

# CURRICULUM VITAE

## Marta SOARE

### Table des matières

<b>1</b>	<b>Parcours</b>	<b>2</b>
<b>2</b>	<b>Synthèse des activités</b>	<b>3</b>
2.1	Activités d'enseignement . . . . .	3
2.2	Activités collectives liées à l'enseignement . . . . .	3
2.3	Activités de recherche . . . . .	5
2.4	Activités collectives liées à la recherche . . . . .	6
<b>3</b>	<b>Activités d'enseignement</b>	<b>7</b>
3.1	En tant que vacataire à l'INP Grenoble et à l'Université Grenoble Alpes . . . . .	7
3.2	En tant que MCF à l'Université d'Orléans . . . . .	7
3.3	Pendant mon postdoctorat - enseignement à l'Université Aalto, Finlande . . . . .	9
3.4	Pendant mon doctorat - enseignement à Polytech Lille . . . . .	9
3.5	Activités collectives liées à l'enseignement . . . . .	9
<b>4</b>	<b>Activités de recherche</b>	<b>11</b>
4.1	En tant que MCF - recherche à l'Univ. Orléans et à l'Univ. Grenoble Alpes . . . . .	11
4.2	Pendant mon postdoctorat - recherche à l'Univ. Aalto, Finlande . . . . .	13
4.3	Pendant mon doctorat - recherche à Inria Lille . . . . .	14
4.4	Encadrement d'activités de recherche . . . . .	15
4.5	Prix et participations aux jurys de thèse et projets avec financement externe . . . . .	17
4.6	Activités collectives liées à la recherche . . . . .	17
4.7	Productions scientifiques . . . . .	18

# 1 Parcours

## Informations personnelles

**Date de naissance** : 05/08/1988 (34 ans)    **Etat civil** : mariée, 1 enfant né en 2019  
**Nationalités** : française, roumaine    **Permis** : B  
**Téléphone** : +33 6 81 46 75 88    **Site web** : <https://martasoare.github.io/>  
**Email** : marta.soare@univ-orleans.fr ; martasoare88@gmail.com

## Expériences professionnelles

### Depuis 09/2017 : Maître de conférences

Université d'Orléans, LIFO EA 4022

*Recherche* : équipe CA (Contraintes et Apprentissage)

*Enseignement* : IUT Orléans, départements Informatique et Qlio (Qualité, Logistique Industrielle, Organisation)

### 03/2023 – 08/2023 : Délégation CNRS

Université Grenoble Alpes, LIG UMR CNRS 5217

*Recherche* : équipe SLIDE (ScaLable Information Discovery and Data Ethics)

### 09/2021 – 02/2023 : Maître de conférences en mise à disposition (100% recherche)

Université Grenoble Alpes, LIG UMR CNRS 5217

*Recherche* : équipe SLIDE (ScaLable Information Discovery and Data Ethics)

*Vacations enseignement* : INP Grenoble (ENSIMAG) et Université Grenoble Alpes (UFR IM2AG)

### 05/2019 – 08/2019 : Congé maternité

### 09/2015 – 09/2017 : Postdoctorat

Helsinki Institute for Information Technology (HIIT), Department of Computer Science, Univ. Aalto, Finlande

*Recherche* : équipe PML (Probabilistic Machine Learning)

*Enseignement* : Department of Computer Science, Université Aalto

### 09/2012 – 09/2015 : Doctorat avec mission d'enseignement

Inria Lille, Université Lille 1, CRISAL UMR CNRS 9189

*Recherche* : équipe SequeL (Sequential Learning)

*Enseignement* : Polytech Lille

## Formation

**Doctorat**, Université Lille 1, 10/2012–09/2015. Soutenu le 14 décembre 2015

*Titre de la thèse* : “Allocation séquentielle de ressources dans le modèle des bandits linéaires”

*Directeur de thèse* : Rémi Munos, DR Inria & Senior Researcher Google DeepMind

*Spécialité* : Informatique

*Mention* : Très Honorable

<i>Jury</i> :	<i>Président du jury</i>	Rémi Gilleron	Professeur, Université Lille 3
	<i>Rapporteurs</i>	Michèle Sebag	DR CNRS, LRI, Paris
		Csaba Szepesvári	Professeur, Université d'Alberta, Canada
	<i>Examineurs</i>	Olivier Cappé	DR CNRS, LTCI, Télécom ParisTech
		Liva Ralaivola	Professeur, Université Aix-Marseille
	<i>Directeur de thèse</i>	Rémi Munos	DR Inria & Senior Researcher Google DeepMind
	<i>Co-encadrant</i>	Alessandro Lazaric	CR Inria

**Visites de recherche** : – Microsoft Research New-England Lab., États-Unis, 1 semaine, février 2014  
 – McGill University, RL Lab., Canada, 1 semaine, février 2014

**École d'été** : Machine Learning Summer School (MLSS), Tübingen, Allemagne, 2 semaines, septembre 2013

**Master**, Université Lille 1, 09/2010 – 09/2012

*Spécialité M2* : Économétrie Appliquée

*Unités d'enseignement M2* : Informatique, Statistiques, Économie Appliquée

*Mention M2* : Très Bien

*Stage de recherche* : Laboratoire EQUIPPE, Université Lille 1, 2012 (6 mois)

**Licence**, ASE (Académie d'études économiques), Bucarest, Roumanie, 10/2007–07/2010

*Bourse Erasmus* pour étudier à l'Université de Lille 1 en 2009–2010

## 2 Synthèse des activités

Cette section est une synthèse de mes activités d'enseignement et de recherche, ainsi que des activités collectives associées. Toutes ces activités sont présentées en détail dans la suite.

### 2.1 Activités d'enseignement

Depuis mon recrutement en tant que maître de conférences à l'Université d'Orléans, ainsi que pendant mon postdoctorat à l'Université Aalto et ma thèse à l'Université Lille 1, j'ai enseigné **1095h EqTD**, à un public varié (DUT, cycle ingénieur, master, licence professionnelle), à des informaticiens et à des non-informaticiens, en français et en anglais. Ci-dessous la liste des matières que j'ai enseignées. J'ai été responsable des matières annotées avec **\***.

**Vacataire (2022-2023), 20h EqTD**, à l'INP Grenoble (ENSIMAG) et à l'Univ. Grenoble Alpes (UFR IM2AG)

- Introduction à l'Intelligence Artificielle – ENSIMAG 2A
- Intelligent Systems – MOSIG M1 (en anglais)
- Bases de données – L3 MIAGE (UFR IM2AG)

**MCF (2017-2021), 924h EqTD**, dans deux départements de l'IUT d'Orléans (détails dans la Table 1)

- Département Informatique
  - Introduction à l'algorithmique et à la programmation : 2018, 2019, 2020
  - Structures des données et algorithmes fondamentaux : 2018
  - Conception de documents et d'interfaces numériques : 2019, 2020
  - Algorithmique avancée (en anglais) : 2017, 2018, 2019, 2020
  - Recherche opérationnelle et aide à la décision : 2017, 2018
- Département Qlio (Qualité, logistique industrielle et organisation)
  - **\*** Outils bureautiques (architecture et fonctionnement d'un poste informatique) : 2020
  - **\*** Algorithmique : 2017, 2018
  - **\*** Système de gestion de bases de données : 2017, 2018, 2019, 2020
  - **\*** Introduction à la programmation : 2017, 2018, 2019, 2020
  - Mise en place d'une GPAO (ERP et systèmes d'information) : 2017
- M2 Informatique (Université Orléans / INSA Centre Val de Loire)
  - Initiation à la recherche : 2018, 2019, 2020

**Postdoctorat (2015 - 2017), 23h EqTD**, à l'Université Aalto - Computer Science Department

- **\*** Machine Learning and Sequential Decision Making (en anglais)
- Machine Learning and Differential Privacy (en anglais)

**Doctorat avec mission d'enseignement**<sup>1</sup> **(2013 - 2015), 128h EqTD**, à Polytech Lille

- Initiation à l'algorithmique et à la programmation, département IC2M
- Introduction aux bases de données, département GB-IAAL, Bac+4 et département IMA, Bac + 3
- Programmation Scilab, département GB-IAAL, Bac + 3

### 2.2 Activités collectives liées à l'enseignement

- **Membre élue du Conseil d'Administration de l'IUT d'Orléans (2018 - 2022).**
- **Responsable projets LP (2019 - 2021)** : En charge des projets tutorés (3 semaines) et projets en entreprise (1 semaine) pour la promotion de la licence professionnelle du département Qlio de l'IUT d'Orléans (30 étudiants).
- **Responsable emplois du temps (2018 - 2019)** : En charge d'établir, saisir et assurer la mise à jour de la totalité des emplois du temps des trois promotions (Bac+1 à Bac+3) du département Qlio de l'IUT d'Orléans.
- **Membre du jury de validation des acquis de l'expérience (2020 - 2021)** pour la licence professionnelle au département Qlio. Étude des dossiers et participation aux auditions des candidats.

1. Statut anciennement connu sous le nom de *monitorat*.

TABLE 1 – Synthèse de mes enseignements en tant que MCF à l'Université d'Orléans. Abréviations : I=département Informatique, Q=département Qlio, M2=Master 2 Informatique à l'Université d'Orléans, CM=Cours Magistral, TD=Travaux Dirigés, TP=Travaux Pratiques. J'ai été responsable des matières annotées avec ★.

ANNÉE	MATIÈRE	DÉP.	CM	TD	TP	BAC+	EFF.
2020–2021 <b>289h</b> EqTD	Introduction à l'algorithmique	I		12	32	1	30
	Conception de docs. et d'interfaces numériques	I		12	24	1	30
	Algorithmique avancée (en anglais)	I		7	21	2	30
	★ Outils bureautiques	Q		8	24	1	51
	★ Système de gestion de bases de données	Q			40	2	27
	★ Programmation	Q			40	3	24
	Suivi apprentis	I,Q		40		2-3	4
	Projet tutoré/soutien/stage	I,Q		20		1-3	5
	Initiation à la recherche	M2	3			5	40
2019–2020 <b>242h</b> EqTD	Introduction à l'algorithmique	I		28	32	1	30
	Conception de docs. et d'interfaces numériques	I			13	1	15
	Algorithmique avancée (en anglais)	I		7	21	2	30
	★ Système de gestion de bases de données	Q			40	2	27
	★ Programmation	Q			20	3	24
	Suivi apprentis	I,Q		60		2-3	6
	Projet tutoré/soutien/stage	I,Q		20		1-3	5
	Initiation à la recherche	M2	5			5	40
2018–2019 <b>265h</b> EqTD + décharge maternité 96h EqTD	Introduction à l'algorithmique	I		12	32	1	30
	Structures des données	I		16	24	1	30
	Algorithmique avancée (en anglais)	I		7	21	2	30
	Recherche opérationnelle et aide à la décision	I		7	42	2	55
	★ Algorithmique	Q	7	10		1	57
	★ Système de gestion de bases de données	Q			40	2	27
	★ Programmation	Q			40	3	30
	Suivi apprentis	Q		20		2-3	2
	Projet tutoré/soutien/stage	I,Q		16		1-3	8
2017–2018 <b>128h</b> EqTD + décharge nouveau MCF 64h EqTD	Initiation à la recherche	M2	3			5	8
	Algorithmique avancée (en anglais)	I		7	21	2	30
	Recherche opérationnelle et aide à la décision	I		7	10.5	2	16
	★ Algorithmique	Q	7	10		1	56
	★ Système de gestion de bases de données	Q			20	2	11
	★ Programmation	Q			22	3	15
	Mise en place d'une GPAO	Q		2	4	3	30
	Encadrement de stage	I		14		2-3	7

- Responsable des cours (2017 – ) : Depuis la rentrée 2017 je suis responsable de deux ou trois cours par an, pour lesquels je fais la préparation des supports de CM/TD/TP/contrôles et j'assure les corrections.
- Présidente du jury LP (2019 – 2021) au département Qlio de l'IUT d'Orléans.
- Suivi d'apprentis (4 par an) et encadrement des stages et des projets (2017 – ).
- Recrutement des étudiants de première année (2018 – ) : Notamment, pour le recrutement en DUT Informatique, j'ai participé à la réalisation d'un test (QCM) indicatif de positionnement et d'un guide d'entretien.
- Accompagnement des étudiants de première année (2018 – ) : J'effectue des oraux de consolidation et participe au suivi personnalisé des étudiants en difficulté.
- Participation à des forums (un par an) et aux Journées Portes Ouvertes (2017 – ).

