant les ressources du LOD pour les services Web mobiles dans la **thèse de Nasreddine Cheniki** (en co-encadrement avec l’Université USTHB d’Alger) et celle de **Fouad Komeiha** (en cotutelle avec l’Université Libanaise), pour inclure tous types d’objets connectés dans la recommandation et répondre ainsi au besoin en services et objets connectés dans des contextes domotiques (Web des objets). Par ailleurs, les mesures sémantiques proposées dans l’équipe ont été couplées à des techniques d’apprentissage pour la gestion des interactions dynamiques entre services dans la **thèse de Hemza Labbaci** (en coencadrement avec l’USTH d’Alger). Enfin, la **thèse de Marwa Boulakbech** (projet régional SmartLoire) a proposé d’enrichir l’ontologie [DATATourisme](#) et de la coupler avec des algorithmes de configuration de mashups de services. Ce travail est en cours d’extension pour inclure l’aspect respect de données personnelles dans le cadre de sélection et composition de services, dans la **thèse d’Amina Brahem** (en cotutelle avec l’Université de Tunis).

Ces travaux ont été mis en œuvre en grande partie dans le cadre de deux projets : **Smart Loire** (APR Centre 2017-2021) et **Mobi’kids** (ANR), donnant lieu à des déclinaisons respectives dans les domaines du tourisme et de la mobilité urbaine. Des déclinaisons dans le domaine de la santé sont également en cours de développement dans le cadre du projet **SQVALD** (APR Centre 2022-2025).

Les nombreuses publications associées à ces thèses sont listées dans le fichier des productions scientifiques du LIFAT (équipe 1), cette fiche visant surtout à montrer comment elles s’articulent en un tout cohérent. Citons tout de même :

[1] Ahmed Abid, Mohsen Rouached, Nizar Messai: Semantic web service composition using semantic similarity measures and formal concept analysis. *Multim. Tools Appl.* 79(9-10): 6569-6597 (2020).

[2] Hamza Labbaci, Nasredine Cheniki, Yacine Sam, Nizar Messai, Brahim Medjahed, Youcef Aklouf: A Linked Open Data Approach for Web Service Evolution. *OTM Conferences 2019*: 265-281

[3] Nasredine Cheniki, Yacine Sam, Nizar Messai, Abdelkader Belkhir: Context-Aware and Linked Open Data Based Service Discovery. *ICWE 2018*: 448-462

[4] Clément Moreau, Thomas Devogele, Cyril de Runz, Verónica Peralta, V., Evelyne Moreau & Laurent Etienne: A Fuzzy Generalisation of the Hamming Distance for Temporal Sequences. In 2021 IEEE International Conference on Fuzzy Systems FUZZ-IEEE 2021 pp. 1-8

### Outils développés et sites Web de présentation

- Git Thèse de Clément Moreau : <https://github.com/Clement-Moreau-Info>
- Git Thèse de Fouad Komeiha : <https://github.com/FouadKom/lds>
- Projet Smart Loire : <https://smartloire.univ-tours.fr/>
- Projet ANR Mobi’kids : <https://www.pacte-grenoble.fr/programmes/mobi-kids>

## Portfolio RFAI

Dans ce portfolio, pour mettre en avant les activités de l'équipe, nous proposons 3 fiches et 3 publications qui illustrent différentes spécificités de l'équipe, que ce soit en termes d'activité soutenue dans le temps, en termes de qualité et/ou en termes de collaborations. Ainsi, la première fiche décrit les résultats de plusieurs projets régionaux sur de l'analyse d'images médicales. Outre les publications, ces travaux ont abouti à la fois à une collaboration régionale forte avec l'INRAE et l'INSERM mais aussi à la création d'un logiciel (en phase de valorisation) et à des actions de diffusion. La deuxième fiche décrit un ensemble d'activités autour du traitement d'images bas niveaux. Ces activités, assez spécifiques au niveau de l'équipe, s'inscrivent dans le temps long et ne s'appuient sur aucun financement lié à des activités de dépôt de projet. Elles reposent essentiellement sur une collaboration suivie avec la Hong Duc University au Vietnam et le concours sur des financements de bourses par le gouvernement Vietnamien, complétés au niveau fonctionnement par le budget récurrent de l'équipe. Ces activités ont abouti à la fois à des publications de très bon niveau, à une collaboration internationale qui prend de l'ampleur et à la réalisation d'une plateforme d'analyse de flux TV. La troisième fiche évoque quant à elle une activité autour des méthodes à base de réseaux de neurones sur graphes, qui s'inscrit dans la thématique graphe qui dans l'ensemble a donné lieu à toute une série de publications de bon niveau. Pour compléter ces éléments, 3 publications sont présentées brièvement. Toutes les trois sont des publications de très bon niveau, que ce soit sur les aspects théoriques d'opérateurs de traitement d'image (article un et deux) ou bien sur l'analyse de vidéos (rythme cardiaque). Cette dernière repose sur une collaboration internationale suivie (et hors projet financé) et a abouti à la mise en place d'une plateforme très utilisée.

### Élément 1 : fiche Projet SILA3D

Tout à fait représentatif des activités de RFAI et du LIFAT, le [projet SILA 3D](#) a été initié en 2016 autour d'un consortium pluridisciplinaire réunissant des chercheurs en informatique, en neuro-anatomie et en biologie animale ainsi que des neurochirurgiens (LIFAT, INRAE, INSERM). Sur la base des travaux de thèse de Gaetan Galisot [4], le projet s'est poursuivi lors des projets Régionaux NeuroGéo et [Neuro2Co](#) pour aboutir à la création de [SILA3D](#), une plateforme logicielle, en accès libre permettant la segmentation et représentation des données anatomo-fonctionnelles sous forme de **graphes anatomiques locaux (GAL)**. Une nouvelle méthode de **segmentation sémantique** exploitant une **combinaison de multiples classificateurs locaux** (plutôt qu'un seul comme habituellement) liés les uns aux autres à l'aide d'un graphe topologique permettant d'apprendre et de représenter les relations spatiales existantes entre les éléments/objets à détecter constitue le cœur scientifique de la méthode proposée (cf fig 1).

