

CV académique - Dernière mise à jour : avril 2026

Myriam TAMI

Avril 2026

Maître de Conférences CDI en informatique à CentraleSupélec

Laboratoire MICS

Département Informatique

<https://myriamtami.github.io/>

myriam.tami@centralesupelec.fr

Apprentissage statistique & Intelligence Artificielle

Mots clés :

- Recherche :
 - Apprentissage automatique et annotation faible des données
 - Apprentissage statistique avec des données latentes ou incertaines pour l'interprétabilité et la robustesse
 - La causalité pour un apprentissage automatique plus robuste, plus transparent et de confiance
- Enseignement :
 - Cours de 1A à 3A et MSc IA (M2) principalement en apprentissage statistique et IA
 - Responsable/co-responsable de cours : Recherche d'Information (1A), Apprentissage Automatique (3A), Ensemble Learning (MSc IA), Avancées de l'IA (3A)
 - Coordination de projets de recherche du MSc IA et encadrement de projets étudiants
 - Participation à la formation continue en Machine Learning et Deep Learning (CentraleSupélec Exed)
- Responsabilités collectives :
 - Co-directrice du MSc IA de CentraleSupélec ; membre du comité de pilotage HubIA
 - Porteuse d'un projet ANR JCJC et de contrats industriels
 - Service à la communauté scientifique internationale : Senior Area Chair (UAI 2026), reviewer pour conférences majeures en IA (NeurIPS, CVPR, IJCAI, etc.)
 - Responsabilités institutionnelles au MICS (référente développement durable et industrie 4.0)
 - Animation scientifique et diffusion : summer school IA, SFdS (groupe MALIA), workshops fondements mathématiques de l'IA, groupe Causal-Tau, Quarter Causality, actions de médiation scientifique
 - Engagement académique : comités de recrutement EC, jurys de thèses, Comités de suivi de thèse, mentorat femmes en sciences

Table des matières

1	CV court	2
1.1	Parcours professionnel	2
1.2	Cursus académique	2
2	Activités d'enseignement	4
2.1	Responsabilités pédagogiques et collectives à CentraleSupélec	5
2.1.1	Responsabilités d'U.E. et implication pédagogiques	5
2.1.2	Co-responsable académique du MSc IA de CentraleSupélec	6
2.1.3	Co-organisatrice et Référente Académique - Summer School en IA de CentraleSupélec	7
2.1.4	Autres activités collectives pré-pédagogiques	7
2.2	CentraleSupélec (2019 - présent)	8
2.2.1	Activités particulières et initiatives pédagogiques	8
2.2.2	Principaux cours et TD/TP	8
2.2.3	Projets d'étudiants et stages	9
2.2.4	Description et autres initiatives pédagogiques spécifiques à ces cours	10
2.2.5	Implication sur la formation continue CentraleSupélec Exed en tant que vacataire	13
2.3	Université Grenoble Alpes (2017-2019)	14
2.4	Université de Montpellier (UM) (2012-2016)	15
2.5	Bilan et charge globale en HETD	16
3	Activités de recherche	17
3.1	Orientation et axes de recherche	17
3.1.1	Résumé général et contexte	17
3.1.2	Positionnement de mon travail : vers une IA robuste, causale et digne de confiance	17
3.1.3	Axe 1 : Apprentissage statistique sous contrainte de supervision limitée	18
3.1.4	Axe 2 : IA causale pour un apprentissage robuste, interprétable et digne de confiance	19
3.2	Publications et productions scientifiques	20
3.3	Principales collaborations académiques et industrielles	23
3.4	Encadrements de thèses	24
3.5	Encadrements de postdoctorants	25
3.6	Encadrements de stagiaires de M1-M2 ou assimilés	26
3.7	Projets financés	27
3.8	Responsabilités collectives en recherche et animation scientifique :	27
3.8.1	Activités éditoriales	27
3.8.2	Participation à des comités de programmes	28
3.8.3	Participation à un réseau de recherche et consortium	28
3.8.4	Invitations séminaires et séjours de recherche	28
3.8.5	Expertise académique : comités et jurys	29
3.8.6	Diffusion logicielle et science ouverte	29
3.8.7	Diffusion du savoir et médiation scientifique : responsabilités et activités au sein de sociétés savantes ou associations	30
3.8.8	Activités institutionnelles pour mon laboratoire de recherche MICS	30
3.9	Récapitulatif des publications scientifiques	31

1 CV court

Myriam TAMI

Enseignante-chercheure CDI Laboratoire MICS
Ecole CentraleSupélec
Université Paris-Saclay
8-10 rue Joliot Curie
91190 Gif-sur-Yvette
myriam.tami@centralesupelec.fr
<https://myriamtami.github.io/>

Informations personnelles :
Née le 04/06/1989 à Nîmes (30)
Française
35 avenue du Général de Gaulle
D7 Le Clos d'Alençon
91140 Villebon-sur-Yvette

1.1 Parcours professionnel

Fev. 2019 – présent

Maître de conférences (MdC) CDI, HDR (depuis 2025) à CentraleSupélec, Université Paris-Saclay. Laboratoire Mathématiques et Informatique pour la Complexité des Systèmes (MICS) (EA 4037), équipe GEMILA. Département Informatique.

Faits marquants : Porteuse d'un projet ANR JCJC et de contrats industriels; Senior Area Chair UAI 2026; Co-directrice MSc AI CentraleSupélec; Encadrement : 8 doctorants (5 ont soutenus), 2 post-doctorants

Mar. 2017 – Juil. 2018

Vacataire en informatique à l'UFR IM2AG de l'Université Grenoble Alpes.

Mar. 2017 – Jan. 2019

Post-doctorante en IA au LIG, supervision : E. Gaussier (Pr. à l'Université Grenoble Alpes) et de M. Clausel (Pr. à l'Université de Lorraine, IECL)

Sep. 2015 – Aoû. 2016

ATER au département de mathématiques de l'Université de Montpellier.

Sep. 2012 – Juil. 2016

Doctorante en biostatistique allocataire de recherche. Université de Montpellier, école doctorale I2S, Laboratoire IMAG.

Sep. 2012 – Sep. 2015

Monitrice département Mathématique. Université de Montpellier.

1.2 Cursus académique

- **19 Mar. 2026 Qualification aux fonctions de Professeur des universités en section 27.**
- **19 Mar. 2026 Qualification aux fonctions de Professeur des universités en section 26.**
- **21 Nov. 2025 Habilitation à diriger des recherches (HDR) en section 27**, Informatique, soutenue le 21 novembre 2025, Université Paris-Saclay, ED Interfaces. Habilitation effectuée au laboratoire MICS sous le tutorat de M. Clausel (Pr., Univ. de Lorraine). Titre : Vers un apprentissage statistique robuste et de confiance : des labels limités à la causalité.
Jury :

Massih-Reza Amini Pr., HDR, Université Grenoble Alpes, Laboratoire d'Informatique de Grenoble.	Rapporteur
Arthur Charpentier Pr., HDR, Université du Québec à Montréal (UQAM), Faculté des Sciences.	Rapporteur
Amaury Habrard Pr., HDR, Université Jean Monnet (UJM) de Saint-Etienne, Laboratoire Hubert Curien UMR CNRS 5516	Rapporteur
Vincent Mousseau Pr., HDR, Université Paris-Saclay, CentraleSupélec, laboratoire MICS	Président
Marianne Clausel Pr., HDR, Université de Lorraine, laboratoire Institut Elie Cartan de Lorraine (IECL)	Examinatrice
Julie Josse Chercheuse Senior (DR) Inria, Antenne de Montpellier	Examinatrice
Stéphane Chrétien Pr., HDR, Université Lyon 2, Laboratoire ERIC, ERIC Laboratory and UFR ASSP	Examinateur

- **Nov. 2024** Obtention de la **RIPEC C3**¹
- **Feb. 2017** **Qualification aux fonctions de Maître de Conférences en section 26** Mathématiques Appliquées, numéro : 17226285155.
- **Juil. 2016** **Doctorat en biostatistique**, soutenance le 12 juillet 2016, Université de Montpellier, école doctorale I2S. Thèse effectuée à l'IMAG sous la direction de C. Lavergne (Pr., Université Paul-Valéry) et de X. Bry (MdC, Université de Montpellier). Titre : Approche EM pour modèles multi-blocs à facteurs à une équation structurelle.