TABLE 2 – Synthèse de mes articles d’audience **internationale**, tous **avec comité de lecture**. Les conférences CORE **A/A\*** et les revues Scimago **Q1** (selon le ranking valable l’année de publication) sont en gras.

5 REVUES	<b>Transactions on Dependable and Secure Computing</b> [CDLS23] <b>Journal of Artificial Intelligence Research</b> [CLMS22] <b>Bioinformatics</b> [SPM <sup>+</sup> 18] <b>Machine Learning</b> [DPSK17a] <b>Journal of Computer Security</b> [CLLS23]	
12 ACTES DE CONFÉRENCE	MAIN TRACK (7) <b>ICML 2023</b> [RBCS23] <b>TrustCom 2020</b> [CLLS20] <b>IUI 2017</b> [MSM <sup>+</sup> 17] <b>NeurIPS 2014</b> [SLM14] ProvSec 2020 [CDLS20] ISPEC 2019 [CLLS19] ICMLA 2016 [SAK16b]	SHORT PAPER TRACK (2) <b>IJCAI 2023</b> [CLMS23] <b>EDBT 2022</b> [CSA22]  DEMO TRACK (3) <b>ICDE 2022</b> [MCL <sup>+</sup> 22] <b>CIKM 2022</b> [TCSA22] <b>CIKM 2021</b> [DCS <sup>+</sup> 21]
10 WORKSHOPS	<b>NeurIPS</b> (5) : ML4H 2017, 2×FILM 2016, MultiTask 2014, BayesOPT 2013 <b>ICML</b> (3) : HITL 2017, WCB 2017, DEML 2016 <b>ECML–PKDD</b> (1) : IAL 2017 <b>ICMB–ECCB</b> (1) : MLSB 2017	

## 2.3 Activités de recherche

Mes recherches s’inscrivent principalement dans le domaine de l’**apprentissage automatique interactif (interactive machine learning)**. Cette thématique englobe des scénarios où un agent apprenant interagit séquentiellement avec l’environnement (qui peut être un humain) : l’agent choisit une action et l’environnement répond avec une récompense associée à cette action. L’agent peut prendre en compte le **renforcement** donné par l’environnement pour apprendre quelles actions lui permettent de maximiser ses récompenses.

Pendant mon *doctorat* je me suis intéressée surtout à la **théorie des bandits à plusieurs bras**.

Au *postdoctorat*, d’une part j’ai travaillé sur la modélisation des algorithmes de **human-in-the-loop-ML** via des algorithmes pour l’élicitation interactive des retours d’experts. D’autre part, je me suis aussi intéressée à l’application de ces méthodes dans la **médecine personnalisée** et les **systèmes de recommandation**.

Depuis que je suis *MCF*, en plus de la continuation des travaux sur ces thématiques, j’ai fait une mobilité thématique majeure vers la **sécurité informatique**, pour la proposition des algorithmes de machine learning séquentiel avec des garanties de sécurité quand les données et les calculs sont externalisés dans un cloud.

La Table 2 donne une synthèse de mes articles d’audience internationale, tous avec des comités de lecture. Plus de détails pour chaque article sont donnés en Section 4.7. Je reprends ici en ordre chronologiques mes principales thèmes de recherche et les publications les plus représentatives pour chaque thème :

**Doctorat à Inria Lille** (sous la direction de Rémi Munos et Alessandro Lazaric)

- analyse des méthodes d’apprentissage actif [SLM13]
- étude de bornes de performances des algorithmes de bandits à plusieurs bras [SLM14]
- apprentissage séquentiel et par transfert [SALP14] (collaboration avec Joelle Pineau, Univ McGill, Canada)

**Postdoctorat à l’Univ Aalto, Finlande** (collaboration avec Samuel Kaski et des membres de son groupe)

- *interactive user modeling* [SMM<sup>+</sup>16]
- *humain-in-the-loop machine learning (ML)* [MSM<sup>+</sup>17]
- *sequential experimental design for knowledge elicitation* [DPSK17a]

**MCF à l’Univ. Orléans/LIFO (dont mobilité à l’Univ. Grenoble Alpes/LIG 09/2021-09/2023)**

- *human-in-the-loop ML* pour la médecine personnalisée (continuation des travaux du postdoctorat) [SPM<sup>+</sup>18]
- *ML-as-a-service* sécurisé (avec Radu Ciucanu–LIFO et Pascal Lafourcade–LIMOS) [CLLS20]
- exploration des données in-DBMS (avec Sihem Amer-Yahia–LIG et Radu Ciucanu) [DCS<sup>+</sup>21, CSA22, TCSA22]
- apprentissage fédéré sécurisé (avec Radu Ciucanu, Pascal Lafourcade) [MCL<sup>+</sup>22, CLMS22, CDLS23]
- apprentissage par renforcement distribué et *private* (avec Merwan Barlier–Huawei, Igor Colin–Huawei) [RBCS23].

## Encadrement d'activités de recherche

- **Deux doctorants, les deux avec publications A/A\*** : *Gael Marcadet* (2021-2024), sur la sécurisation d'algorithmes d'apprentissage fédéré. Je suis la co-encadrante principale (40%) de la thèse, avec Radu Ciucanu (30%) et Pascal Lafourcade (30%, directeur de thèse). Cette thèse fait suite au stage de M2 de Gael, effectué à l'Université d'Orléans (LIFO), dont j'ai aussi été la co-encadrante principale.  
*Alexandre Rio* (2022-2025), sur l'apprentissage par renforcement distribué et respectueux de la vie privée. Je co-encadre (70%) ce thèse CIFRE avec Sihem Amer-Yahia (30%, directrice de thèse) pour le côté académique, ainsi qu'avec Merwan Barlier et Igor Colin pour le côté industriel (Huawei). J'ai eu le rôle principal dans la définition du sujet de cette thèse et dans l'obtention du financement auprès de l'ANRT.
- **Une ingénieure de recherche avec publication A/A\*** : Dans le cadre de deux chaires MIAI (*Contextual Recommendations in Action* et *MyWay2Health*), en collaboration avec Sihem Amer-Yahia, j'ai co-encadrée Sandrine Da Col (01/2021-01/2022). Ce projet a mené à un article.
- **Six stages de M2, dont 3 avec publications A/A\*** : Trois stages (02-07/2022) au LIG, dont un avec publication A/A\* dans le cadre d'un projet Emergence 2022 dont j'ai été co-PI, avec Radu Ciucanu. Ce projet avait pour but de renforcer les liens entre les équipes SLIDE et APTIKAL du LIG.  
Au LIFO, avec Radu Ciucanu (LIFO) et Pascal Lafourcade (LIMOS), j'ai co-encadré deux stages de M2 sur la sécurisation des algorithmes de bandits (un en 2019-2020, un en 2020-2021), les deux avec publications A/A\*. Pendant mon postdoctorat, j'ai co-encadré un stage de M2 avec Samuel Kaski (Université Aalto, Finlande).
- **Deux stages de M1** : Dans le cadre d'une collaboration avec le Laboratoire Hubert Curien (Saint-Étienne), avec Emilie Morvant et Amaury Habrard j'ai co-encadré deux stages de trois mois (04-06/2017, 04-06/2018).

## Prix

- **Honorable mention at CIKM 2022's Demo Track** <https://www.cikm2022.org/best-paper>.
- **Top 30% best reviewers** de la conférence NeurIPS 2018.
- **Prix du meilleur article JPFDA 2015**, accordé par l'Association Française pour l'Intelligence Artificielle.
- Travel award pour NeurIPS 2014, sponsorisé par Google.
- 2×Travel award pour Workshop Women in ML (WiML), co-localisé avec NeurIPS, en 2013 et 2014.

## Jurys de thèse

- 2023 : Examinatrice de la thèse d'Anastasiia Doynychko, Université Grenoble Alpes, LIG. Directeur de thèse Massih-Reza Amini.
- 2022 : Examinatrice de la thèse de Geovani Rizk, Université Paris Dauphine PSL, LAMSADE. Directeurs de thèse Yann Chevaleyre et Rida Laraki.

## Projets de recherche avec financement externe

- Co-porteuse d'un projet **Emergence** 2022, Laboratoire d'Informatique de Grenoble (LIG).
- Porteuse d'un projet de thèse (2022-2025) financé par l'ANRT (**CIFRE**).
- Core Member dans la Chaire **MIAI Contextual Recommendations in Action**.
- Participante au projet **ANR Herelles** (11/2020 - 11/2024).

## 2.4 Activités collectives liées à la recherche

- **Organisation de conférences** : Web Co-Chair ECML-PKDD 2022, Co-organisatrice JFPDA 2013, Co-organisatrice WiML@ICML 2015, Volontaire pour ICML 2015.
- **Vice-présidente de la Société Savante Francophone d'Apprentissage Machine SSFAM**<sup>2</sup>(2021-).
- **Activités de relecture** :
  - Membre de 15 comité de programme (en moyenne relecture de 5 articles par conférence)

---

2. <http://ssfam.org/>

- Core A\* : ICML 2022, IJCAI 2023/2022/2021/2017, NeurIPS 2018/2016, AAAI 2018 ;
- Core A : ECML-PKDD 2016 ;
- Autres : HITL@ICML2017, TSW@NeurIPS2016, CAA 2023, EGC 2022, RJCIA 2020, JFPDA 2013.
- Relectrice pour des revues : Machine Learning Journal (Core A).
- Relectrice pour des conférences : COLT 2017 (Core A\*), ALT 2016 (Core B), IEA/AIE 2017, AlgoTel 2021.
- **Membre de comités de sélection/suivi :**
  - 2022 : Examinatrice pour *tenure-track associate professor position in ML*, Univ. Bergen (Norvège).
  - 2021 : 3 × Comité de sélection pour des postes MCF 27 (Univ. Orléans/LIFO et 2×Univ. Artois/CRIL).
  - 2018 : Comité de suivi de thèses de l'École Doctorale 551 MIPTIS.
  - 2016 : Comité de sélection des étudiants de doctorat et de master dans des thématiques liées au groupe Probabilistic Machine Learning, à l'Université Aalto (Finlande).

### 3 Activités d'enseignement

J'ai effectué **1095h EqTD** depuis le début de ma carrière. Dans la Section 2.1 j'ai présenté une synthèse des thématiques pour ces enseignements. Dans la suite, je présente plus en détail le contenu et le contexte de ces enseignements, en ordre antichronologique : en Section 3.1 l'enseignement effectué en tant que vacataire à Grenoble, en Section 3.2 l'enseignement effectué en tant que MCF, en Section 3.3 l'enseignement pendant mon postdoctorat et en Section 3.4 l'enseignement fait pendant mon doctorat (*monitrice*). En Section 3.5 je présente ma participation aux activités collectives et administratives liées à l'enseignement.

#### 3.1 En tant que vacataire à l'INP Grenoble et à l'Université Grenoble Alpes

En 2022-2023 j'ai participé à l'enseignement des TD/TP dans trois unités d'enseignement.