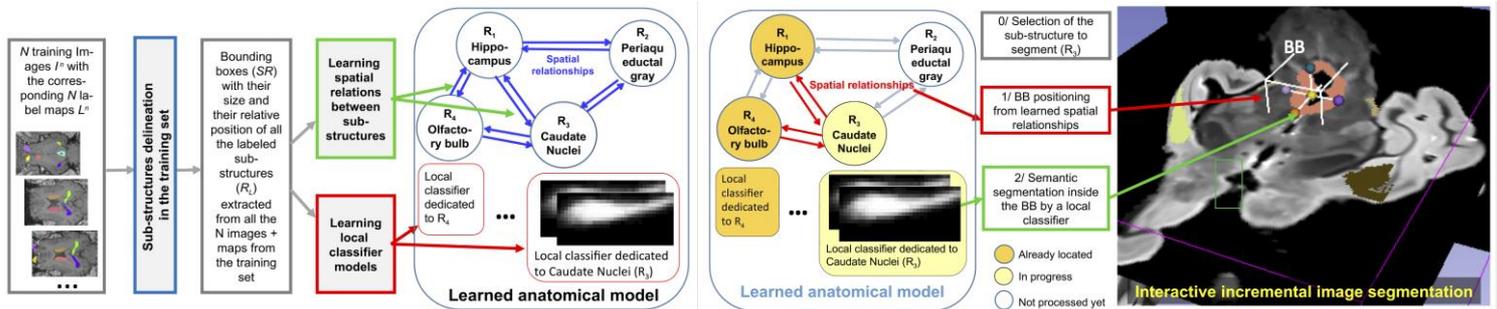


Figure 1 : Vision synthétique de la méthodologie SILA 3D : (a) apprentissage de GAL – (b) segmentation basée GAL

La plateforme SILA3D permet la production rapide de multiples cartographies cérébrales complètes ou partielles insensiblement à la résolution des images. Le site [Lead DBS](#) démontre le très fort besoin de cartographies cérébrales et la difficulté à les obtenir. Les innovations scientifiques mentionnées ci-dessus (création et exploitation de cartographies GAL) rendent possible l'exploitation d'informations d'une résolution de l'ordre de la centaine de  $\mu\text{m}$ . Elles localiseront, in-vivo, des structures anatomiques seulement visibles jusqu'alors sur coupes anatomiques ex-vivo. De telles informations n'étant, pour le moment, pas directement accessibles in vivo, leur prise en compte au sein de logiciels de segmentation ou de planification d'implants cérébraux constituerait une avancée incontestable pour le traitement par Stimulation Cérébrale Profonde de nombreuses pathologies cérébrales (Parkinson, épilepsie...).

Outre les **publications** déjà acceptées [2,3,4,5,6,7] ou en cours de soumission [1], ces travaux ont donné lieu à une **déclaration d'invention** signée en décembre 2018 afin de sécuriser la propriété intellectuelle des méthodes informatiques utilisées (sous le contrôle du SPIV de l'université de Tours). Les droits liés à cette invention sont dévolus à l'université de Tours. Il est prévu de rédiger une nouvelle déclaration d'invention et de faire un **dépôt APP** (Agence pour la Protection des Programmes) pour prendre en compte les évolutions effectuées depuis 2018. Divers modes de valorisation et de pérennisation des résultats produits dans le cadre de ce projet peuvent être envisagés. Certaines voies (création d'une équipe au sein du [CETU ILIAD3](#), création d'une spin-Off, négociation d'accords de licences avec un partenaire ou une entreprise, ...) ont déjà fait l'objet de réflexions en lien avec le SPIV (Service Partenariats Innovations et Valorisation) de l'université de Tours et [CVALO](#). SILA3D a été sélectionné pour participer à plusieurs concours via CVALO visant à identifier les travaux de recherche pouvant donner lieu à la création de startup. Une première levée de fond a permis de lancer une étude de marché flash. Les start-ups [Rebrain](#) et [Neuralink](#) constituent de bons exemples qui démontrent l'intérêt et la viabilité des résultats du projet SILA3D.

Les activités de recherche autour du projet SILA3D se poursuivent toujours et donne actuellement lieu à :

- Projet DeepBrainMap qui vient d'être soumis à l'APR CVL 2021 avec un partenariat de haut niveau (LIFAT porteur, INRAE Nouzilly, INSERM-iBrain, CEA-NeuroSpin, zooparc de Beauval, Siemens) montrant l'intérêt porté à ces travaux à un niveau au moins national.
- 2 bourses doctorales ont été obtenues pour sept 2022 : la première dédiée à l'étude de l'encéphale ovin: co-financement INRAE-Region CVL pour un projet doctoral codirigé par JY Ramel (LIFAT) et E Chaillou (partenaire INRAE-PRC); la seconde pour la création d'un atlas multi-échelles de l'anatomie et de la connectivité des ganglions de la base chez l'homme et le macaque (INSERM).

Notons que le projet SILA3D s'inscrit pleinement dans les domaines d'excellence soutenus par la région CVL, dans les domaines de la [SFR FED Neuro-Imagerie](#) Fonctionnelle et du [RTR \(Réseaux Thématiques de Recherche\) DIAMS](#). Ces objectifs recherche de SILA3D ont également été repris dans le cadre du **projet Loire Val-Health qui vient d'être soumis par l'UT dans le cadre du PIA4**. La région Centre Val de Loire bénéficie de la présence, sur son territoire, de centres d'expertise renommés dans les domaines des neurosciences et de l'imagerie cérébrale (IBrain-INSERM, PRC-INRAE, GIS Autisme, ...).

Le projet SILA3D est aussi l'occasion de réaliser des **actions de diffusion de la culture scientifique** :

- Elodie Chaillou (INRAE) et JY Ramel (LIFAT) ont été [marraine et parrain scientifiques](#) d'une classe du collège Corneille-Tours dans le cadre de l'action La main à la pâte pour un projet comparant cerveau humain et cerveau artificiel ;
- Le projet [Neuro2co](#) avait vocation à diffuser la culture scientifique au travers d'une formation scientifique intitulée "à la découverte du cerveau" menée en collaboration avec la Maison pour la Science Centre Val de Loire ;
- les chercheurs du LIFAT; de l'INRAE et d'INSERM participent régulièrement pour des conférences grand public ou des tables rondes organisées par différentes associations locales et régionales (centre sciences, semaine du cerveau, semaine de la mémoire, ...).