Jury :

Gilbert Saporta Pr. émérite, CNAM, Paris	Rapporteur
Jérôme Saracco Pr., Institut Polytechnique de Bordeaux	Rapporteur
Benoîte de Saporta Pr., IMAG, Université de Montpellier	Présidente
Laura Trinchera MdC, Neoma Business School	Examinatrice

- **Sep.2012** **Master Mathématiques Biostatistique**, Université Montpellier. **Stage de fin d'étude (6 mois), laboratoire IMAG et Danone Research, Montpellier et Palaiseau.** Sujet : Développement méthodologique et computationnel pour l'analyse de données nutritionnelles ainsi que pour la création d'une interface R. L'objectif appliqué était de rapprocher des profils alimentaires défavorables d'une zone cible considérée comme optimale. Le travail s'appuyait sur des méthodes à noyaux et sur des intervalles de confiance. (Encadrants : X. Bry et Pr. A. Gannoun)
- **2008-2010** **Licence de Mathématiques Fondamentales et Appliquées**, Université Montpellier
- **2007 - 2008** **Classes préparatoires en physiques chimie et sciences de l'ingénieur**, Lycée Alphonse Daudet, Nîmes
- **2007** Obtention du **prix de la vocation scientifique et technique des filles**
- **Juin 2007** **Baccalauréat S spécialité physique**, mention B, Lycée Charles Gide, Uzès

Compétences

Langues : Anglais, Marocain dialectal, Espagnol

Programmation : Python, Cython, R, SAS, Scilab/Matlab

Compétences informatiques complémentaires : HTML, LaTeX, Excel, Open office, Word

Formations : Formation contre les violences sexistes et sexuelles (2h) ; Encadrement doctoral (formation dispensée par l'ABG – Association Bernard Gregory – sur 2 jours) ; Prévention, détection et prise en charge des problèmes de santé physique et mentale des doctorants (1h30) ; Interculturalité : mieux comprendre les différences (1h30) ; Innovation et deeptech (1h30) ; Éthique de la recherche et intégrité scientifique (3h) ; Science ouverte (1h30) ; Formation à la pédagogie des maîtres de conférences (32h) avec pour ateliers suivis : Alignement pédagogique, Handicap, Motivation et dynamique de groupe, Rendre les étudiants actifs, Interculturalité et égalité, Évaluation des apprentissages, Positionnement et développement professionnel, Captiver! (animation d'un cours en anglais), Scénariser un enseignement à distance ou hybride.

1. dispositif propre au système français d'enseignement supérieur et de recherche, est une prime individuelle attribuée sur la base d'une évaluation de la qualité des activités de recherche.

2 Activités d'enseignement

Ce chapitre décrit mon activité d'enseignement aux universités de Grenoble Alpes et Montpellier (succinctement) ainsi qu'à CentraleSupélec depuis ma prise de poste de MdC CDI à CentraleSupélec. La description de mes activités d'enseignement est organisée de la manière suivante. Tout d'abord deux tableaux 2.1 et 2.2 synthétisant les Unités d'Enseignements (U.E.) que j'ai dispensées.

La sous section 2.1 présente une partie des responsabilités collectives dans lesquelles je me suis investie et liées à mes activités d'enseignement.

Ensuite, la sous section 2.2 présente de façon plus exhaustive chaque U.E. dans laquelle j'interviens ou bien je suis intervenue à CentraleSupélec dont celles liées aux enseignements par projets et décrites dans le tableau 2.3 ainsi que mes implications en tant que tutrice pédagogique ou bien participation à des soutenances de projets dans le tableau 2.4. Un résumé des contenus pédagogiques de chaque U.E. y est présent ainsi que le nombre d'heures et la nature de l'enseignement délivré (TD, TP, cours, examens...). La sous-section suivante 2.3 présente les deux U.E.s dans lesquelles je suis intervenue à l'Université Grenoble Alpes en tant que vacataire en informatique. Puis, la sous-section 2.4 présente chaque U.E. dans laquelle je suis intervenue à l'Université de Montpellier en tant qu'ATER et monitrice. Enfin, la dernière sous-section 2.5 fait un bilan sur ma charge d'enseignements.

Années	Activité	Lieu d'exercice
2019 - 2026	MdC	Départements d'informatique et de mathématiques, CentraleSupélec
2021 - 2026	Vacataire	Formation continue, CentraleSupélec Exed
2017 - 2018	Vacataire	UFR IM2AG (Informatique, Mathématiques, Mathématiques Appliquées de Grenoble), Université Grenoble Alpes
2015 - 2016	ATER	Département de mathématiques, UFR Sciences, Université de Montpellier
2012 - 2015	Monitrice (docteurante)	Département de mathématiques, UFR Sciences, Université de Montpellier

TABLE 2.1 – Synthèse globale de mon activité d'enseignement en fonction des lieux d'exercices.

Dans la formation initiale du cursus CentraleSupélec où la 1A, 2A et 3A correspondent respectivement aux années de formation du parcours ingénieur de niveau L3, M1 et M2. Une mention correspond à un parcours de spécialisation des étudiants. CM signifie Cours Magistral et TD/TP signifie Travaux Dirigés ou Pratiques. Pour faciliter la lisibilité du tableau 2.2, les enseignements par projets délivrés à CentraleSupélec ainsi que les cours de la formation continue pour CentraleSupélec Exed feront l'objet d'autres tableaux dans les sous sections des enseignements associés à CentraleSupélec.

Matière	Niveau	Volume	Nombre	Années
Apprentissage automatique	3A mention IA	144h(CM), 66h(TD/TP)	60 à 80(CM), 24(TD)	2020-prés.
Ensemble Learning	M2 MSc IA	288h	100(CM), 34(TD) (4 en 2025-26)	2020-prés.
Recherche d'Information	1A	27h(CM), 58.5h(TD/TP)	120(CM), 40(TD)	2019-prés.
Coding Weeks	1A	96h	20 à 40	2021-2022
Statistique et Apprentissage	1A	18h	40 TD	2023
Algorithmes pour le traitement de données	M1 Informatique	3h TD, 10h TP	25	2017 - 2018
Informatique et Aléatoire	L3 Informatique	5.5h CM, 2h TD, 7.5h TP	25	2017 - 2018
Probabilités et Statistique	L3 Mathématiques Fondamentales	39h TD	20	2015 - 2016
Biostatistique	L2 Polytech	15.5h TD, 10h TP	40	2015 - 2016
Analyse	L2 Polytech	36h TD	40	2015 - 2016
Probabilités élémentaires	L2 Mathématiques Fondamentales	13.5h TD	20	2015 - 2016
Mathématiques et Statistique pour l'écologie	L2 Géoscience	22h CM, 75.5h TD, 78h TP	30	2012-2016
Statistiques descriptives et probabilités	L2 Informatique	36h CM, 51h TD	25 TD, 50 CM	2015-2016

TABLE 2.2 – Principaux cours donnés dans les différents lieux d'exercices (séparés par des doubles lignes).

2.1 Responsabilités pédagogiques et collectives à CentraleSupélec

2.1.1 Responsabilités d'U.E. et implication pédagogiques

Je suis **responsable de trois enseignements en IA de niveau M2, dont deux dans une formation internationale (MSc)**. Ces responsabilités comprennent :

- la **création des ressources pédagogiques** (cours, TD, TP, examens, projets),
- la **mise à jour annuelle des supports** et leur diffusion sur la plateforme Edunao,
- l'organisation pédagogique (gestion des devoirs et projets, questionnaires à but pédagogique ou bien d'évaluation, corrections des TDs, TPs et des examens),
- la structuration des modalités d'évaluation (création de **grilles d'évaluation** utilisées en équipe pédagogique afin d'évaluer à la fois par une note et par un système de validation ou non de compétences. J'actualise annuellement ces grilles pour à la fois faciliter le travail d'évaluation en équipe et permettre un échange constructif avec les élèves après évaluation).

En tant que responsable d'UE, j'ai également assuré :

- la **constitution d'équipes pédagogiques** (recrutement de vacataires, suivi administratif),
- l'accompagnement des intervenants pour la préparation des enseignements,
- la participation lors de trois rentrées scolaires consécutives à la **construction des enseignements du nouveau cursus CentraleSupélec** suite à la fusion de Centrale Paris et Supélec lors de l'année de mon recrutement. Voici quelques exemples de cours pour lesquels je me suis investie : (ateliers innovation, cours RI, EI), Apprentissage Automatique, Ensemble Learning (cf. section "Activités particulières et initiatives pédagogiques" plus bas et le descriptif des UE).

Coordination de l'UE Lab Research Projects (MSc IA) Je coordonne l'UE **Lab Research Projects**, dispositif immersif visant à placer les étudiants en situation réelle de recherche académique.

Cette responsabilité comprend :

- la **constitution d'un portefeuille de projets** en lien avec des chercheurs en IA de mon réseau,

- l'accompagnement scientifique des encadrants (calibration des sujets, définition des jalons avec difficulté progressive),
- l'organisation pédagogique (séance de lancement, affectation des groupes),
- le **suivi des projets sur six mois** avec soutenance intermédiaire (afin de suivre l'avancement et réorienter les travaux si nécessaire.),
- l'organisation de la soutenance finale et la coordination des évaluations avec les chercheurs encadrants (où les livrables sont un rapport, le code et une présentation/poster),
- la mise en place d'une **session poster ouverte aux partenaires académiques et industriels** pour valoriser les travaux des étudiants.

Cette UE représente une responsabilité importante de coordination scientifique et pédagogique et contribue activement à l'émergence de vocations pour la recherche.

2.1.2 Co-responsable académique du MSc IA de CentraleSupélec

Depuis septembre 2024, j'assume la **co-direction du MSc international en Intelligence Artificielle de CentraleSupélec**. La prise de fonction, dans un contexte de restructuration de l'équipe administrative, s'est traduite par une charge de travail particulièrement soutenue. Cette responsabilité a représenté la première année environ 50% de mon temps de travail de l'année, bien qu'assortie d'une décharge de 38 heTD par an de mon service d'enseignement.