- **Introduction à l'Intelligence Artificielle** à l'INP Grenoble (ENSIMAG 2A), niveau Bac+4. Le contenu porte sur l'introduction à l'apprentissage automatique (machine learning) et à l'intelligence artificielle symbolique. 6 séances de 1.5h pour 30 étudiants.
- **Intelligent Systems** à l'Université Grenoble Alpes (UFR IM<sup>2</sup>AG), niveau Bac+4, dans le cadre du master MOSIG (Master of Science in Informatics at Grenoble). Le contenu porte sur l'introduction à l'intelligence artificielle symbolique et est enseigné en anglais. 2 séances de 1.5h pour 30 étudiants.
- **Base de Données** à l'Université Grenoble Alpes (UFR IM<sup>2</sup>AG), niveau Bac+3, dans la L3 MIAGE. Le contenu porte sur l'introduction aux bases de données relationnelles (algèbre relationnelle et requêtes SQL, normalisation). Enseignement pour un groupe de 24 étudiants, dans 5 séances de 1.5h.

#### 3.2 En tant que MCF à l'Université d'Orléans

J'effectue mon service d'enseignement majoritairement à l'IUT d'Orléans, 50% au département Informatique et 50% au département Qlio (Qualité, Logistique Industrielle et Organisation). La Table 1 donne une synthèse de mon service d'enseignement en tant que MCF. En 2021-2022 je suis en mise à disposition 100% recherche et je n'effectue donc pas de service d'enseignement.

Dans la suite, je donne des détails sur le contenu des cours que j'ai enseignés pendant cette période. Pour les deux années de DUT, ces descriptions suivent le programme pédagogique national (PPN 2013) pour l'Informatique<sup>3</sup> et pour Qlio<sup>4</sup>. Je participe aussi aux enseignements de la Licence Professionnelle (Bac+3) donnée par le département Qlio et dans le cours d'*Initiation à la recherche* pour les étudiants de M2 Informatique.

3. [https://cache.media.enseignementsup-recherche.gouv.fr/file/25/09/7/PPN\\_INFORMATIQUE\\_256097.pdf](https://cache.media.enseignementsup-recherche.gouv.fr/file/25/09/7/PPN_INFORMATIQUE_256097.pdf)

4. [http://cache.media.enseignementsup-recherche.gouv.fr/file/25/09/9/PPN\\_QLIO\\_256099.pdf](http://cache.media.enseignementsup-recherche.gouv.fr/file/25/09/9/PPN_QLIO_256099.pdf)



## IUT Orléans - Modules enseignés au Département Informatique

**Introduction à l’algorithmique et à la programmation - M1102** : les premières notions de qualité (nommage des variables, assertions, documentation, etc.), la notion de types et de données, structures algorithmiques fondamentales (choix, répétitions), introduction à la programmation orientée objet. L’implémentation des algorithmes a été faite en Python.

**Structures des données et algorithmes fondamentaux - M1103** : les algorithmes fondamentaux avec des structures simples (recherche d’un élément, parcours, tri, etc.), une première notion des performances des algorithmes utilisés, écriture et lecture dans des fichiers. Les travaux pratiques ont été faits en Python.

**Conception de documents et d’interfaces numériques - M1105** : la séparation contenu-structure-présentation, les technologies du Web pour la description de documents et d’interfaces : HTML, CSS, la conformité des sites Web aux standards d’accessibilité W3C. Nous avons utilisé le moteur de template Jinja (Python).

**Algorithmique avancée - M3103** : les structures de données récursives - arbres, dictionnaires, ensembles ; algorithmes récursifs et itératifs sur ces structures, arbres lexicographiques. J’ai enseigné ce module en anglais. Pour la programmation, nous avons utilisé Python.

**Recherche Opérationnelle et Aide à la décision - M4202C** : présentation d’outils de base pour l’analyse des données (classifieur de Bayes, classifieur des plus proches voisins, fouille de texte). La mise en œuvre avec R.

## IUT Orléans - Modules enseignés au Département Qlio

**Outils bureautiques - M1206** : mon intervention dans ce module porte sur l’introduction à l’architecture et le fonctionnement d’un poste informatique et sur la présentation d’outils pour le traitement de texte.

**Algorithmique - M2106** : introduction à l’algorithmique et implémentation des algorithmes simples, analyse des problèmes liés à l’exercice de sa profession et modélisation à l’aide d’un algorithme. Nous avons utilisé le langage de programmation structurée Visual Basic.

**Système de gestion de bases de données - M3105** : savoir exploiter et faire évoluer un système de gestion de base de données, implémenter des outils complémentaires d’interface homme-machine (système de navigation, boutons, etc.), programmer des requêtes simples. Ce travail a été fait en utilisant Access.

**Programmation (Licence Professionnelle)** : introduction à l’algorithmique et à la programmation via des projets de programmation inspirés des tâches et problématiques qu’ils rencontreront dans leur vie professionnelle. Nous avons utilisé Visual Basic.

**Mise en place d’une GPAO (Licence Professionnelle)** : ce module est assuré majoritairement par mes collègues logisticiens, j’ai eu une intervention pour discuter des ERP (progiciel de gestion intégrée, en anglais : Enterprise Resource Planning) et des systèmes d’information. Nous avons utilisé SAP.

## Suivi apprentis, stages, soutien

Pour les deux départements (Informatique et Qlio), une partie de mon service consiste dans l’encadrement des étudiants, sous différentes formes :

- suivi des apprentis (en moyenne 4 apprentis par an, 2 par département) : rencontres avec les maîtres d’apprentissage, visites en entreprise et discussions avec les apprentis après chaque période en entreprise ;
- soutien des étudiants en première année : oraux de consolidation, séances de TD/TP supplémentaires pour les étudiants en difficulté, suivi-personnalisé ;
- participation à des projets tutorés : proposition et encadrement des projets, participation aux soutenances ;
- suivi des stages : une visite en entreprise par stage et participation aux soutenances.

## M2 Informatique

En plus des heures effectuées à l’IUT d’Orléans, depuis 2018 je participe au cours d’*Initiation à la recherche* du M2 Informatique, pour les étudiants de Bac+5 de l’Université d’Orléans et de l’INSA Centre Val de Loire. Ce cours présente aux étudiants de M2 les thématiques de recherche des membres du LIFO. J’ai eu des interventions de 3h ou de 5h de cours (en 2019-2020), dédiées à un tutoriel sur les techniques et applications autour de la thématique de l’**apprentissage par renforcement et la prise de décision dans l’incertain**, notamment des algorithmes de bandits stochastiques et leur sécurisation.



### 3.3 Pendant mon postdoctorat - enseignement à l'Université Aalto, Finlande

Afin de diversifier mes activités d'enseignement j'ai fait le choix de donner des cours pendant mon postdoctorat (2015–2017). J'ai accumulé ainsi une fructueuse expérience en Finlande, dans un système d'enseignement réputé pour son excellence. J'y ai participé dans l'enseignement de deux cours avancés pour des étudiants de master/doctorat, liées à mes thématiques de recherche en **machine learning**. Le contenu des cours est organisé autour de l'étude d'articles scientifiques et/ou chapitres d'ouvrages.

- **Machine Learning and Sequential Decision Making** : j'ai co-dirigé ce cours avec mon collègue Tomi Peltola (qui était aussi postdoc dans l'équipe PML), entre octobre et décembre 2016. Nous avons conçu le planning, organisé la liste de articles à discuter, et conduit les discussions. C'était un cours de 12h, réparties en six séances devant un public de 40 étudiants.
- **Machine Learning and Differential Privacy** : pendant les trois premiers mois de mon postdoctorat, j'ai participé à ce cours optionnel organisé pour les étudiants en master et doctorat. Mon intervention (en tant que "guest lecturer") a été notamment de donner un tutoriel sur les algorithmes de bandits et de discuter l'utilisation des techniques de *differential privacy* pour garder la confidentialité des utilisateurs des systèmes de recommandation. C'était une intervention de 3h, répartie en deux séances, devant un public de 20 étudiants.

Les enseignements à l'Université Aalto m'ont permis de découvrir des étudiants d'un bon niveau, très passionnés et autonomes. La plupart étaient des étudiants en doctorat, directement intéressés à voir comment ils pourraient utiliser les résultats et outils discutés pendant les cours dans des problèmes de recherche liés à leurs thèses. Cependant, j'ai remarqué un niveau d'intérêt également élevé aux étudiants de master, malgré leur parcours très divers. Dans leurs projets ou présentations, ils sont allés plus loin que demandé, en lisant aussi des articles connexes, ou en faisant l'implémentation de plusieurs algorithmes que prévu.

### 3.4 Pendant mon doctorat - enseignement à Polytech Lille

Pendant ma thèse, j'ai eu une mission d'enseignement à l'Université Lille 1, dans la composante Polytech Lille (2013–2015). J'y ai donné au total 128h EqTD d'enseignement (64h EqTD/an). J'ai eu l'occasion d'enseigner des cours à des futurs informaticiens, mais j'ai aussi enseigné des cours introductifs de programmation et des bases de données à un public qui découvrait l'informatique.

- **Informatique (Programmation C)** : pour des étudiants en cycle ingénieur première année, dans le département Instrumentation et Commercialisation des systèmes de Mesure (IC2M) ; initiation à la programmation et aux concepts informatiques de base.
- **Informatique (Programmation Scilab)** : pour des étudiants en cycle ingénieur première année, dans le département Génie Biologique et Alimentaire (GB-IAAL) ; initiation à la programmation et aux concepts informatiques de base.
- **Bases de données** : pour des étudiants en cycle ingénieur en première année au département Informatique Micro-électronique Automatique (IMA) et en deuxième année au département GB-IAAL ; aspects formels des bases de données, modélisation des besoins en UML, requêtes SQL, création des applications web en PHP.

Les enseignements que j'ai effectués m'ont permis d'apprendre à mieux communiquer et à adapter mon discours en fonction du parcours hétérogène de mes étudiants. J'ai eu l'occasion d'enseigner des cours destinés à des futurs informaticiens, mais j'ai aussi enseigné à des publics pour qui l'informatique est uniquement un bon outil pour faciliter les traitements nécessaires aux métiers d'ingénieurs ou techniciens auxquels ils se préparaient. Toutes ces expériences m'ont appris l'importance de trouver des exemples simples et intuitifs, de me construire une approche plus claire et pédagogique.

### 3.5 Activités collectives liées à l'enseignement

Depuis mon recrutement en tant que MCF je participe aux diverses activités collectives et administratives de l'IUT d'Orléans, en tant que volontaire ou en portant des responsabilités récompensées par des primes de charges pédagogiques.

- **Membre élue du Conseil d'Administration de l'IUT d'Orléans (06/2018 – 06/2022)** : Je représente les MCF de la composante au sein du conseil (participation à 3-4 réunions par an). Pendant ce mandat, je suis aussi le représentant des départements Informatique et Qlio dans la *Commissions des moyens* (2-3 réunions par an) et la *Commission des statuts* de la composante (2 réunions).