## REFERENCES

1. Galisot G, Ramel JY, Brouard T, Serres B., Chaillou E. 2021. *Visual and Structural Feature Combination in an Interactive Machine Learning System for Medical Image Segmentation, Accepted for publication (MLWA-D-21-00426R1) Journal Machine Learning with Applications, Elsevier.*
2. Zhuang X, Galisot G, Ramel JY et al. 2019 Evaluation of algorithms for Multi-Modality Whole Heart Segmentation: An open-access grand challenge. *Medical Image Analysis* 58
3. Galisot G, Brouard T, Ramel JY, Chaillou E. 2019. A Comparative Study on Voxel Classification Methods for Atlas based Segmentation of Brain Structures from 3D MRI Images. *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. SCITEPRESS. doi:10.5220/0007358503410350
4. Gaetan Galisot: Segmentation interactive d'images 3d multimodales. Université de Tours Thèse 2018
5. Galisot G., T. Brouard et J.-Y. Ramel. E. Chaillou. Image segmentation using local probabilistic atlases coupled with topological information. 12th International Conference on Computer Vision Theory and Applications. VISAPP 2017.
6. Galisot G., T. Brouard, JY Ramel: Local Probabilistic Atlases and a Posteriori Correction for the Segmentation of Heart Images. STACOM@MICCAI 2017: 207-214
7. Galisot G., T. Brouard et J.-Y. Ramel. Apprentissage et exploitation d'un graphe topologique d'atlas probabilistes locaux pour la segmentation d'images IRM. 20e Congrès RFIAP, Clermont-ferrand. 2016.
8. [http://www.rfai.li.univ-tours.fr/PublicData/3D\\_Brain\\_Seg/home.html](http://www.rfai.li.univ-tours.fr/PublicData/3D_Brain_Seg/home.html)

## Elément 2 : Fiche activités autour du traitement temps-réel d'images et vidéos

L'équipe RFAI a mené par le passé différents travaux sur l'optimisation des algorithmes de traitement d'images sans apprentissage (e.g. opérateur de détection de jonctions en complexité quasi-linéaire, correction d'artefacts de compression dans l'espace DCT). Depuis 2017, cet axe a principalement évolué vers le traitement temps-réel d'images et vidéos. On peut noter trois activités saillantes.

**Opérateur temps-réel de détection de texte dans les scènes naturelles** : ces travaux sont en lien avec la thèse de C.D. Nguyen qui s'est déroulée de 2016 à 2021. Durant cette thèse, un nouvel opérateur de traitement d'images temps-réel RT-LoG a été proposé pour la détection de texte dans les images de scènes naturelles. Les travaux autour de cet opérateur ont fait l'objet de plusieurs publications dont une revue à l'international JRTIP<sup>[1]</sup> et des publications dans la conférence à l'international VISAPP<sup>[2]</sup>. L'opérateur proposé est le plus performant de la littérature sous contraintes de précision et de temps de traitement. Il atteint des performances de 30 FPS sur des vidéos en quatre fois haute résolution à partir d'un CPU en architecture standard.

**Station TV et détection temps-réel de segments vidéo** : dans la continuité de ses activités sur le traitement temps-réel d'images et de vidéos, l'équipe RFAI développe depuis 2018 une station de calcul parallèle dédiée à la télévision. Cette station, dite Station TV, a fait l'objet d'un investissement de 34k€ de l'Université de Tours. Elle a été présentée à l'occasion de différentes publications en conférences et workshops (AI4TV2019, CBMI2021, ORASIS2021, ICIAP2021)<sup>[3]</sup>. Elle permet, à partir de cartes matérielles, le traitement temps-réel de 32 flux vidéo TV en simultané à la fois dans le domaine spatial et compressé H.264. Le traitement est supporté par des processeurs à très hautes performances.

La thèse de doctorat de H. Van Le a été engagée en 2020 en lien avec la station TV sur la problématique de détection temps-réel de segments vidéo. Les travaux ont été orientés en première partie sur l'évaluation de performance. Un nouveau protocole pour la constitution de base de données a été proposé. Ce protocole garantit une capture à bruit minimal et à l'échelle ainsi qu'une génération robuste de vérité terrain à coût d'interaction utilisateur réduit. La base constituée, nommée STVD, a fait l'objet de différentes publications dans les conférences ICIAP2021 et ORASIS2021<sup>3</sup>. Elle est aujourd'hui la plus importante de la littérature avec plus de dix mille heures de vidéos. Elle a été mise à disposition sur un serveur de l'Université de Tours<sup>[4]</sup> sous contrôle d'un CGU établi par les services juridiques.

Les travaux de recherche autour de la Station TV couvrent de multiples perspectives en coopération entre les équipes du laboratoire<sup>[5],[6]</sup>.

**Collaboration FICT / HDU** : au-delà de son bilan scientifique, les travaux de recherche sur l'axe se sont inscrits dans le cadre d'une coopération à l'international soutenue avec la faculté FICT de l'Université de Hong Duc<sup>[7]</sup> (Thanh Hoa, Vietnam). Cette coopération a abouti tout d'abord à la mise en place d'une thèse en codirection en 2016 de C.D. Nguyen, lecteur de la FICT, sur programme vietnamien VP911. Afin de consolider la collaboration, un MoU<sup>[8]</sup> a été signé en 2018 entre les Universités de Tours et de Hong Duc. Un second recrutement doctoral d'un lecteur de la FICT (H. Van Le) a été effectué en 2020 sur bourse d'établissement de l'Université de Tours. Un accord de recherche LIFAT / FICT, faisant suite au MoU, devrait être établi d'ici fin 2022. Des perspectives de recrutements doctoraux sont à l'étude pour 2023, 2024 sur les programmes VIED 89 et de l'ambassade de France au Vietnam.

### REFERENCES

<sup>[1]</sup> Journal of Real-Time Image Processing (JRTIP)

<sup>[2]</sup> Conference on Computer Vision Theory and Applications (VISAPP)

<sup>[3]</sup> International Workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV), Conference Content-Based Multimedia Indexing (CBMI), Journées francophones des jeunes chercheurs en vision par ordinateur (ORASIS), International Conference on Image Analysis and Processing (ICIAP)

<sup>[4]</sup> <http://dataset-stvd.univ-tours.fr>

<sup>[5]</sup> **RFAI-BDTLN** : (i) analyse de contenu audio/vidéo (traitement du langage parlé et reconnaissance de visage) pour la vérification de faits politiques

(ii) enrichissement de métadonnées TV par recherche d'entités nommées et qualité de vidéos temps-réel pour l'annotation automatique texte / image

<sup>[6]</sup> **RFAI-ROOT** : détection temps-réel de segments de vidéos dans l'espace H.264/AVC et ordonnancement de machines parallèles pour la capture

<sup>[7]</sup> <http://www.hdu.edu.vn/>

<sup>[8]</sup> Memorandum of Understanding

### Élément 3 : Fiche travaux autour des réseaux de neurones sur graphes

La plupart des objets d'intérêt de notre vie d'aujourd'hui sont basés sur des objets discrets avec des relations séquentielles (chaînes) ou plus complexes (graphes). On peut évoquer les liaisons entre les atomes dans une molécule, les relations entre les personnes dans les images ou dans les vidéos pour n'en citer que quelques-uns. La prédiction des propriétés de tels objets entre dans le cadre de la reconnaissance de formes structurelles. Dans ce contexte, le Deep Learning a permis des avancées majeures notamment avec les réseaux de neurones sur graphes (Graph Neural Network : GNN).