Responsabilités académiques et pédagogiques Mes missions couvrent l'ensemble des dimensions académiques du programme :

- pilotage pédagogique (plannings, coordination des enseignements, évolution de la maquette),
- coordination des stages, validation des sujets, signatures et suivi académique des étudiants (avec une partie affectée de façon collégiale à des membres de l'équipe enseignante),
- création d'une **table ronde annuelle en IA en anglais** avec partenaires industriels et alumni,
- participation à l'organisation des cérémonies académiques (remise des diplômes, cérémonie de clôture),
- organisation annuelle du **conseil de perfectionnement** réunissant un professeur extérieur et l'ensemble de l'équipe pédagogique, et j'assume le traitement et la prise en compte des retours issus des évaluations des étudiants dans une logique d'amélioration continue de la formation.
- organisation du **comité de pilotage (COPIL)** afin de garantir l'alignement stratégique du MSc IA avec les orientations de l'établissement.

Je coordonne également l'UE Lab Research Projects (cf. section précédente).

Responsabilités administratives et admissions

- gestion des admissions (8 sessions de jury par an),
- traitement direct d'environ **150 candidatures en 2024–2025** dans le cadre d'un processus encore peu structuré (absence de pré-sélection administrative systématique), impliquant de nombreux échanges directs avec les candidats,
- organisation et conduite des entretiens d'admission,
- coordination avec les services administratifs et présentation du MSc aux nouvelles équipes recrutées en 2025 et 2026,
- prise en charge temporaire de missions de scolarité en 2025–2026 (relevés de notes, attestations, suivi étudiants) à la suite du départ de la personne en charge du suivi de la scolarité. J'ai ensuite plus récemment été renforcée par Géraldine Aude depuis sa nouvelle prise de fonction en tant que Directrice des Etudes des programmes de Masters fin janvier 2026.

Développement et rayonnement international

- participation au développement de partenariats académiques internationaux,
- actions de promotion (webinaires, vidéos, communication internationale),
- collaboration avec un cabinet de communication pour structurer la stratégie internationale,
- représentation institutionnelle (Inde, Maroc, Chine) pour la création de partenariats et de double diplômes avec de nouvelles institutions,
- interview prévue dans le Figaro Étudiant (avril 2026).

Structuration et professionnalisation du programme Depuis juin 2025, je collabore avec Guillaume De Lacoste et son équipe afin de structurer le fonctionnement du MSc et mieux l'articuler avec les services institutionnels (admissions, scolarité, relations internationales, entreprises).

L'objectif est de faire évoluer une organisation historiquement artisanale vers un modèle plus robuste adapté à la croissance du programme.

Résultats structurants obtenus (2024–2025) Cette implication a contribué à plusieurs évolutions significatives :

- **+30% d'admissions**,
- augmentation de la part des **étudiantes** (de **18,2% à 33%**),
- diversification internationale (de 7 à 10 nationalités),
- hausse des étudiants internationaux (de 50% à 64%),
- lancement de **trois doubles diplômes internationaux**,
- création d'un cours AI for Finance,
- renforcement des enseignements AI & Ethics,
- évolution du règlement des études en collaboration avec les juristes de l'École.

Ces actions ont également permis le montage d'un projet de recherche international avec des partenaires indiens.

Initiatives pour l'attractivité et l'inclusion J'ai également contribué à renforcer la dynamique de la formation via :

- organisation d'événements structurants (session posters, rencontre alumni, journée d'intégration),
- mise en place, en collaboration et avec le soutien de Franck Richecoeur, d'un **dispositif de bourses basé sur l'excellence et les critères sociaux**. Cette démarche répond à ma volonté de renforcer l'accessibilité d'une formation payante à des étudiants brillants, y compris disposant de ressources limitées, en cohérence avec les orientations stratégiques de l'École.
- collaboration avec Campus France pour l'accompagnement d'étudiants internationaux,
- accompagnement individuel d'étudiants internationaux (ex : Grèce, Kenya),
- démarches auprès de l'Ambassade de France au Maroc pour identifier des financements (déplacements et bourses pour les futurs candidats marocains).

2.1.3 Co-organisatrice et Référente Académique - Summer School en IA de Centrale-Supélec

L'été 2023, j'ai co-organisé la summer school en IA de l'École destinée à un public international de niveau L3 : conception du programme, identification d'intervenants vacataires, coordination de l'équipe pédagogique, sélection des candidats, promotion de l'événement et enseignement d'un cours d'introduction à l'IA. J'ai aussi contribué au renforcement des collaborations internationales grâce à l'implication d'intervenants issus d'universités partenaires.

2.1.4 Autres activités collectives pré pédagogiques

J'ai été sollicité pour **participer en avril 2026 à une table ronde organisée par le groupe de travail sur la pédagogie** de CentraleSupélec intitulée "**Atelier pédagogique : Evaluer à l'ère de l'IA a-t-il encore du sens ?**".

J'ai également commencé à contribuer à des actions collectives de promotion et de diffusion autour de l'intelligence artificielle. Par exemple, j'ai **participé en janvier 2026 à une table ronde intitulée « Placer l'humain au centre : comment accompagner les transformations induites par l'intelligence artificielle ? »**, aux côtés de Frédéric Pascal, Martin Labrousse, Pierre-Frédéric Rouberties, François Dangu et Vincent Lauriat. Ces activités s'inscrivent naturellement dans mes engagements autour des questions d'IA responsable, de formation et de place des femmes dans les sciences, qui sont aussi des thématiques importantes pour les initiatives collectives en IA.

2.2 CentraleSupélec (2019 - présent)

Depuis mon recrutement en février 2019 en tant que Maître de conférences à CentraleSupélec (CS), j'ai assuré un volume total de **1218 h équivalent TD (heqTD)** en formation initiale, complété par **166,5 h** en formation continue (CS Exed), avec une décharge totale de **263 heqTD** liée à mes responsabilités pédagogiques et à l'obtention d'un projet ANR JCJC. Mon service annuel moyen s'élève à **225 h heqTD**. J'enseigne en français et en anglais, aux niveaux **L3 à M2**, dans les départements d'informatique et de mathématiques, auprès de publics variés (cycle ingénieur, MSc Artificial Intelligence, formation continue). Mon investissement pédagogique reflète ma double compétence à la fois en Intelligence Artificielle et en mathématiques. En effet, j'ai pu prendre part, proposer ou construire différentes Unités d'Enseignements (U.E.) dans ces deux champs disciplinaires. A la suite de la fusion de Centrale Paris et de Supélec, en 2019, une nouvelle maquette du cursus ingénieur a été mise en place et cela sur les trois années du cursus. Dans ce cadre, j'ai participé à la création et la mise en place de plusieurs Unités d'Enseignements (U.E.) et activités pédagogiques. Je précise que mes activités pédagogiques en lien avec mes activités de recherche seront présentées de façon spécifique dans la section portant sur mes activités de recherche.

2.2.1 Activités particulières et initiatives pédagogiques

- En 2019, j'ai aidé Wassila Ouerdane et Céline Hudelot (responsables pédagogiques) à la mise en place et coordination d'un ensemble d'UE avec une nouvelle modalité pédagogique, intitulée « séquence thématique (ST) Data@web : valorisation des données du web » en L3. Celle-ci dure de 7 à 8 semaines et englobe 3 U.E.s : un cours classique nommé « Recherche d'Information et traitement de données massives » (RI) avec des Cours Magistraux (CM) et des TD /TP (31,5h) mais aussi, une U.E. d'introduction (12h) et un module nommé Enseignement d'Intégration (EI) (27h). Le module d'introduction permet une mise en contexte du cours et de la problématique ingénieur qui est abordée à travers différentes modalités : conférences introductives, ateliers d'innovation, etc. Le module EI, a pour format un Web Dataathon et, est organisé sur une semaine avec l'appui d'un partenaire industriel, il confronte les étudiants à une problématique d'ingénieur face à un client. Dans ce cadre, j'ai piloté un atelier d'innovation de la partie introductive (2020-2023) ; j'ai organisé et encadré l'un des EI (2020-2022) sous la forme d'un Web Dataathon (avec Céline Hudelot : « e-marketing » de 2020 à 2021 avec pour partenaire Procter & Gamble, puis seule : « e-marketing » avec Rakuten en 2022) ; J'ai construit deux chapitres de cours (contenu, slides, chapitres manuscrits), trois TPs sous le formats jupyter notebook, deux TD, les parties d'examen associés à ces chapitres et TD/TPs ainsi que les rattrapages et leurs corrections.
- En 2020, à CS, j'ai mis en place une nouvelle U.E. nommée « Apprentissage Automatique » (36h) en dernière année du cursus ingénieur (M2) dans la spécialité Intelligence Artificielle. Enfin, dans le cadre du Msc Artificial Intelligence (M2) de CS, à partir de 2020, j'ai mis en place deux nouvelles U.E., Intitulées « Ensemble Learning » (36h) et « Lab research projects » (15h de responsabilité + 6h de suivi de projet par groupe d'étudiants). Ces enseignements sont anglais pour un public international.