- **Responsable projets tutorés (09/2019 – 09/2021)** pour la licence professionnelle donnée par le département Qlio de l'IUT d'Orléans, qui a des promotions de 20-30 étudiants. Cette charge est récompensée par une prime de responsabilités pédagogiques de l'équivalent à 5h EqTD (incluse dans la fiche de service). Dans le cadre de cette charge pédagogique, les principales missions sont :
  - organiser les trois semaines de projets pour les étudiants (définir le cahier de charges avec les collègues enseignants, définir et évaluer l'avancement des équipes, assurer l'organisation logistique) ;
  - gérer l'équipe pédagogique participant aux projets (animer l'équipe des enseignants (5-6 membres), définir l'ordre d'intervention et assurer la gestion des évaluations).
- **Responsable projets en entreprise (09/2019 – 09/2021)** pour les apprentis de la licence professionnelle donnée par le département Qlio de l'IUT d'Orléans. Cette charge n'est pas rémunérée. Dans le cadre de cette charge pédagogique, les principales missions sont :
  - donner le cadre pour une mission d'une semaine à effectuer par les apprentis dans l'entreprise, sur une mission (i) différente de la mission principale de leur contrat d'apprentissage et (ii) qui met les apprentis dans la position de management ou de pilotage d'un projet ;
  - coordonner l'équipe pédagogique (constituée des tuteurs universitaires des apprentis) pour la validation et l'évaluation des projets.
- **Responsable emplois du temps (03/2018 – 04/2019)** : J'ai pris cette responsabilité au sein du département Qlio de l'IUT d'Orléans. Cette mission de planification est récompensée par une **prime de responsabilités pédagogiques** (PRP) équivalente à 13h EqTD. Ce type de prime n'apparaît pas sur la fiche de service.

Les principales tâches de cette mission sont :

  - établir les emplois du temps hebdomadaires, semestriels de toutes les formations initiales et relevant de la formation continue du département Qlio (soit 3 promotions : DUT1 Qlio, DUT2 Qlio et LP Qlio) ;
  - saisir les emplois du temps sur ADE (logiciel des emplois du temps) et mettre en œuvre le séquençage pour les enseignants du département ;
  - contacter les enseignants extérieurs du département pour trouver des séquençages conformes aux maquettes et à leurs disponibilités, puis saisir leur emplois du temps sur ADE ;
  - réaliser les mises à jour hebdomadaires pour prendre en compte les contraintes non prévues des enseignants titulaires ou vacataires ;
  - organiser les semaines de devoir, projets, avec l'équipe pédagogique et le secrétariat ;
  - participer à la gestion des salles communes avec les collègues et les secrétariats des autres départements ;
  - réaliser les extractions ADE en relation avec le secrétariat et le chef de département et corriger les erreurs éventuelles qui surviennent lors des extractions ADE et des bascules ADE vers GOEPEC (pour la génération des fiches de services).

J'ai arrêté cette responsabilité au début de mon congé de maternité (04/2019).
- **Membre du jury de validation des acquis de l'expérience (2020 – 2021)** pour la licence professionnelle au département Qlio. Étude des dossiers et participation aux auditions pour les demandes de validation des acquis de l'expérience pré-sélectionnées.
- **Responsable de 2-3 modules par an (2017 – )** : coordination des équipes de 3-4 enseignants, préparation du matériel de cours et des évaluations, correction et saisie des notes.
- **Présidente du jury LP (2019 – 2021)** au département Qlio de l'IUT d'Orléans décidant des résultats finaux des étudiants de la promotion.
- **Recrutement des étudiants de première année (2018 – )** : En tant que volontaire, j'ai participé au recrutement en DUT, à savoir, j'ai étudié et noté des dossiers de candidature (≈300 dossiers par an) en DUT Informatique et DUT Qlio. De plus, pour le recrutement en DUT Informatique, j'ai participé à la réalisation d'un test indicatif de positionnement et d'un guide d'entretien, aussi qu'aux entretiens de recrutement.
- **Accompagnement des étudiants de première année (2018 – )** : J'ai fait des *heures de soutien* et des *oraux de consolidation* dans les départements Informatique et Qlio, destinées aux étudiants de première année en difficulté, afin de les aider à mieux comprendre le contenu des cours. J'ai aussi fait un suivi personnalisé des étudiants en difficulté (aide méthodologique) en première année en DUT Qlio et DUT Informatique.
- **Participation à des forums pour les lycéens et aux Journées Portes Ouvertes (2017 – )** : En tant que représentant des départements Informatique et Qlio de l'IUT d'Orléans, j'ai présenté les formations aux candidats potentiels. Participation à un forum par an dans la métropole orléanaise.

## 4 Activités de recherche

Mes contributions se situent principalement dans le domaine du Machine Learning et portent notamment sur l'**apprentissage séquentiel et interactif**. J'ai eu de **contributions théoriques** (*théorie de bandits à plusieurs bras, modélisation d'apprentissage human-in-the-loop, étude de bornes de performance des algorithmes et plan d'expériences*) inspirées par des **applications** telles que la *médecine personnalisée* et les *systèmes de recommandation*. Plus récemment, je m'intéresse aussi à la **sécurité des données** utilisés dans des algorithmes de machine learning, lorsque les données et les calculs sont externalisés dans un cloud et dans le cadre de l'apprentissage fédéré.

Dans cette section, je décris mes activités de recherche depuis mon recrutement en tant que MCF (Section 4.1), celles pendant mon postdoctorat (Section 4.2) et celles pendant ma thèse (Section 4.3). Ensuite, je présente mes activités d'encadrement de la recherche (Section 4.4), les prix et les participations aux jurys de thèse et aux projets de recherche avec financement externe (Section 4.5), ainsi que mes activités collectives liées à la recherche (Section 4.6). Enfin, je donne la liste de mes publications et communications (Section 4.7).

### 4.1 En tant que MCF - recherche à l'Univ. Orléans et à l'Univ. Grenoble Alpes

Depuis septembre 2017 je suis membre du *Laboratoire d'Informatique Fondamentale d'Orléans (LIFO<sup>5</sup>, EA 4022)* en tant que maître de conférences (MCF). Dans cette période, j'ai eu un **congé de maternité** qui a commencé en avril 2019 et j'ai aussi fait une **mobilité thématique majeure** vers la communauté de sécurité informatique, un thème central du LIFO. De plus, entre septembre 2021 et septembre 2023, je suis en mise à disposition, suivie par une *délégation CNRS* dans l'équipe SLIDE du *Laboratoire d'Informatique de Grenoble (LIG<sup>6</sup>, UMR 5217)*. Pendant mon intégration au LIG, j'élargis mes connaissances et compétences, en travaillant sur encore **deux nouvelles thématiques** : l'exploration interactive des données et les interactions entre l'intelligence artificielle et la gestion des données. Ci-dessous les détails sur mes thématiques principales de recherche et les résultats obtenus depuis mon recrutement en tant que MCF.

### Sécurisation des algorithmes de bandits stochastiques

Avec Radu Ciucanu (LIFO) et Pascal Lafourcade (LIMOS), je travaille actuellement sur un projet qui vise à sécuriser des algorithmes d'apprentissage séquentiel. Nous nous intéressons à la sécurisation (dans le sens de protection contre l'accès non-autorisé) des données et des calculs nécessaires dans des problèmes de bandit stochastique à plusieurs bras (*multi-armed bandits*). Ce cadre formalise des problèmes d'allocation de ressources dans des environnements incertains où un agent choisit ses actions séquentiellement. Initialement, l'agent n'a aucune information sur la valeur associée à chaque action, mais après chaque pas, l'environnement fournit une observation bruitée sur la valeur de l'action choisie. L'agent doit utiliser ces observations pour allouer ses ressources de façon optimale.

Notre article [CLLS20] publié à **TrustCom 2020** (Core A) traite le cadre classique des bandits stochastiques dans un objectif de maximisation de la récompense cumulée. Le défi que nous avons considéré est de retourner exactement le même résultat que l'algorithme standard (non-sécurisé), tout en garantissant (i) la protection des données et calculs via le chiffrement des données et l'utilisation du calcul multipartite sécurisé et (ii) un temps de calcul raisonnable. La version longue de ce travail est parue dans la revue **JCS** [CLLS23].

Dans [CLLS19] publié à **ISPEC 2019** nous proposons une version sécurisée pour un des problèmes classiques dans la thématique des bandits à plusieurs bras, celle de *l'identification du meilleur bras avec un budget fini*. Nous avons développé un protocole distribué qui permet de faire des calculs sécurisés dans un cloud public *honest-but-curious* et de renvoyer le meilleur bras à un utilisateur, avec une confiance d'identification corrélée au budget de l'utilisateur. L'algorithme sécurisé a la garantie qu'aucun nœud ne peut apprendre suffisamment d'informations pour déduire le meilleur bras et que seulement l'utilisateur ayant payé le budget aura cette information. Pour arriver à ce résultat, nous avons conçu un protocole qui exploite les propriétés du cryptosystème partiellement homomorphe de Paillier. Suite à l'encadrement d'un stage de M2, dans [CDLS23], paru en **TDSC** (Scimago Q1), nous avons proposé trois autres protocoles pour sécuriser les algorithmes d'identifications du meilleur bras dans un cadre fédéré. Ces nouveaux protocoles permettent de prendre en compte des récompenses stochastiques venant de plusieurs types de distributions et d'utiliser un chiffrement symétrique avancé (AES), ce qui a augmenté significativement les performances computationnelles du protocole initial introduit en [CLLS19]. Lors du même stage de M2 encadré en 2020, nous avons étudié aussi la sécurisations des algorithmes de maximisation de récompense dans le cadre de bandits linéaires. Ce travail a conduit à un article [CDLS20] publié à **ProvSec 2020**.

5. <https://www.univ-orleans.fr/lifo/>

6. <https://www.liglab.fr/>

Le travail sur ce projet de sécurisation d'algorithmes de machine learning a continué en janvier-juillet 2021 lors d'un stage de M2, co-encadré avec Radu Ciucanu et Pascal Lafourcade, sur la sécurisation des algorithmes de bandits dans le cadre d'apprentissage fédéré. Le travail de ce stage est paru dans la revue **JAIR** (Scimago Q1) [CLMS22] et une démonstration du système associé a été présenté à la conférence [MCL+22] à **ICDE 2022** (Core A\*). Nous poursuivons actuellement cette direction de recherche dans le cadre d'une thèse de doctorat à ce sujet.

## Exploration interactive des données

Depuis fin 2020, je participe à une collaboration à cheval entre deux chaires de l'institut MIAI Grenoble Alpes : *Contextual Recommendations in Action - Bridging AI and Real-Life Economics*<sup>7</sup>, porté par Sihem Amer-Yahia (responsable de l'équipe SLIDE du LIG) et *My Way to Health Trajectories Medicine*<sup>8</sup>, porté par Jean-Louis Pépin.

Dans ce travail en collaboration avec Sihem Amer-Yahia (LIG), Radu Ciucanu (LIFO), et Sandrine Da Col, ingénieure au LIG que nous avons co-encadrée, un utilisateur non-expert souhaite explorer un vaste ensemble de données en générant de manière interactive un tableau de bord résumant les données. Plus précisément, un tableau de bord est un ensemble de panels, tandis que chaque panel affiche des statistiques sur une région des données, par exemple le résultat d'une requête de regroupement et d'agrégation, ou le résultat d'un algorithme de clustering. Comme le nombre de panels possibles que l'on peut exprimer sur un ensemble de données est très important, on a besoin d'un moyen efficace pour proposer des panels significatifs à l'utilisateur.

Nous modélisons le problème de proposition de panels à l'utilisateur comme un problème de bandits, afin de trouver un bon compromis entre l'*exploitation* de panels proches de ce que l'utilisateur a déjà sur son tableau de bord et l'*exploration* de panels complètement différents afin de creuser dans des données actuellement inexplorées. D'un autre point de vue, si on externalise une partie de la génération du tableau de bord vers le cloud, il est judicieux d'externaliser des données chiffrées pour les attributs sensibles et ensuite réaliser les agrégats directement dans le domaine chiffré. En résumé, l'objectif est de proposer un **système sécurisé** qui utilise des bandits afin d'aider les utilisateurs non-experts à **explorer efficacement** des données. Nous avons un premier article de démonstration à **CIKM 2021** (Core A) [DCS+21] qui propose un système interactif pour extraire des informations pertinentes des différentes jeux de données.