Partant du constat que les modèles existants contraignent les graphes à être projeté (localement) dans des espaces vectoriels, l'équipe RFAI a travaillé à des modèles opérant directement dans l'espace des graphes. En particulier, la convolution est remplacée par un problème de mise en correspondance de graphes. Plus précisément, en tout nœud du graphe d'entrée, un sous-graphe est extrait puis mis en correspondance avec un graphe filtre afin de capturer la structure locale du graphe. Ces travaux ont donné lieu à 2 publications dans des revues [1, 2] et une communication dans un workshop [3]. Des premiers modèles ont été implémentés et des applications sur des images d'insectes ont été réalisées dans le cadre du projet régional CARAMBA.

Une deuxième avancée a eu lieu lors de l'adaptation des Transformers aux graphes. Dans le prolongement de cette idée, nos travaux s'orientent vers l'étude de la mise en correspondance de graphes dans l'objectif de construire un mécanisme d'attention structurelle.

#### REFERENCES

- [1] Maxime Martineau, Romain Raveaux, Donatello Conte, Gilles Venturini:  
Graph matching as a graph convolution operator for graph neural networks. *Pattern Recognit. Lett.* 149: 59-66 (2021)
- [2] Maxime Martineau, Romain Raveaux, Donatello Conte, Gilles Venturini:  
Learning error-correcting graph matching with a multiclass neural network. *Pattern Recognit. Lett.* 134: 68-76 (2020)
- [3] Romain Raveaux, Maxime Martineau, Donatello Conte, Gilles Venturini:  
Learning Graph Matching with a Graph-Based Perceptron in a Classification Context. *GbrPR 2017*: 49-58

### Élément 4 : Publication 1

**C.D. Nguyen, M. Delalandre, D. Conte and T.A. Pham. Fast RT-LoG operator for scene text detection. *Journal of Real-Time Image Processing (JRTIP)*, 18(1), pp. 19-36, 2021.**

**Mots clés :** détection de texte, Laplacien de Gaussienne (LoG), amas, points d'intérêt, filtrage gaussien, représentation en échelle, modèle de trait, opérateur RT-LoG, temps-réel, prédictibilité

Cet article propose un nouvel opérateur de traitement d'images temps-réel RT-LoG. Cet opérateur est dédié à la détection de texte dans les images de scènes naturelles. Il est basé une fonction de Laplacien de Gaussienne. Il utilise une approche en deux temps pour la sélection des filtres de convolution dans le domaine spatial et en échelle. L'opérateur est évalué sur la base de test ICDAR2017 RRC-MLT en termes de robustesse et temps de traitement. Les résultats sont comparés avec les opérateurs temps-réel de la littérature. Il se présente comme l'opérateur le plus performant sous contraintes de précision et de temps de traitement. Il est en moyenne trois fois plus rapide que l'opérateur RT-LoG de base. Il atteint des performances de 30 FPS sur des vidéos en quatre fois haute résolution à partir d'un CPU en architecture standard.

### Élément 5 : Publication 2

**HIDANE, Moncef, EL GHECHE, Mireille, AUJOL, Jean-François, et al. Image zoom completion. *IEEE Transactions on Image Processing*, 2016, vol. 25, no 8, p. 3505-3517**

**Mots-clés :** super-résolution, régularisation non locale, variation totale, éclatement proximal

Nous considérons le problème de l'estimation d'une image haute résolution à partir d'un couple composé d'une image basse résolution complète et d'une partie de la haute résolution. Nous discutons les différents contextes dans lesquels ce cadre peut survenir et introduisons une stratégie de régularisation non locale. Les détails concernant l'optimisation numérique de l'énergie correspondante sont fournis ainsi qu'une discussion concernant ses avantages et ses inconvénients. Nous détaillons également deux algorithmes basés sur la variation totale et évaluons les performances des méthodes proposées sur un ensemble d'images naturelles et texturées.

### Elément 6 : Publication 3

**Giuseppe Boccignone, Donatello Conte, Vittorio Cuculo, Alessandro d'Amelio, Giuliano Grossi, et al.. An Open Framework for Remote-PPG Methods and their Assessment. IEEE Access, IEEE, 2020, 8, pp.216083-216103. DOI: 10.1109/ACCESS.2020.3040936**

**Mots-clés :** Vision par Ordinateur, Remote Photoplethysmography (rPPG), Signal Processing, Statistical Analysis, Python Library

Cet article présente une étude complète des méthodes d'estimation de la fréquence cardiaque par la photopléthysmographie à distance (rPPG). Les techniques de rPPG ont connu un développement remarquable au cours des dernières années, et plusieurs études ont également été publiées, mais une évaluation solide de leurs performances a été, au mieux, négligée, voire inexistante.

Afin d'étudier, de développer et de comparer de nouvelles méthodes de rPPG, les conditions suivantes doivent être réunies : i) une chaîne de traitement structurée et bien définie pour contrôler les entrées, la sortie et les principaux paramètres des algorithmes ; ii) la disponibilité de multiples ensembles de données avec un protocole d'évaluation fixé et unique ; iii) une évaluation statistique solide des performances des méthodes.

Le cadre proposé est réalisé sous la forme d'une librairie Python appelé pyVHR (abréviation de Python tool for Virtual Heart Rate), qui est disponible gratuitement sur GitHub. La librairie est facile à installer et à utiliser, et surtout elle est conçue de manière à faciliter l'ajout de nouvelles méthodes et nouvelles bases de données.

Pour étayer notre approche, nous évaluons huit méthodes bien connues de rPPG, par le biais d'expériences approfondies sur cinq ensembles de données publiques, et d'une analyse statistique non paramétrique approfondie.

Deux contributions majeures ont été apporté avec cet article :

- i. proposer une plateforme open-source pour l'estimation de la fréquence cardiaque par analyse de vidéos et un cadre unique défini (données, méthodes, performances), pour la comparaison de méthodes ;
- ii. analyser statistiquement l'efficacité des meilleurs algorithmes existants pour l'estimation de la fréquence cardiaque par analyse de vidéos.