Comme chaque enseignant, j'ai été confrontée à la crise de la Covid-19 et ses contraintes lourdes sur les modalités d'enseignement. Les périodes d'enseignement à distance ou hybride (la moitié des étudiants étant présents, l'autre étant en ligne) m'ont amené à la mise en place de sessions en visioconférences. La démarche de classe inversée, où les étudiants préparent à l'avance un sujet et ont l'initiative de la séance de travaux dirigés s'est avérée souvent pertinente dans cette configuration. Plus précisément, pour le cours de RI, j'ai mis en place pour deux chapitres : des vidéos de cours avec montage vidéo, des wooclaps et des padlets pour rendre le suivi du cours plus aisé. J'ai aussi mis en place des vidéos et un certain nombre d'outils numériques (slack, groupes Teams avec salles virtuelles) pour les cours où je suis responsable : Apprentissage Automatique, Ensemble Learning et Lab Research Projects. J'ai fait de même pour les UE où j'intervenais tels que Coding Weeks.

2.2.2 Principaux cours et TD/TP

Les principaux cours et TD/TP dans lesquels je me suis impliquée depuis ma prise de poste à Centrale-Supélec sont décrits dans les 5 premières lignes du tableau 2.2 et complétés par les tableaux 2.3 et 2.4 par mes implications respectives dans des suivis de projets et en tant que tutrice pédagogique ou bien encadrante

pour des groupes d'étudiants. Il est à noter que l'investissement pédagogique pour un suivi de projet est plus élevé que celui d'un tutorat. Les descriptions des contenus de ces U.E.s sont présentes dans la section 2.1.3 qui suivra les tableaux.

2.2.3 Projets d'étudiants et stages

J'encadre chaque année plusieurs travaux de groupes projets en tant que tutrice pédagogique ou bien encadrante. En voici le détail dans les tableaux 2.3 et 2.4.

Année	Lab re- search pro- jects (MSc IA, M2)	Projets tuto- rés IA (M1)	Projets tutorés DataS- cience (M1)	Projets tutorés Modé- lisation et santé (M1)	Digital Tech Year (M1)	Avancés et réa- lités de l'IA (M2)	Ensei- gnement d'in- tégration Data@Web (L3)	Ateliers inno- vation Data@web (L3)
Fev.-Juil. 2019	NA	0	0	0	0	NA	9	6
2019-20	NA	2	0	1	0	8	9	6
2020-21	1	1	0	2	10	8	9	6
2021-22	1	0	0	0	0	8	9	6
2022-23	0	1	3	1	0	NA	0	6
2023-24	1	0	0	0	0	NA	0	0
2024-25	1	0	0	0	0	NA	0	0
2025-26	1	0	0	0	0	0	NA	0
Volume (heqtd) Total	75h (orga.) 30h (suivi)	30h	20h	60h	15h	63h	108h	33h

TABLE 2.3 – Encadrement de groupes d'étudiants (2 à 3) en fonction des U.E.s de type suivis de projets dans le cursus CentraleSupélec. NA signifie que le cours n'était pas encore ouvert ou bien qu'il a fermé.

Année	SIA/SIP/SIE (M1)	Tuteur pro- jets alternants (M2)	Tuteur stage fin d'étude (MSc IA, M2)	Jury soute- nance projets (M1, M2)	Tuteur pro- jet entreprise (M2)
Fev.-Juil. 2019	0	0	0	14	0
2019-20	0	0	0	4	0
2020-21	0	0	2	4	0
2021-22	3	2	2	0	5
2022-23	3	3	0	4	4
2023-24	3	3	2	5	0
2024-25	1	0	3	5	0
2025-26		2	3	0	0
Volume (heqtd) Total	9h	96h	19h	21h	24h

TABLE 2.4 – Principales U.E.s où j'ai réalisé des activités de tutorats pédagogique et académique dans le cadre des cursus de CentraleSupélec.

2.2.4 Description et autres initiatives pédagogiques spécifiques à ces cours

Je souhaiterais ici mettre en lumière, la description synthétique du contenu de mes différents enseignements ainsi que certaines initiatives pédagogiques que j'ai déployées, qu'il s'agisse de ressources, d'activités mises en place, d'éléments d'évaluation ou d'organisation structurelle.

Cours Apprentissage Automatique (responsable, M2, CM, TD/TP, effectif : entre 70 et 90 (CM), 35(TD/TP), volume horaire total : 210h)

- **Contenu** : Je suis responsable de ce cours de niveau M2 consacré aux fondements et aux méthodes avancées de l'apprentissage supervisé et non supervisé. Il vise à fournir aux étudiants une compréhension théorique des principales méthodes du machine learning ainsi que la capacité à les mobiliser pour résoudre des problèmes réels issus de la science des données. Le cours introduit les notions essentielles (types d'apprentissage, classification, régression, réduction de dimension) ainsi que des méthodes incontournables telles que les méthodes d'ensembles (bagging, random forests, boosting) et les approches non supervisées. Une attention particulière est portée à l'évaluation des modèles et à leur analyse et comparaison. L'enseignement combine cours théoriques, TD/TP en Python et projets sur données réelles afin de développer les compétences pratiques des étudiants. À l'issue du cours, ils doivent être capables d'identifier la méthode adaptée à un problème donné, de l'implémenter et d'en analyser les performances. Ce cours s'inscrit dans la formation de futurs data scientists en les préparant à extraire des connaissances à partir de données complexes et à déployer des solutions d'apprentissage automatique dans des contextes applicatifs variés.
- **Initiatives pédagogiques** : j'assure l'ensemble des cours magistraux ainsi que, selon les années, tout ou partie des TD/TP en Python. J'ai structuré l'évaluation autour d'un examen individuel intermédiaire et d'un projet final en groupe (4 à 5 étudiants) basé sur des données réelles issues de data challenges. Afin de favoriser leur engagement, je leur propose de soumettre des sujets, que je présélectionne avant un vote final des étudiants. Ils disposent ensuite d'un mois et demi pour réaliser leur projet, qui donne lieu à une soutenance finale avec présentation des résultats et du code produit sur Git. Afin de les préparer aux pratiques professionnelles, j'ai introduit l'usage de Git et proposé, les premières années, une séance de formation ainsi qu'une fiche pratique que j'ai conçue. Les livrables (rapport, présentation et code) sont évalués via une grille que j'ai élaborée afin d'harmoniser les évaluations entre les chargés de TDs co-impliqués dans l'évaluation et de fournir des retours clairs aux étudiants. Ces initiatives pédagogiques ont été progressivement améliorées en tenant compte des retours des étudiants lors des premières années d'enseignement. Enfin, j'ai adapté les modalités du cours aux contraintes sanitaires (distanciel en 2020, hybride en 2021, présentiel depuis 2022) tout en maintenant une pédagogie par projet centrée sur la mise en pratique.

Cours Ensemble Learning (responsable, M2, CM, TD, TP, effectif : 100(CM) puis 5(CM) en 2025, 35(TD/TP) puis 5(TD/TP) en 2025, volume horaire total : 288h)

- **Contenu** : Ce cours dispensé en anglais et destiné aux étudiants internationaux du MSc AI. Ce cours porte sur les méthodes d'agrégation de modèles (ensemble learning), qui consistent à combiner plusieurs modèles simples ou qualifiés de faibles afin d'améliorer les performances en classification ou en régression. Le cours couvre les principales approches basées sur les arbres de décision (CART), le bagging, les Random Forests et le boosting, en abordant à la fois leurs fondements théoriques, leurs propriétés pratiques ainsi que leurs limites. Une attention particulière est également portée aux enjeux actuels liés à ces méthodes et à leur utilisation sur des données réelles. L'enseignement combine cours magistraux, travaux pratiques en Python et projets appliqués afin de permettre aux étudiants de mettre en œuvre ces méthodes dans des situations concrètes. Les prérequis incluent des bases en probabilités, en algèbre linéaire, en programmation Python ainsi qu'une première exposition au machine learning. À l'issue du cours, les étudiants doivent être capables d'identifier les problèmes relevant de l'ensemble learning, de choisir et d'implémenter les algorithmes appropriés, d'évaluer leurs performances sur des données réelles et d'analyser de manière critique leurs avantages et leurs limites.
- **Initiatives pédagogiques** : Je suis responsable de cette unité d'enseignement dispensée en anglais, initialement ouverte aux étudiants du MSc IA de CentraleSupélec ainsi qu'à ceux du MSc DSBA de

l'ESSEC (cette ouverture ayant pris fin il y a deux ans). J'assure l'ensemble des cours magistraux ainsi que les TD/TP Python que j'ai conçus afin de garantir une cohérence entre les aspects théoriques et pratiques. L'évaluation repose sur un ou deux examens individuels intermédiaires ainsi qu'un projet final. Initialement réalisé en équipe (4 à 5 étudiants) sur des données issues de data challenges, ce projet est désormais réalisé individuellement depuis la fin de l'ouverture au MSc DSBA, afin de mieux s'adapter aux effectifs et de permettre une évaluation plus individualisée des compétences. La première année, j'avais proposé un projet en partenariat avec l'industriel Rakuten, portant sur l'analyse de données multimodales (texte et images). Par la suite, j'ai introduit un système de choix des projets par vote afin de renforcer la motivation des étudiants et d'adapter le niveau de difficulté. Les projets donnent lieu à une soutenance finale avec présentation des résultats et du code développé sous Git. Les livrables (présentation, rapport et code) sont évalués à l'aide d'une grille que j'ai élaborée pour assurer la transparence et l'harmonisation des évaluations. Enfin, j'ai adapté les modalités pédagogiques aux contraintes sanitaires (distanciel en 2020, hybride en 2021, présentiel depuis 2022).