## Interactions machine learning–data management

Un nouvel axe de recherche que j'ai commencé depuis ma mise à disposition au sein du LIG porte sur les interactions entre *machine learning* (et IA plus en général) et les systèmes de gestion des données. En effet, avec l'avènement de l'IA et de son utilisation dans un large éventail d'applications, la communauté gestion de données s'intéresse de plus en plus à l'étude de ce que la gestion de données peut faire pour l'IA (DB4AI) et de ce que l'IA peut faire pour la gestion de données (AI4DB). Si dans l'exploration interactive des données on approche la partie AI4DB, dans cet axe on s'intéresse à la partie DB4AI et on étudie plus précisément la question de recherche suivante : *Comment les techniques de bases de données peuvent-elles exploiter la structure intrinsèque des données pour obtenir des implémentations plus efficaces des algorithmes de ML ?* Motivés par la popularité croissante du federated learning, avec Sihem Amer-Yahia et Radu Ciucanu, nous nous intéressons à l'implémentation d'algorithmes de bandits linéaires dans SQLite, qui est le moteur de base de données le plus utilisé au monde, intégré dans tous les téléphones portables et la plupart des ordinateurs. A notre connaissance, notre travail est le premier qui s'attaque à l'implémentation d'algorithmes de machine learning dans SQLite et dans les bases de données embarquées en général. Nous avons un premier article court accepté à **EDBT** (Core A) [CSA22], où nous montrons que LinUCB, un algorithme bien connu de maximisation de la récompense cumulative pour les bandits linéaires, peut être implémenté en SQLite standard. De plus, notre étude empirique préliminaire a mis en avant des cas où notre implémentation SQLite est plus rapide qu'une implémentation en Python.

Plus récemment, nous avons aussi co-encadré le stage de M2 de Sotirios Tzamaras (MOSIG, UGA) dans lequel nous avons proposé une version fédéré de l'algorithme Q-learning, un des algorithmes classiques d'apprentissage par renforcement. Nous avons aussi proposé une implémentation SQLite de l'algorithme Q-learning fédéré et à nouveau, nous avons observé que celle-ci est plus rapide qu'une implémentation en Python, notamment pour le cas où les données ont un partitionnement vertical. Ce travail a obtenu une **honorable mention** à la conférence **CIKM 2022** (Core A), où il a été publié [TCSA22] dans le track demo.

7. <https://miai.univ-grenoble-alpes.fr/research/chairs/machine-learning-and-reasoning/contextual-recommendations-in-action-bridging-ai-and-real-life-economics-851786.htm?RH=6499587813011763>

8. <https://miai.univ-grenoble-alpes.fr/research/chairs/health/my-way-to-health-trajectories-medicine--851203.htm?RH=6499587813011763>

## Apprentissage actif et interactif

J'ai initié et animé un groupe de travail portant sur l'apprentissage actif et interactif entre mars et juillet 2018. Ce groupe s'adressait aux membres du LIFO qui travaillent en apprentissage automatique et était composé de dix membres. Le but du groupe de travail était d'étudier les principes et les techniques de base de l'apprentissage actif et interactif et de continuer par la suite avec nos projets de recherche à mener dans ces thématiques.

Avec trois autres membres de ce groupe de travail, je participe au **projet ANR Herelles** (2020 – 2024), notamment sur les objectifs de ce projet collaboratif qui visent : (i) le développement des mécanismes d'interaction moins chronophages avec l'utilisateur expert et (ii) la proposition des méthodes d'extraction et capitalisation des connaissances des utilisateurs.

## Continuation des collaborations d'avant mon recrutement en tant que MCF

- J'ai continué la collaboration avec mon ancienne équipe de postdoctorat à l'Univ. Aalto. Suite à cette collaboration, un article [SPM<sup>+</sup>18] a été accepté à ISMB 2018 et paru dans le journal **Bioinformatics** (Scimago Q1), dans l'édition dédiée aux actes de la conférence.
- J'ai continué également la collaboration avec Emilie Morvant (Mcf. Université Jean Monnet–Laboratoire Hubert Curien), avec qui j'ai co-encadré deux stages de M1, portant sur des techniques d'élicitation d'information des experts et de l'intégration de ces informations aux algorithmes d'apprentissage.

## 4.2 Pendant mon postdoctorat - recherche à l'Univ. Aalto, Finlande

Entre septembre 2015 - septembre 2017, j'ai été postdoctorante au département d'informatique de l'Université Aalto, dans l'équipe *Probabilistic Machine Learning (PML)*<sup>9</sup>, sous l'encadrement du Professeur Samuel Kaski. Mon postdoctorat s'est déroulé dans le cadre d'un projet portant sur la médecine personnalisée, réalisé en collaboration avec l'Institut Finlandais de Médecine Moléculaire (FIMM) qui a fourni les données réelles sur lesquelles nous avons testé une partie de nos algorithmes. Le groupe de recherche auquel j'ai appartenu faisait aussi partie de *Helsinki Institute for Information Technology (HIIT)* et notamment du *Finnish Centre of Excellence in Computational Inference Research (COIN)*, qui a financé mon postdoctorat.

À l'Université Aalto, j'ai travaillé sur plusieurs projets dont le dénominateur commun vient d'un côté (i) de l'intérêt pour des solutions adaptées à la médecine personnalisée et (ii) de l'autre côté, de l'apprentissage par interaction, où les algorithmes doivent être adaptatifs et prendre en compte toute nouvelle information obtenue sur l'environnement. On retrouve dans (ii) l'aspect de prise de décision séquentielle sur lequel j'ai travaillé aussi pendant le doctorat. Ce mariage des deux aspects, à la fois appliqué et théorique correspond bien au type de recherche que je souhaite aussi mener dans la suite de ma carrière.

Un exemple de thématique sur laquelle j'ai travaillé pendant mon postdoctorat porte sur **l'amélioration des prédictions à l'aide du feedback d'un utilisateur expert**, avec un nombre très limité d'interrogations possibles. Le problème ici peut être résumé ainsi : quand on travaille avec des **données de large dimension** et que les observations sont difficilement obtenables, de nouvelles sources d'information sont nécessaires afin de pouvoir obtenir des prédictions viables. Parfois on a des informations sur la structure des données qui peuvent être prises en compte pour une meilleure prédiction ou recommandation. Par exemple, dans le cas des **données médicales**, on peut avoir des informations sur des régularités dans les données (parcimonie, la structure linéaire dans le DNA, etc.). Cependant, quand le nombre d'observations est beaucoup inférieur au nombre des variables à estimer ou quand la structure des données ne permet pas d'obtenir suffisamment d'information exploitable a priori, de nouvelles sources d'informations sont nécessaires pour arriver à des prédictions pertinentes. Les algorithmes que nous avons proposés **élicitent de manière efficace** et utilisent les **retours (implicites ou explicites) des experts humains** en tant que source complémentaire d'informations pour les prédictions ou recommandations. Ce type d'information, si disponible, vient d'habitude à un grand coût, ainsi il est très important de choisir intelligemment l'information à demander aux experts, de façon à obtenir des réponses permettant de réduire au maximum l'incertitude.

Avec Prof. Samuel Kaski et mes collègues doctorants et postdoctorants, nous avons proposé des approches de **modélisation interactive** (*probabilistic interactive modeling*) pour guider le choix séquentiel des questions à poser aux experts afin de maximiser l'information que l'on obtient à partir de leurs réponses. Nos méthodes et leurs applications ont amené à plusieurs publications, dont [DPSK17a] dans le journal international **Machine Learning** (Scimago Q1) et [MSM<sup>+</sup>17] publié dans les actes de la conférence **IUI 2017** (Core A). J'ai également présenté à la communauté francophone [SAK16a] une des contributions pendant le postdoctorat.

9. <http://research.cs.aalto.fi/pml/>

### 4.3 Pendant mon doctorat - recherche à Inria Lille

J'ai réalisé ma thèse dans le domaine de l'**apprentissage automatique**, au sein de l'équipe *Sequential Learning* (*Sequel*<sup>10</sup>), commune à l'Inria Lille et au laboratoire CRISTAL UMR CNRS 9189. Ma thèse a été financée par une bourse du Conseil Régional du Nord - Pas de Calais et par Inria, à travers le projet européen CompLACS (Composing Learning for Artificial Cognitive Systems).

J'ai réalisé ma thèse sous la direction de Rémi Munos (DR Inria et Senior Researcher à Google DeepMind) et le co-encadrement de Alessandro Lazaric (CR Inria).

Ma thèse aborde le problème de la **prise de décision séquentielle**. J'ai étudié notamment le développement et l'analyse des stratégies d'allocation de ressources, en vue d'optimiser de différents critères de performance. Dans ce type de problème, un agent interagit avec un *bandit à plusieurs bras* (*multi-armed bandit*), où les bras représentent plusieurs actions qui s'offrent à l'agent et dont il doit en choisir une à chaque pas. Après le tirage de l'un des bras, l'agent observe une récompense aléatoire provenant d'une distribution inconnue associée au bras tiré. L'agent doit se servir des observations pour décider quel bras tirer par la suite, sachant que le but est typiquement de maximiser la somme des récompenses observées. Ce problème a plusieurs applications et peut formaliser tout problème où un agent cherche à prendre la meilleure décision face à un environnement inconnu. La motivation initiale pour le problème des bandits à plusieurs bras vient de l'allocation optimale des traitements où le but est de maximiser le nombre des patients guéris en choisissant parmi plusieurs traitements possibles dont l'effet est inconnu. D'autres applications plus récentes visent les systèmes de recommandation, où on veut découvrir rapidement les préférences des utilisateurs pour leur proposer les articles qui maximisent la probabilité d'achat.

Pendant la thèse, je me suis intéressée à un problème de bandit où il existe une structure globale de l'espace des bras (ou actions), tel que chaque observation d'un bras peut être utilisée pour estimer les valeurs de l'ensemble des actions de l'environnement. Plus précisément, nous proposons des stratégies d'allocation de ressources qui sont efficaces et adaptées à un environnement caractérisé par une *structure linéaire globale*. Dans ce cadre, avec mes collaborateurs, nous nous sommes focalisés sur les problématiques suivantes :

(i) *L'identification de la meilleure action* (*best-arm identification*) avec une précision donnée et en utilisant un nombre minimum d'observations. Pour ce problème, nous caractérisons la complexité de la tâche d'identification, nous offrons une interprétation géométrique pour savoir quand la précision demandée dans l'identification du meilleur bras a été atteinte. De plus, nous proposons des algorithmes statiques et adaptatifs d'allocation de ressources, dont nous analysons les performances en termes de nombre d'observations nécessaires avant d'atteindre la précision demandée. Ce travail a donné lieu à une publication [SLM14] à *Neural Information Processing Systems NeurIPS 2014* (Core A\*), une des plus compétitives et prestigieuses conférences de la communauté. Ce travail a été réalisé en collaboration avec mes encadrants, Rémi Munos et Alessandro Lazaric.

(ii) Un deuxième problème d'intérêt a été celui de formaliser des allocations séquentielles de ressources (ou bien, des séquences de tirages des bras), dont les observations permettent d'atteindre des objectifs venant de la littérature sur les **plans d'expériences** (*optimal experimental design*), par exemple, maximiser la précision de l'estimation de la valeur des actions possibles. Ce travail a aussi été fait en collaboration avec mes deux encadrants et a donné lieu à l'article [SLM13] dans le workshop international avec comité de sélection **BayesOPT**, co-localisé avec NeurIPS 2013. Dans ma thèse, j'ai étendu l'étude théorique de cette problématique de l'estimation uniforme des valeurs des actions pour minimiser l'erreur moyenne dans l'estimation de tous les bras. Nous avons aussi mené une analyse de performance, en considérant plusieurs modèles de bruit à variance inconnue et dépendante du bras tiré.