## Portfolio ROOT

Le portfolio de l'équipe ROOT contient 3 fiches de projets de recherche ciblés thématiquement et 2 publications. Les 3 projets ont été sélectionnés car ils répondaient à au moins 3 critères : ils se sont déroulés sur toute la période 2016-21, ils ont donné lieu à plusieurs publications de très bon niveau, et ils ont été menés en lien avec des thèses de doctorat. Le premier projet porte sur l'étude et la résolution de problème intégrés d'ordonnancement et de transport aussi bien d'un point de vue théorique qu'appliqué (au milieu de la santé). Les travaux de recherche de ce projet ont été soutenus par deux financements (ANR et Région). Le deuxième projet aborde l'algorithmique modérément exponentielle pour la résolution de problèmes d'ordonnancement. Ces travaux de recherche ont donné lieu à des résultats théoriques très intéressants en se basant sur de nouvelles techniques d'énumération, de mémorisation ou encore d'arborescences. Le troisième projet se focalise sur des problèmes de transports avec des véhicules électriques. Les travaux ont été financés principalement via un projet ANR JCJC et se poursuivent via une thèse en cotutelle. Ils abordent de nouvelles problématiques de transports en raison des spécificités liées à la technologie des véhicules électriques. Enfin les publications ont été sélectionnées de manière à compléter un aperçu global des travaux de recherche de l'équipe. La première présente des résultats de complexités de problèmes d'ordonnancement multi-agents obtenus dans le cadre d'une thèse. La deuxième présente une méthode de résolution pour un problème de plus court chemin multicritère capable d'obtenir les meilleurs résultats sur les benchmarks de la littérature. Cette méthode a été développée dans une la cadre d'une thèse CIFRE et d'une collaboration avec Géovélo, jeune entreprise cofondée par un ex-doctorant de l'équipe et proposant un moteur de calcul d'itinéraires à vélo.

### Élément 1 du portfolio ROOT : Problèmes intégrés d'ordonnancement de la production et de tournées de véhicules

La famille des problèmes d'ordonnancement est étudiée depuis le début des années 1950. Une très grande variété de problèmes a été identifiée, les problèmes polynomiaux les plus « purs » ont rapidement été identifiés, ainsi que les problèmes les plus difficiles. Certains problèmes sont encore à la frontière et font l'objet de recherches. Pour les problèmes difficiles, des méthodes de résolution de plus en plus efficaces sont proposées. Les solvers de programmation linéaire ont également accru leur efficacité sur un grand nombre de problèmes. De la même façon, les problèmes de tournées de véhicules sont étudiés depuis la fin des années 1950. Là aussi, les variantes les plus classiques ont largement été abordées, et les variantes les plus complexes sont encore à l'étude. Les solvers exacts peuvent résoudre des instances allant jusqu'à 200 clients par des techniques de branch-and-cut ou de génération de colonnes. Les méthodes approchées de type algorithme génétique, recherche tabou ou recherches à voisinage variable sont particulièrement efficaces sur de grandes instances.

Très souvent, ces deux problèmes sont abordés de façon séparée. Il y a plusieurs raisons pour cela. Tout d'abord, les transporteurs sont des entreprises indépendantes des entreprises de production, les producteurs font appel à plusieurs transporteurs, et les transporteurs livrent pour plusieurs producteurs. Chacun gère donc son problème d'optimisation de son côté. L'hypothèse implicite est que les entreprises disposent de zones de stockage suffisantes pour permettre aux produits d'attendre le départ en livraison. Dans certains cas, quand par exemple la production et la distribution sont assurées par la même entreprise, ou bien lorsque la zone de stockage est extrêmement réduite et que les produits doivent partir très rapidement en livraison (secteur de l'alimentation ou de la production pharmaceutique par exemple), le problème de l'ordonnancement de la production et celui de la livraison sont intimement liés. Il est inutile de produire rapidement un produit, si son lieu de livraison va engendrer une très grande tournée de livraison, et donc retarder les livraisons de tous les autres produits. On doit alors résoudre un problème qui doit décider de l'ordonnancement des tâches sur les machines de production, mais aussi de la constitution des batches de livraison, et enfin des tournées de livraison pour chacun des batches. On intègre donc dans un unique problème différents niveaux de complexité, où chaque sous-problème est déjà un problème difficile à résoudre. Sur ces problèmes, les solvers sont aujourd'hui inefficaces, et seules des méthodes approchées peuvent être envisagées, même pour des instances de taille modeste.

Une partie des activités de recherche de l'équipe s'est intéressée à l'étude de plusieurs problèmes intégrés, notamment dans le cadre du projet ANR ATHENA [1]. Les premiers problèmes étudiés étaient des cas particuliers classiques ou simplifiés. L'objectif était d'apporter plusieurs résultats théoriques (preuves de complexités et algorithmes exactes) sur ces types de problèmes [2]. La plupart de ces résultats sont présentés dans la thèse de doctorat d'Azzedine Cheref [3]. Les travaux de recherche se sont ensuite portés sur les relations qui existent entre les deux acteurs de ce problème : le producteur et le livreur (thèse de doctorat de Hugo Chevrotton [4]). Plusieurs problèmes ont été étudiés dans le cas où l'un des acteurs domine l'autre et vice-versa [5,6]. Une autre partie des activités de recherche de l'équipe en lien avec ce type problème s'est déroulée dans un contexte hospitalier et plus particulièrement dans la production et la livraison de chimiothérapies, via le projet région DISTRI-CHIMIO [7]. Ce contexte implique la prise en compte de nouvelles données et contraintes supplémentaires venant enrichir la difficulté du problème (comme la stabilité chimique des chimios). Afin de développer de nouveaux outils d'aide à la décision pour le service de production de chimiothérapies du CHRU, la thèse de doctorat d'Alexis Robbes [8] s'est focalisée sur la résolution de ce type de problème dans ce contexte. Plusieurs méthodes de résolution ont été proposées pour des cas avec un livreur [9] ou plusieurs livreurs [10].

D'autres problèmes peuvent être abordés, qui intègrent d'autres problèmes difficiles. Par exemple, l'ordonnancement avec la prise en compte de ressources additionnelles (production de chimiothérapies avec prise en compte des limitations de ressources en cytotoxiques ou ordonnancement de processus avec prise en compte de la mémoire utilisée) intègre la complexité des problèmes d'ordonnancement et la complexité des problèmes de packing (bin packing, sac-à-dos).

L'ensemble de ces travaux ont donné lieu à 3 thèses, 5 articles publiés dans des revues internationales, 10 communications dans des conférences internationales et 6 communications dans des conférences nationales.