Cours Lab research projects (responsable, M2, CM+TD/TP, effectif : 15 à 24, volume horaire total : 105h)

- **Contenu :** Je suis responsable de cette unité d'enseignement de formation à la recherche dans laquelle les étudiants internationaux du MSc IA réalisent, sur plusieurs mois, un projet de recherche encadré par un enseignant-chercheur. L'objectif est de les initier aux méthodes de la recherche en intelligence artificielle à travers un travail approfondi sur une problématique scientifique actuelle. Cette UE vise à développer des compétences essentielles telles que la lecture critique d'articles, la formulation d'un problème de recherche, la mise en œuvre d'expérimentations, l'analyse de résultats ainsi que la communication scientifique écrite et orale. L'organisation repose sur un travail en petits groupes, des interactions régulières avec les encadrants, une soutenance intermédiaire et une soutenance finale. Les livrables comprennent un rapport de type article scientifique, un poster ou une présentation, ainsi que le code associé. À l'issue du cours, les étudiants doivent être capables de conduire un projet de recherche complet, de positionner leur contribution par rapport à l'état de l'art et de produire des résultats reproductibles.
- **Initiatives pédagogiques :** Je suis responsable de cette unité d'enseignement (Lab Research Project) dispensée en anglais au MSc IA de CentraleSupélec. Mon objectif est de placer les étudiants dans des conditions proches d'un véritable travail de recherche académique. J'organise une première séance durant laquelle des chercheurs viennent présenter des sujets de recherche que j'ai préalablement collectés, auxquels j'apporte une aide à la calibration des différents jalons pédagogiques. Il s'agit en effet très souvent d'une première initiation à la recherche. Les étudiants classent ensuite les sujets par ordre de préférence et j'organise leur affectation à des groupes de 2 à 3 étudiants. Ils travaillent pendant environ six mois sur leur projet, avec des créneaux hebdomadaires dédiés dans leur emploi du temps. J'organise une soutenance intermédiaire permettant de faire un point d'avancement et, si nécessaire, de réorienter les travaux. Le projet se conclut par une soutenance finale devant un jury composé des encadrants et de moi-même, avec présentation des résultats et du code développé sous Git. Les livrables comprennent un rapport, une présentation et le code, évalués conjointement avec les superviseurs. Cette organisation permet à certains groupes de vivre une expérience de recherche complète, allant parfois jusqu'à la soumission d'un article scientifique ou à l'émergence d'un projet doctoral. Depuis un an, j'ai également introduit une session poster ouverte à un public élargi (enseignants-chercheurs, partenaires industriels et étudiants intéressés par l'IA) organisée au Hub IA de l'École. Cette initiative vise à renforcer la valorisation des travaux, favoriser les interactions avec l'écosystème académique et industriel et contribuer à l'émergence d'opportunités professionnelles pour les étudiants.

Cours Avancées et réalité de l'IA (M2, CM+TP, effectif : 35, volume horaire total : 63h)

- **Contenu :** Cette unité d'enseignement, orientée projets, vise à exposer les étudiants de dernière année en spécialité IA à des thématiques récentes de l'intelligence artificielle ainsi qu'à leurs limites méthodologiques et applicatives. L'objectif est de les amener à développer une compréhension critique

des avancées actuelles de l'IA tout en les confrontant à des problématiques concrètes. Le cours s'articule autour de la réalisation d'un projet en petits groupes, encadré par des enseignants-chercheurs, combinant travail bibliographique, analyse d'articles récents et mise en œuvre d'une approche méthodologique. Des apports thématiques viennent structurer ce travail, notamment sur l'apprentissage avec peu de données annotées (Learning with limited annotated data), l'IA hybride, la robustesse des modèles (Robust ML) et les enjeux de sécurité (AI and safety). L'organisation alterne séances d'autonomie, accompagnement méthodologique et présentations intermédiaires. Le cours se conclut par une restitution orale des travaux permettant d'évaluer la compréhension scientifique, la capacité d'analyse critique et la qualité de la mise en œuvre. À l'issue du cours, les étudiants doivent être capables d'analyser une problématique récente en IA, de mobiliser des méthodes adaptées, et de présenter de manière structurée leurs résultats et leurs limites.

- **Initiatives pédagogiques :** Cet enseignement, réalisé en spécialité IA, combine des exposés courts sur des méthodes récentes de l'intelligence artificielle et un encadrement de projets réalisés en petits groupes de 3 à 4 étudiants. L'objectif est de les confronter rapidement à des thématiques de recherche actuelles et à leurs applications concrètes. En début de module, la responsable du cours, Céline Hudelot, et moi-même proposons une liste de sujets portant sur des problématiques d'IA de pointe, issus soit de la littérature scientifique récente soit de nos propres propositions. Les étudiants sélectionnent ensuite leur sujet et nous les accompagnons dans la lecture et l'analyse critique d'articles scientifiques récents. Nous les guidons également dans le choix d'une méthode pertinente et dans sa mise en œuvre à travers une étude de cas ou un mini-projet applicatif. Le cours se conclut par une séance de restitution orale et de discussion scientifique collective, favorisant la prise de recul et l'esprit critique. Les présentations font l'objet d'une évaluation à laquelle je participe. Enfin, nous avons adapté les modalités pédagogiques aux contraintes sanitaires (distanciel en 2020, hybride en 2021, présentiel depuis 2022) tout en maintenant une pédagogie active par projet.

Cours ST Data@Web : regroupant le cours Recherche d'Information et traitement de données massives, le module introductif et un Enseignement d'Intégration (L3, CM+TD/TP, effectif : 120 (CM), 40 (TD/TP), volume horaire total : 229,5h)

- **Contenu :** Ce cours présente les fondements scientifiques de la recherche d'information (RI), en mettant l'accent sur les méthodes d'indexation, les modèles de recherche, les modèles probabilistes, l'évaluation des systèmes ainsi que les applications modernes comme les moteurs de recherche, la recommandation et la personnalisation. Il vise à donner aux étudiants une compréhension globale du cycle de vie des données du Web, depuis leur collecte jusqu'à leur exploitation. Dans le cadre de la séquence thématique Data@Web, le cours aborde également les spécificités des données web (volume, vitesse, hétérogénéité) et les défis associés à leur traitement. Les objectifs incluent aussi l'acquisition de compétences pratiques en analyse de données, machine learning et développement d'applications en Python. Les enseignements d'intégration sous forme de mini-datakathon permettent aux étudiants de mettre en œuvre ces concepts sur des cas réels (e-réputation, web search, recommandation). Enfin, le cours vise à développer une réflexion critique sur les enjeux sociétaux du numérique (éthique, protection des données, impact social du Web) et à former les étudiants à concevoir des solutions responsables et innovantes.
- **Initiatives pédagogiques :** Je contribue activement à l'évolution pédagogique de cette UE à travers la mise en place de dispositifs favorisant l'apprentissage par projet et l'immersion dans des problématiques réelles. Depuis 2019, j'ai participé, avec les responsables pédagogiques, à la mise en place de la séquence thématique Data@Web : valorisation des données du web, combinant cours théoriques, ateliers d'innovation et un enseignement d'intégration sous forme de Web Datakathon. J'ai notamment piloté un atelier d'innovation (2020-2023) visant à sensibiliser les étudiants aux enjeux éthiques et sociétaux du Web. J'ai également organisé et encadré un Web Datakathon (2020-2022) en partenariat avec des acteurs industriels (Procter & Gamble puis Rakuten), permettant aux étudiants de travailler sur des cas réels de e-marketing. Par ailleurs, j'ai conçu deux chapitres de cours avec une rédaction au format polycopié, plusieurs TD/TP sous Jupyter Notebook ainsi que les évaluations associées. Mon enseignement privilégie une approche équilibrée entre théorie et pratique, renforcée par un accompagnement régulier des étudiants. Enfin, l'adaptation aux formats présentiel, distanciel et hybride m'a conduit à intégrer des outils numériques interactifs pour maintenir l'engagement et

l'autonomie des étudiants (dont la construction et le montage vidéo de cours).

Cours Statistique et Apprentissage (L3, TD, effectif : 40, volume horaire total : 18h)

- **Contenu** : J'ai effectué 18h de TD de mathématiques sous la responsabilité d'Arthur Tenenhaus. Dans un premier temps, ce cours vise à faire acquérir aux futurs ingénieurs les formalismes, concepts et résultats élémentaires de la statistique mathématique. Cela inclut en particulier la définition de modèles statistiques, les principes de la théorie de l'estimation et de la théorie des tests d'hypothèses. Dans un second temps, les élèves se familiarisent avec les méthodes et algorithmes d'apprentissage statistique à partir des données, dans le cadre de l'apprentissage supervisé pour la statistique prédictive ou non supervisé pour la statistique descriptive. Dans ce cadre, ils sont en particulier sensibilisés à la problématique de la grande dimension. J'ai aussi participé à la correction des copies.

Cours Coding-Weeks (L3, TP, effectif : 40, volume horaire total : 96h)

- **Contenu** : J'ai participé pendant 3 ans à cette U.E. qui dure deux semaines à temps pleins sous le format « mini bootcamp » de programmation. L'objectif, au travers de la réalisation de projets de développement informatique, est de permettre aux étudiants de consolider leurs connaissances en programmation et développement d'applications informatiques d'une part, et d'autre part de s'initier aux pratiques et aux méthodologies du développement informatique en groupe de 4 à 5 élèves et en mode projet. J'avais la charge du thème : « Programmation de jeux ». Mon rôle était de guider les élèves dans l'utilisation de Git, dans leur implémentations Python, dans la préparation orale de leur soutenance et de les évaluer. Le bootcamp était en présentiel en 2019 et 2021.