(iii) Un troisième problème sur lequel j'ai travaillé pendant ma thèse est celui du *transfert des observations à travers plusieurs problèmes de bandit linéaire*. J'ai commencé ce travail suite à ma **visite de recherche à McGill University**, Montréal, Canada, où j'ai travaillé avec Prof. Joelle Pineau et son étudiant Ouais Alsharif. Sur ce thème du transfert dans les bandits linéaires, avec Alessandro Lazaric, Joelle Pineau et Ouais Alsharif, nous avons un article [SALP14] dans le workshop international avec comité de sélection **MultiTask**, co-localisé avec NeurIPS 2014.

En plus d'avoir publié les résultats de ma thèse dans des manifestations internationales avec comité de sélection, j'ai considéré que c'était important de partager mes résultats avec la communauté francophone. Par conséquent, j'ai aussi présenté une partie de mes contributions aux *Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes (JFPDA 2015)* : l'article sur l'apprentissage par transfert [SALP15] et l'article sur l'identification du meilleur bras [SLM15a], qui a obtenu le **Prix du meilleur article**. De plus, l'article sur l'identification du meilleur bras a aussi été sélectionné pour une présentation orale [SLM15b] à la *Conférence nationale d'Apprentissage (CAp 2015)*.

10. <https://sequel.lille.inria.fr/>



## 4.4 Encadrement d'activités de recherche

Je donne d'abord la liste de mes encadrements *officiels* (par exemple, où mon nom apparaît dans la convention de stage), regroupés par niveau d'expérience en Section 4.4.1 pour les doctorants, en Section 4.4.2 pour les ingénieurs de recherche, en Section 4.4.3 pour les étudiants de M2, et en Section 4.4.4 pour les étudiants de M1. Ensuite, je regroupe en Section 4.4.5 mes autres expériences qui ont mené à des publications en collaboration avec des étudiants pour qui je n'ai pas un rôle officiel d'encadrement. Enfin, en Section 4.4.6, je donne des détails sur mes expériences de coordination de groupes de travail.

### 4.4.1 Doctorants (2, les deux avec publications rank A/A\*)

- *Gael Marcadet* (2021-2024), doctorant à l'Université Clermont Auvergne, LIMOS.  
Sujet : Sécurisation d'algorithmes d'apprentissage fédéré.  
Co-encadrement à 40%, les co-encadrants étant Radu Ciucanu (LIFO) pour 30% et Pascal Lafourcade (LIMOS) pour 30%, qui assure aussi la direction de la thèse. La thèse de Gael fait suite à son stage de M2 effectué à l'Université d'Orléans (LIFO), dont j'ai aussi été la co-encadrante principale.  
Articles ensemble : [MCL<sup>+</sup>22] **ICDE 2022** (Core A\*), [CLMS23] **IJCAI 2023** (Core A\*).
- *Alexandre Rio* (2022-2025), doctorant à l'Université Grenoble Alpes, LIG.  
Sujet : Apprentissage par renforcement distribué et respectueux de la vie privée.  
Co-encadrement à 70%, le co-encadrement académique et la direction de la thèse étant faits par Sihem Amer-Yahia (LIG). Les co-encadrants en entreprise sont Merwan Barlier (Huawei) et Igor Colin (Huawei). J'ai eu le rôle principal dans la définition du sujet de cette thèse et dans l'obtention du financement auprès de l'ANRT.  
Article ensemble : [RBCS23] **ICML 2023** (Core A\*).

### 4.4.2 Ingénieurs de recherche avec publication A/A\* (1, avec publication rank A/A\*)

- *Sandrine Da Col* (01/2021-01/2022), LIG/MIAI (Chairs *Contextual Recommendations in Action* et *MyWay2Health*).  
Sujet : Interactive exploratory data analysis  
Co-encadrement à 50%, avec Sihem Amer-Yahia.  
Article ensemble : [DCS<sup>+</sup>21] **CIKM 2021** (Core A).

### 4.4.3 Stagiaires de M2 (6, dont 3 avec publications rank A/A\*)

- *Sotirios Tzamaras* (02-09/2022), Université Grenoble Alpes, Master of Science in Informatics at Grenoble (MOSIG). Stage au LIG.  
Sujet : Implementing Multi-Armed Bandits in DBMS.  
Co-encadrement à 33%, les co-encadrants étant Sihem Amer-Yahia (LIG) et Radu Ciucanu (LIG).  
Article ensemble : [TCSA22] **CIKM 2022** (Core A).
- *Théo Nommay* (02-07/2022), Université Grenoble Alpes, Master of Science in Informatics at Grenoble (MOSIG). Stage au LIG.  
Sujet : Secure Bandits Algorithms for Recommendations.  
Co-encadrement à 50%, le co-encadrant étant Massih-Reza Amini (LIG).
- *Judy Said* (02-06/2022), Université Grenoble Alpes, Master of Science in Industrial and Applied Mathematics (MSIAM). Stage au LIG.  
Sujet : Bandit Algorithms for Feature Selection.  
Co-encadrement à 33%, les co-encadrants étant Massih-Reza Amini (LIG) et Emilie Devijver (LIG).
- *Gael Marcadet* (01-07/2021), Université d'Orléans, Master Informatique Mobile Intelligente et Sécurisée. Stage au LIFO en collaboration avec le LIMOS.  
Sujet : Sécurisation d'algorithmes d'apprentissage fédéré.  
Co-encadrement à 40%, les co-encadrants étant Radu Ciucanu (LIFO) et Pascal Lafourcade (LIMOS).  
Article ensemble : [CLMS22] **JAIR** (Scimago Q1).



- *Anatole Delabrouille* (11/2019 - 05/2020), Université de Bordeaux, Master Cryptologie et Sécurité Informatique. Stage au LIMOS en collaboration avec le LIFO.  
Sujet : Sécurisation d’algorithmes de Machine Learning séquentiel distribués dans le cloud.  
Co-encadrement à 33%, les co-encadrants étant Radu Ciucanu (LIFO) et Pascal Lafourcade (LIMOS).  
Articles ensemble : [CDLS20] ProvSec 2020 et [CDLS23] TDSC (Scimago Q1).
- *Karel Kaurila* (03-09/2016), Université Aalto (Finlande), Master’s Programme in Life Science Technologies, Major in Bioinformatics.  
Sujet : Interactive learning in personalized medicine.  
Co-encadrement à 80%, le co-encadrant étant Samuel Kaski.  
Karel a continué avec une thèse de doctorat à l’Univ. Helsinki, Department of Mathematics and Statistics.

#### 4.4.4 Stagiaires de M1 (2)

- *Loujain Liekah* (04-06/2018), Université Jean Monnet (Saint-Étienne), International Master on Machine Learning and Data Mining. Stage au Lab. Hubert Curien (LHC) en collaboration avec le LIFO.  
Sujet : Combining Expert Feedback.  
Co-encadrement à 50%, la co-encadrante étant Emilie Morvant (LHC).  
Loujain a continué avec une thèse de doctorat au LIRIS.
- *Arunava Maulik* (04-06/2017), Université Jean Monnet (Saint-Étienne), International Master on Machine Learning and Data Mining. Stage au LHC en collaboration avec l’Université Aalto.  
Sujet : Majority Vote as a Source of Additional Knowledge for Regression.  
Co-encadrement à 33%, les co-encadrants étant Emilie Morvant (LHC) et Amaury Habrard (LHC).

#### 4.4.5 Autres publications de rang A/A\* faites en collaboration avec des doctorants (4)

- Pendant mon postdoctorat (09/2015-09/2017), j’ai travaillé étroitement avec **3 doctorants à l’Université Aalto** : *Muhammad Ammad-Ud-Din*, *Pedram Daei*, *Iiris Sundin* (pour tous le directeur de thèse était Samuel Kaski). Nos travaux ensemble ont conduit à 3 publications de rang Core A/A\*, respectivement Scimago Q1 ([DPSK17a] **Machine Learning Journal**, [SPM<sup>+</sup>18] **Bioinformatics**, [MSM<sup>+</sup>17] **IUI 2017**) et à 9 autres articles dans des conférences ou ateliers internationaux avec comités de lecture.
- Depuis septembre 2017 quand j’ai eu mon poste de MCF, en plus de mes co-encadrements officiels, j’ai travaillé de près avec *Marius Lombard-Platet*, **doctorant à l’Université Clermont Auvergne** (son directeur de thèse étant Pascal Lafourcade). Nous avons 3 publications ensemble, dont 1 de rang Core A : [CLLS20] **TrustCom 2020**.

#### 4.4.6 Coordination de groupes de travail (3)

- **Animation de recherche d’un groupe de 9 personnes (septembre 2015 – juillet 2017)**. Le groupe de recherche PML où j’ai fait mon postdoctorat était composé d’une cinquantaine de personnes (professeurs, assistant-professeurs, postdoctorants, doctorants, étudiants de master). Dès mon arrivée dans le groupe, j’ai **dirigé un des quatre sous-groupes thématiques de PML**, portant sur les *techniques d’extraction d’information à partir de la connaissance des experts*. Ce sous-groupe était formé de 9 personnes (4 doctorants, 3 étudiants de master, un autre postdoctorant, et moi-même). Mon rôle était de suivre l’avancement de chaque membre dans ses projets pertinents pour la thématique du sous-groupe, via des réunions informelles (1 ou 2 fois pas mois). C’était l’occasion de discuter de l’état de l’art de notre thématique commune et de s’entraider (partage des bonnes pratiques, suggestions sur des méthodes et jeux de données à utiliser).
- **InterActive Learning** (03/2018 - 07/2018). Comme précisé dans la Section 4.1, j’ai coordonné ce groupe de travail composé par dix membres du LIFO. Mes tâches ont été d’inciter la participation active des membres dans l’étude d’articles liés à l’apprentissage actif et interactif (en relation avec leurs activités de recherche précédentes) et d’organiser les réunions du groupe de travail.

- **HWFA-bio** (HIIT Wide Focus Area on Interactive Personalized Medicine) (02/2017 - 07/2017). Pendant tout mon postdoctorat, j'ai participé au groupe de travail HWFA-bio, un partenariat entre HIIT (Helsinki Institut for Information and Technology), FIMM (Finnish Insitute for Molecular Medicine), l'Université de Helsinki et l'Université Aalto. Le but de ce projet était de *proposer des systèmes utilisant conjointement des méthodes de machine learning et de l'interaction humain-machine pour aider les docteurs dans la prise de décision*. Pour les six dernières mois de mon postdoctorat, j'ai été la coordinatrice de ce groupe de travail composé d'une vingtaine de membres. Mes tâches ont été de conduire les réunions hebdomadaires (en définir les objectifs et le déroulement), planifier et suivre la progression du groupe, envoyer les notes des réunions et les tâches à faire pour les semaines à venir. J'ai continué le travail avec HWFA-bio même après mon recrutement, ce qui a mène notamment à la publication [SPM+18].

## 4.5 Prix et participations aux jurys de thèse et projets avec financement externe

### Prix

- **Honorable mention at CIKM 2022's Demo Track** <https://www.cikm2022.org/best-paper>.
- **Top 30% best reviewers** de la conférence NeurIPS 2018.
- **Prix du meilleur article JPFDA 2015**, accordé par l'Association Française pour l'Intelligence Artificielle.
- Travel award pour NeurIPS 2014, sponsorisé par Google.
- 2×Travel award pour Workshop Women in ML (WiML), co-localisé avec NeurIPS, en 2013 et 2014.

### Jurys de thèse

- 2023 : Examinatrice de la thèse d'Anastasiia Doynychko, *Multiview Learning with Missing Views and Learning Solutions for Cross-Process Modeling in Semiconductor Manufacturing Industry*, Université Grenoble Alpes, LIG. Thèse CIFRE avec Mentor Graphics. Directeur de thèse Massih-Reza Amini.
- 2022 : Examinatrice de la thèse de Geovani Rizk, *Stochastic Graphical Bilinear Bandits*, Université Paris Dauphine PSL, LAMSADE. Thèse CIFRE avec Huawei. Directeurs de thèse Yann Chevaleyre et Rida Laraki.