- [1] ATHENA (Méthodes d'optimisation pour l'étude intégrée de problèmes décisionnels complexes). Partenaires : LIFAT - LAAS-CNRS, LIMOS, Heudiasyc Projet ANR-13-BS02-0006, porteur J.-C.Billaut (ROOT) – 372 k€- 4ans.
- [2] A. Cheref, A. Agnetis, C. Artigues, J.-C.Billaut, Complexity results and algorithms for an integrated single machine scheduling and outbound delivery problem with fixed sequence, *Journal of Scheduling* 20(6) p. 681-693, 2017
- [3] A. Cheref (2017). Méthodes de Recherche Opérationnelle pour la résolution de problèmes intégrés d'ordonnement de la production et de distribution, Thèse de doctorat de l'Université de Tours.
- [4] H. Chevroton (2020). Résolution par des algorithmes exacts et approchés de problèmes intégrés d'ordonnement de la production et de tournées de véhicules, Thèse de doctorat de l'Université de Tours.
- [5] H. Chevroton, J.-C.Billaut, ROHMER Sonja, A production and distribution framework: Manufacturer dominates, *Computers & Industrial Engineering* 155 p. 107162, 2021.
- [6] H. Chevroton, Y. Kergosien, L. Berghman, J.-C.Billaut, Solving an integrated scheduling and routing problem with inventory, routing and penalty costs, *European Journal of Operational Research* 294(2) p. 571-589, 2021
- [7] DISTRI-CHIMIO (Ordonnement et distribution des chimiothérapies), Partenaires : CHRU de Tours, Eticsys. Projet Région – Intérêt Régional 2017-00117170, porteur J.-C.Billaut (ROOT) – 140 k€ - 4ans.
- [8] A. Robbes (2021). Résolution de problèmes intégrés dans un contexte pharmaceutique : ordonnancement et tournées de véhicules, Thèse de doctorat de l'Université de Tours.
- [9] Y. Kergosien, M. Gendreau, J.-C.Billaut, A Benders decomposition-based heuristic for a production and outbound distribution scheduling problem with strict delivery constraints, *European Journal of Operational Research* 262(1) p. 287-298, 2017.
- [10] A. Robbes, Y. Kergosien, V. André, J.-C.Billaut, Efficient heuristics to minimize the total tardiness of chemotherapy drugs production and delivery, *Flexible Services and Manufacturing Journal* Published: 01 October 2021, 10.1007/s10696-021-09431-w

### Élément 2 du portfolio ROOT : Algorithmique modérément exponentielle en ordonnancement

Cette thématique de recherche a débuté au sein de l'équipe ROOT en 2009. Faire de l'AME (Algorithmique Modérément Exponentielle) c'est proposer, pour des problèmes NP-difficiles au sens de la théorie de la complexité, des algorithmes exacts ou heuristiques dont le temps de calcul au pire cas est modérément exponentiel. Lorsqu'on fait de la Recherche Opérationnelle, de l'optimisation combinatoire, on s'intéresse classiquement à des algorithmes efficaces en moyenne lorsqu'ils sont évalués sur des données expérimentales. L'AME permet donc d'apporter un point de vue complémentaire en s'intéressant sur ces situations où les algorithmes efficaces en moyenne explosent face à la combinatoire. Les travaux dans lesquels nous avons fait de l'AME ont tous concerné la résolution de problèmes d'ordonnement ce qui constitue encore aujourd'hui un axe de recherche novateur et peu défriché en dehors de nos travaux [1].

Lors du contrat 2016-2021, nous sommes intéressés à deux axes principaux :

Axe 1 : proposition d'AME exacts ou approchés et établissement de nouvelles bornes théoriques de complexité pour des problèmes d'ordonnement.

Axe 2 : proposition de techniques d'optimisation efficaces en pratique et inspirées des techniques de l'AME.

S'il nous apparaît fondamental d'explorer les champs théoriques pour préciser la complexité au pire cas des problèmes d'ordonnement, il est également très pertinent de voir comment ces champs peuvent améliorer la résolution en moyenne de ces problèmes. Sur la période, ces travaux ont fait l'objet de collaborations avec Mathieu Liedoff (Université d'Orléans), Federico Della Croce et Michele Garrafa (Politecnico di Torino, Italie). Une thèse de doctorat est en passe d'être soutenue : Olivier Ploton [3] qui a travaillé sur les approches de type Inclusion-Exclusion. La thèse de doctorat de Lei Shang [2] a été soutenue en 2017 et a donné lieu à des travaux sur les approches arborescentes et la programmation dynamique.

Dans le cadre de l'Axe 1, nous avons tout d'abord montré [4] que le problème d'ordonnement noté  $1 || \sum_i T_i$  pouvait être résolu au pire cas en temps  $O^*(2^n)$  et espace polynomial par une méthode arborescente incluant une technique de fusion de nœuds. Nous avons également étudié les problèmes d'ordonnement de type flowshop et montré dans un premier temps [5] que le problème à 3 machines et minimisation de la date d'achèvement pouvait être résolu en temps  $O^*(3^n)$  avant que ce résultat ne soit amélioré et généralisé à un très large ensemble de problèmes de flowshop de permutation et critères de type max ou de type somme [6]. Ce dernier résultat, basé sur la technique d'Inclusion-Exclusion est un résultat extrêmement important puisqu'il montre que ces problèmes peuvent être résolus en temps  $O^*(2^n |I|^m)$  et espace mémoire  $O^*(|I|^m)$ , avec  $n$  le nombre de travaux,  $m$  le nombre de machines et  $|I|$  l'amplitude de l'instance. Cette technique a également permis ces mêmes résultats très généraux pour les problèmes à machines parallèles non reliées [7]. La question qui reste en suspens est de savoir si pour ces familles de problèmes il est possible de les résoudre avec un temps de calcul au pire cas qui soit inférieur à  $2^n$ . Toujours concernant la résolution exacte de problèmes d'ordonnement, nous avons montré [8] que certains problèmes à une machine de type Just-in-Time pouvaient être résolus en temps et espace en  $O^*(1.4143^n)$  et montré que la technique Sort & Search pouvait s'appliquer pour résoudre des problèmes multicritères [9]. Enfin, nous avons commencé à nous intéresser à des résultats d'approximation en temps modérément exponentiel pour des problèmes pour lesquels il n'existe pas d'algorithme d'approximation fonctionnant en temps polynomial. C'est notamment le cas de problèmes d'ordonnement à machines parallèles et minimisation du nombre de travaux en retard [10]. Il s'agit là des premiers résultats connus dans le domaine.

Ces travaux, de nature théorique, viennent préciser la difficulté de problèmes d'ordonnement NP-difficile : ils sont novateurs dans la communauté et permettent d'établir une cartographie de la complexité de ces problèmes.