De plus, dans la cadre du cursus CentraleSupélec l'encadrement et le suivis des élèves peut se faire de plusieurs façons et à différents niveaux. Dans le cursus, j'ai suivi des élèves dans le contexte d'autres activités d'enseignements :

- **Digital Tech Year** : encadrement de projets élèves en année de césure à l'École CentraleSupélec (entre le niveau M1 et le niveau M2) pour développer "soft & hard skills".
- **Semestre académique international/professionnel/en entreprise (SIA/SIP/SIE)** : le cursus CentraleSupélec prévoit durant la 2A (M1) le départ pour un Semestre à l'International (SI), soit dans un contexte académique (SIA) ou professionnel (SIP). Un élève en SI est supervisé par un enseignant chercheur. Dans ce cadre, j'ai suivi 7 élèves pour lesquels je réalise aussi une soutenance orale finale.
- **Stage de fin d'étude** : suivi du bon déroulement du stage de fin d'étude d'élèves de l'École CentraleSupélec en spécialité IA ou du Msc IA et également membre du jury de soutenance finale.
- **Encadrement de projets (groupe de 3-5 élèves)** : construction de sujets et encadrement de projets d'élèves en première année d'École d'Ingénieurs (L3) dans le cadre de l'U.E. "projets IA" (20h) ou bien d'élèves de deuxième année (M1) via les U.E.s « projets de modélisation et santé » (50h) et « projets Datascience » (20h).
- **Encadrement de projets avec entreprise** : encadrement de projets d'élèves en dernière année d'École d'Ingénieurs (M2) dans le cadre de tutorat de stage de fin d'études (15h, 8 élèves), ou bien de projets à réaliser pendant l'année sur des sujets entreprise (27 h, 5 groupes d'élèves). Je participe aussi à l'évaluation finale par une soutenance.
- **Tutorat d'élèves apprentis** : suivi d'élèves sur un projet en alternance en entreprise (5 élèves). Je réalise et évalue un minimum de quatre soutenances annuelles par élève.
- **Soutenances de projets** : je participe aux évaluations de projets d'élèves en semestre 8, césure et autre.

2.2.5 Implication sur la formation continue CentraleSupélec Exed en tant que vacataire

En plus de mes activités au sein des cursus de CS, je participe à des U.E. dans le cadre de la formation continue. En 2022, j'ai supervisé un projet de recherche sur la durée de la formation (6 mois) pour SAFRAN (coaching d'un élève, 21h, en distanciel). Le rôle de l'encadrant est de structurer les recherches, de guider et de conseiller sur le plan scientifique pour la réalisation du projet, et de participer aux soutenances finales.

Depuis 2022, je participe à la formation en IA où j'enseigne la deuxième partie des cours magistraux en Deep Learning et Machine Learning (145,5h) en complément des interventions de Frédéric Pascal. Ces enseignements ont lieu en présentiel, auprès de groupes de 5 à 10 personnes. Les publics sont des professionnels provenant par exemple de la DGA, MBDA ou Thalès. L'approche d'enseignement est orientée applications industrielles et transfert de compétences.

2.3 Université Grenoble Alpes (2017-2019)

J'ai effectué à l'Université Grenoble Alpes des enseignements pour des étudiants en Master d'Informatique et Licence 3 d'Informatique. Je suis intervenue dans deux Unités d'Enseignements (U.E.) intitulées "Algorithmes pour le traitement de données" et "Informatique et Aléatoire".

Algorithmes pour le traitement de données (13h)

J'ai effectué des TP et TD pour un groupe d'une vingtaine d'étudiants de M1 en Informatique. Les objectifs de l'U.E. sont d'introduire aux étudiants les principaux modèles et algorithmes utilisés pour le traitement de données ainsi que de les leur faire implémenter entièrement ou en partie via des TP en python. Un intérêt particulier est porté sur,

- l'indexation et la représentation des documents
- la recherche d'information (RI), y compris la RI sur le web
- la catégorisation et la classification (clustering)
- la fouille de données

En TP ou TD sont étudiés l'algorithme d'ordonnement du *PageRank*, celui des *k*-moyennes, les réseaux de neurones artificiels de type Perceptron Multi-Couche, les méthodes agglomératives (algorithmes de classification hiérarchique), SVM, *Naïve Bayes* et le maximum de vraisemblance. J'ai réalisé, en collaboration avec Thibaut Thonet et Éric Gaussier (le responsable de l'U.E.), les sujets de TP, TD et corrigé des rendus de TP (*scripts*). Tous les TP ne sont pas évalués mais les scripts sont relevés à la fin de chaque séance. Certains étudiants viennent de parcours où ils n'ont pas été familiarisés au langage python. Les premières séances ne sont donc pas évaluées pour leur permettre de prendre en main l'outil python. Le livre "Python Data Science Handbook" (J. VanderPlas, 2016) leur a été recommandé. Pour compléter le cours, les livres "Recherche d'information : Applications, modèles et algorithmes" (M.-R. Amini, E. Gaussier, 2017) et "Introduction to Information Retrieval" (C. Manning, P. Raghavan, H. Schütze, 2008) leur ont également été recommandés.

Informatique et Aléatoire (15h)

J'ai enseigné le cours et délivré les TD et TP en python pour un groupe d'une quinzaine d'étudiants en L3 Informatique. Cette U.E. est une option choisie par les étudiants dont l'objectif est de leur délivrer des bases en modélisation probabiliste pour l'informatique afin de leur faciliter l'acquisition des notions qui seront abordées en Master Informatique. L'U.E. est divisée en deux parties. La première, que j'ai enseigné, aborde les notions de probabilités élémentaires, de variables aléatoires, d'inégalités (Markov, Chebychev), de lois de probabilité ainsi que la loi faible des grands nombres. La seconde partie est enseignée par un autre enseignant : Nicolas Gast. Il aborde les notions de convergence et de chaînes de Markov. Le responsable de l'U.E. et auteur de la première partie du cours est Massih-Reza Amini. J'ai réalisé les énoncés de TD et encadré à la fois les séances de TD et de TP. Sur les 5 séances de TP, les étudiants avaient 4 séances pour réaliser un projet python consistant en la construction d'une ou plusieurs fonctions permettant de déterminer automatiquement la langue d'un mot. Ces fonctions devaient respecter un modèle de détection de langue donné dans l'énoncé du projet. Pour commencer, un modèle "simpliste" à implémenter leur est proposé puis en bonus, il leur est proposé d'adapter leur(s) fonctions python à un nouveau modèle un peu plus précis : le modèle *bi-gramme*. Un corpus de mots de 5 langues est mis à leur disposition comme données d'apprentissage puis un dernier fichier test de données est disponible pour leur permettre d'évaluer les performances des deux modèles appris ainsi que de les comparer. A la fin de ces quatre séances il leur est demandé de rendre un rapport accompagnant leur script commenté. La cinquième et dernière séance est ensuite dédiée à des entretiens individuels avec moi-même où l'objectif est la présentation orale du script rendu par les étudiants. Ils sont alors amenés à expliquer leur réflexion, les démarches qu'ils ont entreprises pour aboutir à cette structure d'algorithme. Je vais contribuer à la réalisation du sujet d'examen dans les

prochaines semaines. Pendant l'examen, les étudiants auront droit à une fiche "pense-bête" que j'ai distribuée en début de cours et présentée comme un contrat pédagogique entre eux et l'enseignant. Mon but est de leur permettre de visualiser concrètement les notions "piliers" du cours et de les pousser à se manifester quand ces notions restent confuses pendant leur présentation en cours.

Langue d'enseignement

La langue officielle d'enseignement de ces U.E. est le français. Cependant, dans le cadre d'échanges internationaux, de nombreux étudiants sont étrangers. Ainsi, j'ai par exemple un étudiant américain et une étudiante canadienne parmi mes étudiants en Informatique et Aléatoire puis, quelques étudiants espagnols et asiatiques parmi mon groupe d'étudiants en Algorithmes pour le traitement des données. Ainsi, j'ai été amenée à communiquer en anglais avec eux lors d'explications complémentaires.

2.4 Université de Montpellier (UM) (2012-2016)

J'ai effectué à l'UM un service de 376,5 h d'enseignement : environ 64 h par an pendant 3 ans dans le cadre de monitorats puis un service complet de 192 h en tant qu'ATER. Les publics étaient différents et à des niveaux variés : en Licence de Mathématiques, en Licence d'Informatique, en Licence de Biologie et en École d'Ingénieurs (Polytech Montpellier). Je présente ici l'ensemble des U.E. que j'ai dispensé.

Mathématiques et Statistique pour l'écologie (175.5 h)

J'ai effectué des cours, TD et TP de cette U.E. pour des groupes d'une quarantaine d'étudiants de L2 Géoscience Biologie Environnement durant les trois premières années de ma thèse et seulement des TP la dernière année. La responsable d'U.E. était Elodie Brunel. Les objectifs de l'U.E. sont d'introduire des outils mathématiques et statistique ainsi que leur application pratique sur R pour que les étudiants les comprennent et se les approprient. L'U.E. est composée en trois parties : une première d'algèbre, une deuxième de statistique et en fin de semestre une dernière partie de TP, dont j'ai respectivement dispensé 22 h, 75,5 h et 78 h. J'ai réalisé des sujets de Contrôles Continus (CC), corrigé des copies et des rapports de TP et surveillé les examens.