### Projets de recherche avec financement externe

- Co-porteuse d'un projet **Emergence** 2022, Laboratoire d'Informatique de Grenoble (LIG). Ce projet renforce les liens entre les équipes de gestion des données (SLIDE) et d'apprentissage machine (Aptikal) au LIG.
- Porteuse d'un projet de thèse (2022-2025) financé par l'ANRT (**CIFRE** avec Huawei).
- Core Member dans la Chaire **MIAI Contextual Recommendations in Action** (depuis 2021).
- Participante au projet **ANR Herelles** (11/2020 - 11/2024).

## 4.6 Activités collectives liées à la recherche

Dès ma première année de thèse j'ai participé à diverses activités collectives liées à la recherche qui ont contribué à ma compréhension du milieu académique.

- **Co-organisation de conférences**
  - Web co-chair pour ECML-PKDD 2022<sup>11</sup>, tenue à Grenoble en septembre 2022.
  - Co-organisatrice de la conférence nationale *Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes (JFPDA 2013)*<sup>12</sup>, qui a eu lieu à Lille en juillet 2013.
  - Organisatrice de la rencontre WiML, événement co-organisé avec ICML 2015 à Lille en juillet 2015. Cet événement a attiré environ 100 participants. J'ai coordonné et réalisé presque en totalité les tâches liées à l'organisation de l'événement :
    - financement : obtention des financements locaux de l'Inria Lille et de l'Université de Lille 1 (Laboratoire CRISTAL) ; l'événement a aussi bénéficié des financements de la part des entreprises tel que Google, Intel, Microsoft, obtenus conjointement avec les organisateurs de WiML 2015 ;

11. <https://2022.ecmlpkdd.org/>

12. <http://pfia2013.univ-lille1.fr/doku.php%3Fid=fr:jfpda.html>

- logistique, programme de l'événement, communication.
- Organisatrice volontaire de la conférence internationale ICML 2015 : participation à l'accueil des participants et à la mise en place de la signalétique.
- **Vice-présidente de la Société Savante Francophone d'Apprentissage Machine (SSFAM)**(2021-).
- **Program Committee Board member** : IJCAI 2022, 2023, 2024 (Core A\*)
- **Membre de 15 comités de programme** (en moyenne, relecture de 5 articles par conférence)
  - ICML 2022 – International Conference on Machine Learning (Core A\*)
  - IJCAI 2017, 2021, 2022, 2023 – International Joint Conference on Artificial Intelligence (Core A\*)
  - NeurIPS 2016, 2018 – Annual Conference on Neural Information Processing Systems (Core A\*)
  - J'ai été parmi les top 30% meilleurs reviewers pour NeurIPS 2018.**
  - AAAI 2018 – AAAI Conference on Artificial Intelligence (Core A\*)
  - ECML-PKDD 2016 – European Conference on Machine Learning (Core A)
  - HITL 2017 – Humain in the Loop Machine Learning Workshop at ICML 2017
  - TSW 2016 – Times Series Workshop at NeurIPS 2016
  - CAp 2023 – Conférence francophone en Apprentissage Automatique
  - EGC 2022 – Conférence francophone sur l'Extraction et la Gestion des Connaissances
  - RJCIA 2020 – Rencontre des Jeunes Chercheur-ses en Intelligence Artificielle
  - JFPDA 2013 – Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes
- **Relectrice pour des revues** : Machine Learning Journal (Core A)
- **Relectrice externe** (relecture d'un article par conférence)
  - COLT 2017 – Annual Conference on Learning Theory (Core A\*)
  - ALT 2016 – Algorithmic Learning Theory (Core B)
  - IEA/AIE 2017 – International Conference on Applied Intelligent Systems
  - AlgoTel 2021 – Rencontre Francophone sur les Aspects Algorithmiques des Télécommunications
- **Membre de comités de sélection/suivi**
  - 2022 : Examinatrice pour *tenure-track associate professor position in ML*, University of Bergen (Norvège).
  - 2021 : 3 × Comité de sélection pour des postes MCF 27 (Univ. Orléans/LIFO, 2×Univ. Artois/CRIL).
  - 2018 : Comité de suivi des doctorants de l'Ecole Doctorale 551 MIPTIS.
  - 2016 : Comité de recrutement des doctorants FIMM-EMBL/HIIT (appel joint aux candidatures fait par Finnish Institute in Molecular Medicine et par Helsinki Institute for Information Technology), en tant que représentante du HIIT, groupe Probabilistic Machine Learning.
  - 2016 : Coordinatrice du recrutement des stagiaires (stages d'été et stage de fin d'études pour la rédaction du mémoire de master) dans le domaine *Probabilistic Machine Learning*, à l'Université Aalto.

## 4.7 Productions scientifiques

Dans la communauté d'**apprentissage automatique (machine learning)** les conférences phare sont NeurIPS<sup>13</sup> (Annual Conference on Neural Information Processing Systems) et ICML (International Conference on Machine Learning), les deux de rang Core<sup>14</sup> A\*, considérées qualitativement égales aux journaux, dont les plus importants sont Machine Learning et JMLR, les deux rang Scimago Q1. Parmi les articles réalisés pendant mon parcours, j'ai notamment un article publié à **NeurIPS**, un article publié à **ICML**, un article dans le journal **Machine Learning** et huit articles dans les workshops internationaux co-localisés avec NeurIPS ou ICML. Typiquement, les articles dans les workshops de la communauté *machine learning* ont des comités de lecture et des actes informels, le but étant surtout de présenter des projets de recherche devant les collègues de la communauté (d'où l'intérêt que les workshops soient co-localisés avec les meilleures conférences du domaine).

J'ai également publié dans des revues/conférences de rang Core A/A\* dans d'autres communautés :

13. En 2018, la conférence a changé de nom de NIPS en NeurIPS. Dans le dossier, j'utilise à chaque fois le nom actuel, même si je fais référence aux éditions d'avant 2018.

14. Pour chaque article, le rang affiché est celui de l'année de publication.

- Sécurité informatique : un article à **TrustCom** et un autre dans la revue **TDSC**.
- Gestion de données : deux articles demo à **CIKM**, un article demo à **ICDE**, un article court à **EDBT**.
- Bioinformatique et biologie computationnelle : un article à ISMB (International Conference on Intelligent Systems for Molecular Biology) dont les actes sont publiés dans le journal **Bioinformatics**.
- Interfaces intelligentes : un article à **IUI**.

La Table 2 donne une vue de synthèse de mes articles d'audience internationale, ci-dessous je donne la liste complète des références pour les papiers d'audience internationale, ainsi que pour ceux présentés aussi dans de conférences francophones. Tous les articles ont été acceptés après évaluation par des comités de lecture. Certains articles sont accompagnés du **code open source**, dont le but principal est de permettre de reproduire les résultats expérimentaux de nos articles.

### Revues internationales avec comité de lecture

1. [CDLS23] R. Ciucanu, A. Delabrouille, P. Lafourcade, and M. Soare. Secure Protocols for Best Arm Identification in Federated Stochastic Multi-Armed Bandits. *Transactions on Dependable and Secure Computing (TDSC)*, 20(2):1378–1389, 2023.  
**TDSC 2023, Scimago Q1**  
DOI : <https://doi.org/10.1109/TDSC.2022.3154585> (accepté en 2022, paru en 2023)  
Code open source : <https://github.com/anatole33/SR-secure>
2. [CLLS23] R. Ciucanu, P. Lafourcade, M. Lombard-Platet, and M. Soare. Secure Protocols for Cumulative Reward Maximization in Stochastic Multi-Armed Bandits. *Journal of Computer Security*, 31(1):1–27, 2023.  
**JCS 2023, Scimago Q2**  
DOI : <https://doi.org/10.3233/JCS-210051> (accepté en 2022, paru en 2023)  
Code open source : <https://github.com/radu1/secure-ucb>
3. [CLMS22] R. Ciucanu, P. Lafourcade, G. Marcadet, and M. Soare. SAMBA: A Generic Framework for Secure Federated Multi-Armed Bandits. *Journal of Artificial Intelligence Research (JAIR)*, 73:737–765, 2022.  
**JAIR 2022, Scimago Q1**  
DOI : <https://doi.org/10.1613/jair.1.13163>  
Code open source : <https://github.com/gamarcad/paper-samba-code>
4. [SPM<sup>+</sup>18] I. Sundin, T. Peltola, L. Micalef, H. Afrabandpey, M. Soare, M. Majumder, P. Dae, C. He, B. Serim, A. Havulinna, C. Heckman, G. Jacucci, P. Marttinen, and S. Kaski. Improving genomics-based predictions for precision medicine through active elicitation of expert knowledge. *Bioinformatics (OUP), Special Issue dedicated to the proceedings of ISMB 2018*, 34(13):i395–i403, 2018.  
**Bioinformatics 2018, Scimago Q1**  
DOI : <https://doi.org/10.1093/bioinformatics/bty257>  
Code open source : <https://github.com/AaltoPML/knowledge-elicitation-for-precision-medicine>
5. [DPSK17a] P. Dae, T. Peltola, M. Soare, and S. Kaski. Knowledge Elicitation via Sequential Probabilistic Inference for High-Dimensional Prediction. *Machine Learning (ML), Special Issue of the ECML-PKDD 2017 Journal Track*, 106(9-10):1599–1620, 2017.  
**Machine Learning 2017, Scimago Q1**  
DOI : <https://doi.org/10.1007/s10994-017-5651-7>  
Code open source : <https://github.com/HIIT/knowledge-elicitation-for-linear-regression>

### Actes des conférences internationales avec comité de lecture

1. [RBCS23] A. Rio, M. Barlier, I. Colin, and M. Soare. Multi-Agent Best Arm Identification with Private Communications. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, page to appear, 2023.  
**ICML 2023, Core A\***
2. [CLMS23] R. Ciucanu, P. Lafourcade, G. Marcadet, and M. Soare. SAMBA: A Generic Framework for Secure Federated Multi-Armed Bandits (Extended Abstract). In *Proceedings of the 32nd International Joint*