Dans le cadre de l'Axe 2, nous avons abordé un versant plus expérimental et mis en avant l'impact expérimental que peuvent avoir les travaux plus théoriques menés dans l'Axe 1. En effet, mettre au point un algorithme modérément exponentiel nécessite d'identifier une technique permettant de réduire l'explosion combinatoire au pire cas. La technique qui nous a tout d'abord intéressée, et initiée dans la thèse de Lei Shang, est la technique de memorization ou node merging dans sa version polynomiale en espace utilisée de façon théorique dans [4] au sein d'une méthode arborescente. Cette technique permet d'obtenir sur certains problèmes de bonnes garanties au pire cas sur le temps d'exécution. En pratique, une version à espace mémoire limitée peut être utilisée pour éliminer des formes de redondances rencontrées lors du parcours de l'arbre de recherche. Nous avons montré que cette technique, même dégradée par l'absence d'une mémoire infinie sur un ordinateur, permet d'obtenir d'excellents résultats en pratique sur des problèmes d'ordonnement [11]. Par exemple, pour le problème  $1 || \sum_i T_i$ , les résultats expérimentaux menés montrent qu'une méthode exacte arborescente sans mémorisation pour résoudre ce problème est limitée à des instances contenant 500 travaux tandis qu'avec cette technique, la méthode devient capable dans la même limite de temps de résoudre toutes les instances contenant jusqu'à 1200 travaux. Sur tous les problèmes d'ordonnement sur lesquels cette technique a été utilisée, celle-ci s'est avérée améliorer les méthodes arborescentes de la littérature. Dans le cadre de la thèse de Olivier Ploton nous avons commencé à nous intéresser à des intégrations possibles de la technique d'Inclusion-Exclusion au sein de méthodes arborescentes et notamment du calcul de bornes inférieures pour des problèmes d'ordonnement. Les résultats préliminaires obtenus sont plutôt très satisfaisants mais nécessitent d'être poursuivis.

Les méthodes exactes efficaces expérimentalement, en cas moyen, présentent le défaut d'avoir un comportement au pire cas pénalisant. Les travaux que nous menons dans ce second axe visent donc à y intégrer des techniques issues de l'AME et à enrichir l'arsenal des méthodes de la Recherche Opérationnelle pour la résolution des problèmes d'ordonnement.

L'ensemble de ces travaux ont donné lieu à 2 thèses, 6 articles publiés dans des revues internationales, 5 communications dans des conférences internationales et 6 communications dans des conférences nationales.

- [1] V. T'kindt, F. Della Croce, M. Liedloff (2022). Revisiting scheduling problems in the light of exponential-time algorithms, 4'OR, Invited survey, en cours.
- [2] L. Shang (2017). Exact algorithms with worst-case guarantee for scheduling: from theory to practice, Thèse de doctorat de l'Université de Tours.
- [3] O. Ploton (2022). Inclusion-Exclusion based algorithms for solving, Thèse de doctorat de l'Université de Tours.
- [4] M. Garraffa, L. Shang, F. Della Croce, V. T'Kindt (2018). An exact exponential branch-and-merge algorithm for the single machine total tardiness problem. *Theoretical Computer Science*, 745 (1), pp.133 - 149. <10.1016/j.tcs.2018.05.040>.
- [5] L. Shang, C. Lenté, M. Liedloff, V. T'Kindt (2018). Exact exponential algorithms for 3-machine flowshop scheduling problems. *Journal of Scheduling*, <10.1007/s10951-017-0524-2>.
- [6] O. Ploton, V. T'kindt (2022). Moderate worst-case complexity bounds for the permutation flowshop scheduling problem using Inclusion-Exclusion, *Journal of Scheduling*, soumis, 2nde revision.
- [7] O. Ploton, V. T'kindt (2022). Exponential-time algorithms for parallel machine scheduling problems, *Journal of Combinatorial Optimization*, soumis, 3ème revision (mineure).
- [8] V. T'Kindt, L. Shang, F. Della croce (2020). Exponential time algorithms for just-in-time scheduling problems with common due date and symmetric weights. *Journal of Combinatorial Optimization*, 39 (3), pp.764-775. <10.1007/s10878-019-00512-z>.
- [9] L. Shang, V. T'Kindt (2019). A Sort & Search method for multicriteria optimization problems with applications to scheduling theory. *Journal of Multi-Criteria Decision Analysis*, 26 (1-2), pp.84-90. <10.1002/mcda.1660>.
- [10] F. Della Croce, V. T'Kindt, O. Ploton (2021). Parallel machine scheduling with minimum number of tardy jobs: approximation and exponential algorithms. *Applied Mathematics and Computation*, 397, <10.1016/j.amc.2020.125888>.
- [11] L. Shang, V. T'Kindt, F. Della Croce (2021). Branch & Memorize Exact Algorithms for Sequencing problems: Efficient embedding of Memorization into Search Trees. *Computers and Operations Research*, 128, <10.1016/j.cor.2020.105171>.

### Élément 3 du portfolio ROOT : Problèmes de transport avec véhicules électriques

Les véhicules électriques constituent une technologie prometteuse pour réduire la dépendance au pétrole et l'émission de gaz à effet de serre. Ceci est particulièrement vrai si l'on considère l'utilisation de ce type de véhicule pour le transport de marchandises, qui serait responsable de 30% des émissions de CO<sub>2</sub> de l'Union Européenne. L'utilisation massive des véhicules électriques pour le transport de marchandises est freinée par des contraintes techniques telles que l'autonomie limitée des batteries et de longs temps de recharge. Depuis plusieurs années, de nombreuses recherches ont visé (et visent encore) à relâcher ces contraintes via des avancées technologiques concernant la capacité et la vitesse de recharge des batteries, ainsi que la consommation électrique des véhicules. Cependant, un effort considérable reste nécessaire pour développer des outils permettant de gérer des flottes de véhicules électrique en tenant compte de ces contraintes particulières. L'un des champs dans lequel le besoin est critique est celui des techniques d'optimisation de tournées de véhicules électriques (e-PTV). Les problèmes de tournées de véhicules (PTV) consistent à concevoir des tournées efficaces pour livrer (ou collecter) des marchandises, ou fournir des services à des clients à partir d'un dépôt en satisfaisant un certain nombre de contraintes. Durant les 50 dernières années, de nombreux travaux de recherche ont traité la résolution de différentes variantes de PTV. Malheureusement, la plupart des méthodes de résolution proposées sont difficiles, voire