Biostatistique (25.5 h)

J'avais pour charge de dispenser des cours, TD (15,5 h) et TP (10 h) de statistique descriptive univariée dans un premier temps puis bivariée pour un groupe d'une quarantaine d'étudiants de l'école d'ingénieurs Polytech Montpellier et des étudiants de L2 Biologie de l'UM. Le responsable d'U.E. était Lionel Cucala. J'ai également corrigé des copies et surveillé l'ensemble des examens.

Probabilités élémentaires (13.5 h)

J'ai effectué des TD pour un groupe d'une vingtaine d'étudiants de L2 Mathématiques Fondamentales. Le cours était délivré par Lionel Cucala (responsable de l'U.E.). Malgré les changements de programmes au lycée ces dernières années, il est difficile pour eux de formaliser une situation réelle via les outils mathématiques de probabilités. Le but était donc de les aider à travers une multitude d'exercices à se familiariser avec tous ces outils et leur usage ; mais aussi d'insister sur la rigueur de rédaction. J'ai corrigé des copies et surveillé les examens.

Statistiques descriptives et probabilités (87 h) (responsable)

J'ai été responsable de cette U.E. dispensée à une centaine d'étudiants de L2 Informatique. J'ai effectué 36 h de cours magistral et 51 h de TD. J'ai construit un polycopié de cours et des fiches de TD (cf. ma page web), et utilisé R pour illustrer des notions de cours (telles que le Théorème Central Limite (TCL)). J'ai également organisé les sessions d'examens (CC, examens terminaux), corrigé les copies et surveillé les examens. Le but était de leur enseigner des probabilités élémentaires, puis des rappels de statistique et enfin faire le lien avec les notions de probabilités élémentaires permettant d'aborder le TCL et le calcul d'intervalles de confiances.

Probabilités et Statistique (39 h)

J'ai effectué 39 h de TD pour une vingtaine d'étudiants de L3 Mathématiques Fondamentales sous la responsabilité d'U.E. de Benoîte De Saporta (qui enseignait le cours). Après des rappels de probabilités, les premiers objectifs sont de leur présenter les limites sup et inf afin de leur apprendre à manipuler les différentes convergences (presque sure, en probabilité, en loi) plus particulièrement par l'utilisation du Lemme de Borel Cantelli. Les objectifs suivants consistent à leur présenter la loi des grands nombres, le théorème de Lévy, le TCL et enfin leur apprendre à calculer des intervalles de confiances. L'U.E. se termine sur l'enseignement des notions de vecteurs gaussiens, leurs propriétés et lois satellites. J'ai corrigé les copies de CC et surveillé les examens.

Analyse (36 h)

J'ai dispensé 36 h de TD à un groupe d'une quarantaine d'étudiants de deuxième année de l'école d'ingénieurs Polytech Montpellier. Le cours était délivré par Benjamin Charlier (responsable d'U.E.). Le programme est partagé en une partie de probabilités élémentaires et une seconde partie d'analyse. La première partie est composée des notions de dénombrement et de manipulation des lois usuelles. Quant à la partie d'analyse, elle a pour but de faire le lien entre les outils d'algèbre déjà acquis et l'analyse. Le programme de cette dernière partie aborde la géométrie vectorielle, les courbes paramétrées, les normes et espaces euclidiens. J'ai corrigé les copies des CC et des examens terminaux ainsi que surveillé les examens.

Activités d'enseignement supplémentaires :

Assistante d'enseignement (M2 Biostatistique, 15 étudiants)

J'ai assisté mon co-superviseur de thèse, responsable du cours Méthodes d'estimation pour modèles à variables latentes, en contribuant à la préparation du support de cours en méthodes d'estimation pour modèles à variables latentes (à partir de mes travaux de thèse), accompagnement des étudiants en séance, partage de notes de recherche et mise à disposition de scripts R pour travaux pratiques.

Membre de jury de soutenance (niveau Licence)

J'ai contribué à l'évaluation d'un projet en probabilités (« ruine du joueur »), analyse du rapport et de la soutenance, participation à l'attribution de la note finale en coordination avec l'équipe pédagogique.

Cours particuliers (Master & Doctorat) Enfin, je me suis investie dans l'accompagnement en mathématiques et statistique (écologie, psychologie) part un appui à des analyses de données sociales et comportementales (préparation, modélisation et interprétation).

2.5 Bilan et charge globale en HETD

Au bilan, ma contribution à l'enseignement à CentraleSupélec est restée globalement stable, avec environ 225 heTD par an en formation initiale, complétées par 166,5 heures en formation continue pour Centrale-Supélec Exed. **Depuis septembre 2024, je bénéficie d'une décharge d'enseignement d'environ 200 h eqTD (soit un demi-service), accordée au titre de mes responsabilités pédagogiques et du financement, en grande partie, par un projet ANR JCJC dont je suis lauréate. Mes enseignements, dispensés en français et en anglais, couvrent les niveaux L3 à M2 au sein des départements d'informatique et de mathématiques. Ils s'adressent à des publics variés, incluant les élèves du cycle ingénieur, les étudiants du MSc Artificial Intelligence ainsi que des professionnels en formation continue.**

Avant mon recrutement à CentraleSupélec, j'ai assuré des enseignements en informatique et en mathématiques à l'Université Grenoble Alpes et à l'Université de Montpellier. Durant mon postdoctorat à Grenoble, j'ai été vacataire pour des étudiants de Licence 3 et de Master d'informatique, où j'ai dispensé des enseignements en algorithmique et en informatique probabiliste (28heqTD). Pendant ma thèse, j'ai effectué un service d'enseignement conséquent à l'Université de Montpellier, d'abord dans le cadre de trois monitorats puis en tant qu'ATER, auprès de publics variés (licences de mathématiques, d'informatique et de biologie, ainsi qu'en école d'ingénieurs).

3 Activités de recherche

3.1 Orientation et axes de recherche

3.1.1 Résumé général et contexte

Lors de mon recrutement en tant qu'enseignante-chercheure contractuelle au laboratoire MICS (dirigé par PH. Cournède puis C. Hudelot), j'ai intégré l'équipe LOGIMICS (LOGIcal approaches for Modelling, data Interpretation and Complex Systems) de P. Le Gall. Depuis, le découpage en équipes du laboratoire MICS a été actualisé. Je suis à ce jour membre de l'équipe GEMILA de G. Fay où mes activités de recherche portent sur le domaine de l'Intelligence Artificielle (IA) et plus précisément l'apprentissage statistique. Cette section détaille mes orientations de recherche, d'abord de manière résumée puis plus développée pour chacun de mes 3 axes.

Résumé : La trajectoire scientifique de mes travaux de recherche est principalement consacrée à la conception de modèles d'apprentissage statistique alliant prédiction, compréhension et fiabilité, dans des contextes de données réelles, hétérogènes, imparfaites ou temporelles. L'ensemble des travaux s'inscrit dans une évolution méthodologique et conceptuelle allant des approches boîtes blanches, interprétables et guidées par la connaissance, vers les boîtes noires issues de l'apprentissage profond, pour aboutir aujourd'hui à une vision boîte grise, conciliant puissance prédictive et transparence causale. Mes premières contributions s'inscrivaient dans le paradigme de l'apprentissage supervisé et en modélisation statistique interprétable, centrées sur l'intégration de la connaissance experte dans la modélisation de phénomènes complexes, parfois bruités ou partiellement observés. Le deuxième axe de mes contributions est quant à lui consacré au développement de méthodes d'apprentissage avec peu ou pas de labels, d'adaptation de domaine et de Few-Shot Learning, principalement appliquées à des données d'images, dans l'objectif d'améliorer la robustesse et la transférabilité des modèles d'apprentissage. Enfin, le troisième axe introduit et formalise un cadre d'intelligence artificielle "boîte grise", fondé sur la fusion entre apprentissage profond et causalité. Les contributions présentées portent notamment sur l'apprentissage de représentations causales, l'estimation d'effets de traitement individuels, ainsi que sur la robustesse causale contrefactuelle face aux décalages de distribution. Ces travaux, issus notamment du projet ANR JCJC CLEAR Deep (A Causal Explanation-oriented DEEP Structured framework adaptable to temporal data for trustability and robustness), ouvrent la voie à des modèles hybrides combinant efficacité prédictive et interprétabilité, où la causalité joue un rôle structurant dans la représentation et l'inférence. Ce cadre ambitionne de contribuer à une IA de confiance, dont l'interprétabilité est intrinsèque, au service d'une prise de décision éclairée et responsable.

3.1.2 Positionnement de mon travail : vers une IA robuste, causale et digne de confiance

Le développement récent de l'intelligence artificielle, et en particulier du deep learning, a profondément transformé de nombreux domaines scientifiques et industriels. Ces avancées reposent en grande partie sur le paradigme de l'apprentissage supervisé, qui a permis des progrès majeurs en vision par ordinateur, en traitement automatique du langage et plus généralement en science des données. Toutefois, malgré leurs performances remarquables, ces approches présentent plusieurs limitations structurelles qui restreignent leur applicabilité dans de nombreux contextes réels.