- Conference on Artificial Intelligence (IJCAI), Journal Track*, page to appear, 2023.  
**IJCAI 2023, Core A\***
3. [MCL<sup>+</sup>22] G. Marcadet, R. Ciucanu, P. Lafourcade, M. Soare, and S. Amer-Yahia. SAMBA : A System for Secure Federated Multi-Armed Bandits. In *Proceedings of the 38th IEEE International Conference on Data Engineering (ICDE), Demo Track*, pages 3154–3157, 2022.  
**ICDE 2022, Core A\***  
 DOI : <https://doi.org/10.1109/ICDE53745.2022.00286>  
 Code open source : <https://github.com/gamarcad/samba-demo>
  4. [TCSA22] S. Tzamaras, R. Ciucanu, M. Soare, and S. Amer-Yahia. FeReD: Federated Reinforcement Learning in the DBMS. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM), Demo Track*, pages 4989–4993, 2022.  
**CIKM 2022, Core A**  
 DOI : <https://dl.acm.org/doi/10.1145/3511808.3557203>  
 Code open source : <https://github.com/sotostzam/FeReD>  
**Honorable Mention** <https://www.cikm2022.org/best-paper>
  5. [CSA22] R. Ciucanu, M. Soare, and S. Amer-Yahia. Implementing Linear Bandits in Off-the-Shelf SQLite. In *Proceedings of the 25th International Conference on Extending Database Technology (EDBT), Short Paper Track*, pages 388–392, 2022.  
**EDBT 2022, Core A**  
 DOI : <https://doi.org/10.48786/edbt.2022.27>  
 Code open source : <https://github.com/radu1/linucb-sqlite>
  6. [DCS<sup>+</sup>21] S. Da Col, R. Ciucanu, M. Soare, N. Bouarour, and S. Amer-Yahia. DashBot: An ML-Guided Dashboard Generation System. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM), Demo Track*, pages 4696–4700, 2021.  
**CIKM 2021, Core A**  
 DOI : <https://doi.org/10.1145/3459637.3481968>  
 Code open source : <https://github.com/sandrinedacol/DashBot>
  7. [CLLS20] R. Ciucanu, P. Lafourcade, M. Lombard-Platet, and M. Soare. Secure Outsourcing of Multi-Armed Bandits. In *Proceedings of the 19th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pages 202–209, 2020.  
**TrustCom 2020, Core A**  
 DOI : <https://doi.org/10.1109/TrustCom50675.2020.00038>  
 Code open source : <https://github.com/radu1/secure-ucb>
  8. [CDLS20] R. Ciucanu, A. Delabrouille, P. Lafourcade, and M. Soare. Secure Cumulative Reward Maximization in Linear Stochastic Bandits. In *Proceedings of the 14th International Conference on Provable and Practical Security (ProvSec)*, pages 257–277, 2020.  
**ProvSec 2020, Core B**  
 DOI : [https://doi.org/10.1007/978-3-030-62576-4\\_13](https://doi.org/10.1007/978-3-030-62576-4_13)  
 Code open source : <https://github.com/anatole33/LinUCB-secure>
  9. [CLLS19] R. Ciucanu, P. Lafourcade, M. Lombard-Platet, and M. Soare. Secure Best Arm Identification in Multi-Armed Bandits. In *Proceedings of the 15th International Conference on Information Security Practice and Experience (ISPEC)*, pages 152–171, 2019.  
**ISPEC 2019, Core B**  
 DOI : [https://doi.org/10.1007/978-3-030-34339-2\\_9](https://doi.org/10.1007/978-3-030-34339-2_9)  
 Code open source : <https://gitlab-sds.insa-cvl.fr/vciucanu/secure-bai-in-mab-public-code>
  10. [MSM<sup>+</sup>17] L. Micaleff, I. Sundin, P. Marttinen, M. Ammad-ud-din, T. Peltola, M. Soare, G. Jacucci, and S. Kaski. Interactive Elicitation of Knowledge on Feature Relevance Improves Predictions in Small Data Sets. In *Proceedings of the 22nd Annual Meeting of the Intelligent User Interfaces Community (IUI)*, pages 547–552, 2017.  
**IUI 2017, Core A**  
 DOI : <https://dl.acm.org/doi/10.1145/3025171.3025181>



11. [SAK16b] M. Soare, M. Ammad-ud-din, and S. Kaski. Regression with  $n \rightarrow 1$  by Expert Knowledge Elicitation. In *Proceedings of the 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 734–739, 2016.  
**ICMLA 2016, Core C**  
 DOI : <https://doi.org/10.1109/ICMLA.2016.0131>
12. [SLM14] M. Soare, A. Lazaric, and R. Munos. Best-Arm Identification in Linear Bandits. In *Proceedings of the 28th Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 828–836, 2014.  
**NeurIPS 2014, Core A\***  
<http://papers.nips.cc/paper/5460-best-arm-identification-in-linear-bandits>

#### Workshops internationaux avec comité de lecture

1. [DPSK17b] P. Daeë, T. Peltola, M. Soare, and S. Kaski. Probabilistic Expert Knowledge Elicitation of Feature Relevances in Sparse Linear Regression. In *Proceedings of the Workshop and Tutorial on Interactive Adaptive Learning, co-located with ECML-PKDD*, 2017.  
**IAL@ECML-PKDD2017**
2. [SPM+17a] I. Sundin, T. Peltola, M. Majumder, P. Daeë, M. Soare, H. Afrabandpey, C. Heckman, S. Kaski, and P. Marttinen. Ask the doctor - Improving drug sensitivity predictions through active expert knowledge elicitation. In *Workshop on Machine Learning for Health, co-located with NeurIPS*, 2017.  
**ML4H@NeurIPS2017**
3. [DPSK17c] P. Daeë, T. Peltola, M. Soare, and S. Kaski. Probabilistic Expert Knowledge Elicitation of Feature Relevances in Sparse Linear Regression. In *Workshop on Human-in-the-Loop Machine Learning, co-located with ICML*, 2017.  
**HITL@ICML2017**  
**Selected for a contributed talk**
4. [SPM+17b] I. Sundin, T. Peltola, M. Majumder, P. Daeë, M. Soare, H. Afrabandpey, C. Heckman, S. Kaski, and P. Marttinen. Ask the doctor - Improving drug sensitivity predictions through active expert knowledge elicitation. In *Workshop on Computational Biology, co-located with ICML*, 2017.  
**WCB@ICML2017**
5. [SPM+17c] I. Sundin, T. Peltola, M. Majumder, P. Daeë, M. Soare, H. Afrabandpey, C. Heckman, S. Kaski, and P. Marttinen. Ask the doctor - Improving drug sensitivity predictions through active expert knowledge elicitation. In *The Eleventh International Workshop on Machine Learning in Systems Biology, Special Session at ISMB/ECCB*, 2017.  
**MLSB@ISMB-ECCB2017**
6. [SMM+16] I. Sundin, L. Micalef, P. Marttinen, M. Ammad-ud-din, T. Peltola, M. Soare, G. Jacucci, and S. Kaski. Interactive User Intent Modeling for Eliciting Priors of a Normal Linear Model. In *Workshop on Future of Interactive Learning Machines, co-located with NeurIPS*, 2016.  
**FILM@NeurIPS2016**
7. [DPSK16] P. Daeë, T. Peltola, M. Soare, and S. Kaski. Probabilistic Expert Knowledge Elicitation of Feature Relevances in Sparse Linear Regression. In *Workshop on Future of Interactive Learning Machines, co-located with NeurIPS*, 2016.  
**FILM@NeurIPS2016**
8. [SAK16c] M. Soare, M. Ammad-ud-din, and S. Kaski. Regression with  $n \rightarrow 1$  by Expert Knowledge Elicitation. In *Workshop on Data-Efficient Machine Learning, co-located with ICML*, 2016.  
**DEML@ICML2016**  
**Selected for a contributed talk**
9. [SALP14] M. Soare, O. Alsharif, A. Lazaric, and J. Pineau. Multi-task Linear Bandits. In *Workshop on Transfer and Multi-task Learning: Theory Meets Practice, co-located with NeurIPS*, 2014.  
**MultiTask@NeurIPS2014**

10. [SLM13] M. Soare, A. Lazaric, and R. Munos. Active Learning in Linear Stochastic Bandits. In *Workshop on Bayesian Optimization in Theory and Practice, co-located with NeurIPS*, 2013.  
BayesOPT@NeurIPS2013

### Conférences d'audience nationale avec comité de lecture

1. [SAK16a] M. Soare, M. Ammad-ud-din, and S. Kaski. Expert Knowledge Elicitation for “large p, small n” Regression. In *11ièmes Journées Francophones sur la Planification, la Décision et l’Apprentissage*, 2016.  
JFPDA2016
2. [SLM15a] M. Soare, A. Lazaric, and R. Munos. Identification du meilleur bras dans le modèle des bandits linéaires. In *10ièmes Journées Francophones sur la Planification, la Décision et l’Apprentissage*, 2015.  
JFPDA2015  
**Prix du meilleur article**, accordé par l’Association Française pour l’Intelligence Artificielle (AFIA).
3. [SALP15] M. Soare, O. Alsharif, A. Lazaric, and J. Pineau. Transfert séquentiel dans le modèle de bandit linéaire. In *10ièmes Journées Francophones sur la Planification, la Décision et l’Apprentissage*, 2015.  
JFPDA2015
4. [SLM15b] M. Soare, A. Lazaric, and R. Munos. Identification du meilleur bras dans le modèle des bandits linéaires. In *Conférence Francophone sur l’Apprentissage Automatique (CAp), Lille, France*, 2015.  
CAp2015  
**Selected for a contributed talk**

### Manuscrit de thèse

- [Soa15] M. Soare. *Sequential Resource Allocation in Linear Stochastic Bandits. (Allocation séquentielle de ressources dans le modèle de bandit linéaire)*. PhD thesis, Lille University of Science and Technology, France, 2015 <https://tel.archives-ouvertes.fr/tel-01249224>

### Présentations

#### — 27 présentations invitées/séminaires :

*DashBot : An ML-Guided Dashboard Generation System*

- Workshop de l’axe SIDCH du LIG, Grenoble, France, 03/2022
- MIAI Midterm Review, Demo session, Grenoble, France, 03/2022

*Secure Multi-Armed Bandit Computations in the Cloud*

- Séminaire de l’équipe SLIDE - LIG, Grenoble, France, 09/2021
- Séminaire de l’équipe AMA - LIG, Grenoble, France, 01/2020

*Introduction to Stochastic Multi-Armed Bandits*

- ANR Herelles, Kick-off meeting, online talk, 11/2020
- Séminaire du LIFO, Orléans, France, 01/2018

*Secure Best-Arm Identification in Multi-Armed Bandits*

- Séminaire de l’équipe SDS - LIFO, Bourges, France, 04/2019

*Expert Knowledge Elicitation of Feature Relevances in Sparse Linear Regression*

- Présentation au groupe de travail Math-Info, LIFO/MAPMO, Orléans, France, 12/2017
- Séminaire Probabilistic Machine Learning Lab, Aalto University, Helsinki, Finlande, 02/2017

*Sequential Decision Making in Linear Bandits*

- Séminaire de l’équipe DM2L - LIRIS, Lyon, France, 03/2017
- Séminaire de l’équipe QARMA - LIF, Marseille, France, 03/2017
- Séminaire du LIFO, Orléans, France, 03/2017



- Séminaire de l'équipe TAO - LRI, Paris Saclay, France, 03/2017
- Séminaire du laboratoire GREYC, Caen, France, 02/2017
- Séminaire de l'équipe Data Intelligence - Lab. Hubert Curien, Saint-Étienne, France, 10/2016
- Séminaire du LIMOS, Clermont-Ferrand, France, 09/2016
- Séminaire de l'équipe AMA - LIG, Grenoble, France, 09/2016
- Sequential Resource Allocation in Linear Stochastic Bandits*
  - Séminaire Probabilistic Machine Learning Lab, Aalto University, Helsinki, Finlande, 12/2015
- Sequential Transfer of Samples in Linear Bandits*
  - Séminaire de l'équipe SequeL, Lille, France, 06/2015
- Best-Arm Identification in Linear Bandits*
  - Présentation invitée aux Journées YSP (Young Statisticians and Probabilists), Paris, France, 01/2015
  - Statistical Machine Learning Seminar (SMILE), Paris, France, 01/2015
  - Data Intelligence Group Meeting, Lille, France, 01/2015
  - Présentation poster à Women in Machine Learning (WiML), Montréal, Canada, 12/2014
  - Séminaire de l'équipe SequeL, Lille, France, 07/2014
- Pure Exploration in Linear Bandits*
  - Séminaire Reasoning and Learning Lab, McGill University, Montréal, Canada, 02/2014
- Active Learning in Linear Bandits*
  - Présentation poster à Women in Machine Learning (WiML), Lake Tahoe, États-Unis, 12/2013
  - Présentation poster à l'école d'été MLSS, Max Planck Institute, Tübingen, Allemagne, 08/2013
- **15 présentations en conférence/workshop :**  
 pour les articles [CLMS23, TCSA22, DCS<sup>+</sup>21, CLLS20, CDLS20, SAK16b, DPSK16, SAK16a, SAK16c, SALP15, SLM15a, SLM15b, SALP14, SLM14, SLM13]