impossibles, à utiliser dans un contexte où les tournées sont effectuées par des véhicules électriques. En effet, les variantes classiques de PTV considèrent que les tournées sont effectuées par des véhicules thermiques, donc ayant une grande autonomie qui peut être rapidement et facilement restaurée. Ainsi ces méthodes se focalisent sur la construction d'une (longue) tournée pour chaque véhicule, sans prévoir d'arrêts dans des stations-services avec des temps de recharges non négligeables et des disponibilités de bornes de recharge limitée. Pour prendre en compte ces spécificités, les algorithmes d'optimisation pour les e-PTV doivent considérer i) des tournées multiples par véhicule (avec un rechargement de batterie entre chaque), ii) des détours par des stations de recharge, et iii) la coordination de l'accès des différents véhicules aux infrastructures de recharge. Afin de répondre à ces challenges, la communauté scientifique a commencé à développer des techniques spécifiques pour les e-PTV. Cependant, la plupart des problèmes soulevés plus haut restaient ouverts, voire inexplorés. Les travaux de recherche de l'équipe ROOT a contribué à cet effort en imaginant, implémentant, testant et disséminant des méthodes d'optimisation pour résoudre les e-PTV, notamment dans le cadre d'un projet ANR JCJC e-VRO [1]. Des travaux de recherche de trois thèses ont porté sur la gestion de véhicules électriques et planification de leurs tournées pour l'optimisation de transport de marchandises (thèse de Alejandro Montoya [2]), dans le cadre de problèmes de transports (thèse de Laura Echeverri [3]) et dans le cas de problèmes dynamiques (thèse de Nicholas Kullman [4] ayant remporté le prix de thèse du GT2L du GdR RO en 2021). Il est à noter que le fait que la propriété de l'infrastructure de recharge soit privée ou publique a un impact sur les contraintes opérationnelles et les paramètres du problème. Plusieurs problématiques ont donc été étudiées. La première concernait l'étude des problèmes apparaissant dans le cas d'infrastructures de recharge privées [5,6]. Dans ce contexte, la flotte de véhicules et les stations de recharge appartiennent au planificateur de tournées. La seconde portait sur l'étude des e-PTV dans le contexte d'infrastructures de recharge publiques [7,8]. Contrairement au premier scénario, le planificateur de tournées ne détient que la flotte de véhicules et doit partager les stations de recharge publiques avec d'autres utilisateurs. D'autres problématiques ont été également étudiées pour des cas plus classiques [9] ou dans des contextes dynamiques [10].

L'ensemble de ces travaux ont donné lieu à 3 thèses, 6 articles publiés dans des revues internationales, 12 communications dans des conférences internationales et 4 communications dans des conférences nationales. Ces travaux se poursuivent actuellement dans un autre contexte que le transport de marchandises. Il s'agit des transports en communs avec bus électriques. L'objectif est de proposer des outils d'aide à la décision permettant de mieux gérer la gestion et les opérations de recharges des bus. Une thèse et plusieurs publications sont en cours.

- [1] e-VRO (Optimisation des Tournées de Véhicules Electriques). Partenaires : LIFAT, Projet ANR-15-CE22-0005, porteur J. Mendoza (ROOT) – 222 k€ - 4ans.
- [2] A. Montoya (2016). Optimization of freight transportation with electric vehicles, Thèse de doctorat de l'Université d'Angers.
- [3] L. Echeverri (2020). The multi-period electric vehicle routing problem: model and solution approaches, Thèse de doctorat de l'Université de Tours.
- [4] N. Kullman (2020). Dynamic decision making under uncertainty in vehicle routing and logistics, Thèse de doctorat de l'Université de Tours.
- [5] A. Montoya, C. Guéret, J.E. Mendoza, J.G. Villegas (2017). The electric vehicle routing problem with nonlinear charging function. *Transportation Research Part B: Methodological*, 103, 87-110.
- [6] A. Froger, J.E. Mendoza, O. Jabali, G. Laporte (2019). Improved formulations and algorithmic components for the electric vehicle routing problem with nonlinear charging functions. *Computers & Operations Research*, 104, 256-294.
- [7] Ç. Koç, O. Jabali, J.E. Mendoza, G. Laporte (2019). The electric vehicle routing problem with shared charging stations. *International Transactions in Operational Research*, 26(4), 1211-1243.
- [8] N. Kullman, J.C. Goodson, J.E. Mendoza (2021). Electric vehicle routing with public charging stations. *Transportation Science*, 55(3), 637-659.
- [9] N. Kullman, A. Froger, J.E. Mendoza, J.C. Goodson (2021). frvcpy: An open-source solver for the fixed route vehicle charging problem. *INFORMS Journal on Computing*, 33(4), 1277-1283.
- [10] N. Kullman, M. Cousineau, J.C., Goodson, J.E. Mendoza (2021). Dynamic ride-hailing with electric vehicles. *Transportation Science*. Online.

#### Élément 4 du portfolio ROOT : publication 1

**Sadi, F., & Soukhal, A. (2017). Complexity analyses for multi-agent scheduling problems with a global agent and equal length jobs. *Discrete Optimization*, 23, 93-104.**

**Mots clés :** ordonnancement, machines parallèles, multi-agent, complexité, algorithmes.

Cet article se focalise sur l'ordonnancement de travaux indépendants où plusieurs agents sont en concurrence pour exécuter leurs travaux sur des machines parallèles identiques communes : le gestionnaire de ressources (agent global) veut minimiser une fonction de coût associée à tous les travaux, tandis que chaque agent  $k$  veut qu'une autre fonction de coût associée à ses travaux ne dépasse pas une valeur donnée  $Q_k$ . Les tâches ont des durées d'exécutions identiques et les fonctions objectifs sont régulières et monotones en fonction des dates de fin d'exécution des travaux. La fonction de coût global de l'agent correspond à la performance globale de l'atelier indépendamment des fonctions objectives des agents. En considérant plusieurs combinaisons de fonctions objectifs, de nouveaux résultats de complexité sont proposés et des

algorithmes polynomiaux sont dérivés pour trouver une solution optimale qui minimise la fonction objectif globale, sous contraintes que les valeurs des fonctions objectifs des autres agents ne dépassent pas un seuil donné.

#### Élément 5 du portfolio ROOT : publication 2

**Kergosien, Y., Giret, A., Néron, E., & Sauvanet, G. An Efficient Label-Correcting Algorithm for the Multiobjective Shortest Path Problem. INFORMS Journal on Computing, Published Online: 26 Aug 2021 <https://doi.org/10.1287/ijoc.2021.1081>**

**Mots clés :** plus court chemin, multiobjectif, front Pareto, méthode exacte.

Cet article propose un algorithme exact pour résoudre le problème du plus court chemin multiobjectif. La solution consiste à déterminer un ensemble de chemins non dominés entre deux nœuds donnés dans un graphe de manière à minimiser plusieurs fonctions objectifs. Cette étude se déroule dans un contexte de calcul d'itinéraires pour cyclistes. L'algorithme proposé est basé sur un algorithme de type « label-correcting » dans lequel plusieurs techniques d'amélioration supplémentaires sont intégrées. Des expériences numériques comparent cet algorithme aux méthodes les plus connues de la littérature afin de valider ses performances sur des graphes de petite et grande taille, instances du « Center for Discrete Mathematics and Theoretical Computer Science » (DIMACS) issues du neuvième défi DIMACS. De nouvelles instances issues du contexte des itinéraires cyclistes sont également proposées. Les résultats obtenus montrent que l'algorithme proposé est capable de surpasser les meilleurs algorithmes de la littérature.