Parmi ces limitations, plusieurs défis scientifiques majeurs peuvent être identifiés :

- la forte dépendance des méthodes d'apprentissage profond à de grandes quantités de données annotées, dont la production est souvent coûteuse ou irréaliste dans de nombreux domaines applicatifs ;
- le manque de robustesse des modèles face aux changements de distribution (distribution shifts) et leur difficulté à généraliser hors des conditions d'apprentissage ;
- le caractère largement opaque des modèles profonds (black-box models), qui limite leur interprétabilité et leur acceptabilité dans des contextes décisionnels sensibles ;
- la difficulté à distinguer corrélations statistiques et relations causales, ce qui limite la capacité des modèles à produire des prédictions fiables ;

- plus généralement, le besoin croissant de méthodes d’IA à la fois performantes, robustes et dignes de confiance.

Dans ce contexte, mes travaux de recherche s’articulent autour de deux axes scientifiques complémentaires. Le premier axe porte sur l’apprentissage statistique en contexte de supervision limitée et vise à concevoir des méthodes d’IA plus robustes et moins dépendantes de grandes quantités de données annotées. Le second axe se situe à l’interface entre apprentissage statistique et IA causale et a pour objectif le développement de méthodes d’apprentissage plus robustes, plus interprétables et plus fiables pour des applications décisionnelles.

Le premier axe concerne donc l’apprentissage statistique en contexte de supervision limitée. Contrairement à l’approche standard consistant à améliorer les performances par l’augmentation des volumes de données annotées ou de la taille des modèles, mes travaux explorent la conception de nouveaux biais inductifs (inductive biases), que ce soit à travers les objectifs d’apprentissage, les architectures ou les protocoles d’entraînement. Ces contributions s’inscrivent notamment dans les cadres de l’apprentissage semi-supervisé, non supervisé, de l’adaptation de domaine et du few-shot learning, avec une attention particulière portée à la robustesse aux changements de distribution et aux contraintes réalistes des applications industrielles.

Le second axe de mes recherches vise quant à lui à intégrer les apports de la causalité à l’apprentissage automatique, afin d’améliorer la robustesse et la transparence des modèles. Dans cette perspective, je m’intéresse en particulier à l’inférence causale, à l’apprentissage de représentations causales et à la découverte de structures causales à partir de données observationnelles. L’idée directrice est que l’introduction de modèles causaux explicites ou de représentations latentes causales permet de dépasser les limites des approches purement statistiques, notamment en améliorant la robustesse aux perturbations, l’explicabilité des prédictions et la fiabilité des systèmes d’aide à la décision.

Ces deux axes présentent une forte cohérence scientifique. Les travaux sur l’apprentissage sous contrainte de données limitées ont naturellement conduit à s’intéresser aux questions de robustesse et de généralisation, qui constituent également des problématiques centrales en causalité. Réciproquement, l’intégration de modèles causaux dans les systèmes d’apprentissage statistique offre de nouvelles perspectives pour concevoir des méthodes plus fiables et mieux adaptées aux contraintes des applications réelles.

Ainsi, l’ensemble de mes travaux contribue à un objectif scientifique commun : le développement de méthodes d’intelligence artificielle capables d’apprendre de manière plus efficace à partir de données imparfaites, tout en offrant de meilleures garanties de robustesse, de généralisation et d’interprétabilité. Ce positionnement s’inscrit pleinement dans les enjeux actuels de l’IA de confiance et de l’IA causale, qui constituent aujourd’hui des directions structurantes de la recherche internationale en intelligence artificielle.

3.1.3 Axe 1 : Apprentissage statistique sous contrainte de supervision limitée

Un premier axe structurant de mes recherches concerne le développement de méthodes d’apprentissage statistique capables de fonctionner dans des contextes où les données annotées sont limitées, bruitées ou coûteuses à obtenir. Cette problématique est devenue centrale avec l’augmentation massive des volumes de données disponibles, pour lesquelles l’annotation humaine constitue désormais le principal facteur limitant.

Mes travaux dans cet axe visent à concevoir des méthodes permettant de réduire la dépendance aux annotations tout en améliorant la robustesse des modèles. Ils s’inscrivent principalement dans les cadres de l’adaptation de domaine, de l’apprentissage semi-supervisé, de l’apprentissage non supervisé et du few-shot learning, avec une attention particulière portée aux contraintes réalistes rencontrées dans les applications industrielles.

Apprentissage adaptatif et robustesse aux changements de distribution. Dans le cadre du co-encadrement de la thèse de Victor Bouvier, je me suis intéressée aux fondements théoriques de l’adaptation de domaine. Nos travaux ont notamment permis d’introduire un nouveau terme d’erreur, appelé *Hypothesis Class Reduction error*, qui formalise le risque de perte d’information lors de la construction de représentations invariantes entre domaines. Cette contribution permet de mieux comprendre les compromis entre invariance et transférabilité, et souligne le rôle central des biais inductifs dans la capacité d’adaptation des modèles.

Ces travaux ont également conduit au développement de nouvelles méthodes d’adaptation non supervisée robustes aux changements de distribution des labels, ainsi qu’à des stratégies d’adaptation active permettant d’identifier les exemples cibles les plus informatifs à annoter. Ils ont aussi permis de mettre en évidence certaines limites des approches cherchant uniquement l’invariance, notamment en termes de robustesse.

Plus récemment, dans le cadre du co-encadrement de la thèse CIFRE d'Étienne Bennequin, je me suis intéressée aux situations où les données disponibles sont limitées dans les deux domaines source et cible. Nous avons notamment introduit le cadre du Few-Shot Learning under Support/Query Shift, qui formalise des situations réalistes où les distributions des données d'apprentissage et de test diffèrent, tout en proposant des protocoles expérimentaux de référence pour ce problème.

Apprentissage parcimonieux en labels pour la vision par ordinateur. Dans le cadre du co-encadrement de la thèse de Yassine El Ouali, j'ai contribué au développement de méthodes d'apprentissage efficaces en données annotées pour des tâches de segmentation d'images. Ces travaux ont porté sur différents paradigmes d'apprentissage, allant du semi-supervisé au non supervisé, et ont permis de proposer de nouveaux cadres méthodologiques exploitant la structure intrinsèque des données pour améliorer l'efficacité de l'apprentissage.

Ces contributions s'inscrivent dans une démarche plus large visant à concevoir des méthodes capables d'opérer à différents niveaux d'abstraction (image entière ou pixel) et sur différentes modalités de données, tout en maintenant des propriétés de robustesse et de généralisation.

Apprentissage méta et adaptation rapide à de nouvelles tâches. Mes travaux dans le domaine du few-shot learning ont également porté sur le développement de méthodes capables de s'adapter rapidement à de nouvelles classes à partir d'un nombre très limité d'exemples. Ces recherches, menées notamment dans un contexte de supervision d'une thèse CIFRE (où le doctorant était E. Bennequin) et en collaboration internationale, ont permis de proposer de nouveaux protocoles d'évaluation plus réalistes ainsi que des méthodes mieux adaptées aux contraintes industrielles.

L'ensemble de ces travaux contribue à un objectif commun : concevoir des systèmes d'apprentissage plus flexibles, capables d'apprendre efficacement à partir de données imparfaites ou limitées, tout en améliorant leur robustesse.

3.1.4 Axe 2 : IA causale pour un apprentissage robuste, interprétable et digne de confiance

Un second axe majeur de mes recherches concerne l'intégration des approches causales au machine learning, dans le but d'améliorer la robustesse, l'interprétabilité et la fiabilité des systèmes d'intelligence artificielle. Cet axe s'inscrit dans le cadre plus large de l'IA de confiance, qui constitue aujourd'hui un enjeu majeur pour le déploiement de l'IA dans des contextes décisionnels sensibles.

Mes contributions dans ce domaine couvrent plusieurs aspects complémentaires, allant de l'inférence causale à l' et à la découverte de structures causales.

Apprentissage statistique avec variables latentes et données incertaines. Mes premiers travaux sur ces questions remontent à ma thèse, au cours de laquelle j'ai proposé une méthode d'estimation par EM pour des modèles d'équations structurelles avec variables latentes. Ces travaux ont notamment trouvé des applications en oncologie, où ils ont permis de simplifier l'analyse longitudinale de données de qualité de vie en réduisant la dimension des variables observées tout en améliorant leur interprétabilité.

Dans la continuité de ces travaux, j'ai également contribué au développement de modèles d'apprentissage interprétables et robustes aux incertitudes, notamment à travers la généralisation probabiliste des arbres de régression. Ces travaux ont permis de proposer des modèles conciliant performance prédictive, interprétabilité et robustesse au bruit, tout en conservant des garanties théoriques de consistance.

Inférence causale et estimation d'effets de traitement individualisés. Dans le cadre du co-encadrement de la thèse CIFRE de Mouad El Bouchattaoui, j'ai contribué au développement de nouvelles méthodes d'inférence causale pour des données longitudinales et de grande dimension. Ces travaux ont notamment conduit à la proposition de modèles génératifs causaux dynamiques permettant d'estimer les effets de traitement individualisés en tenant compte de variables d'ajustement latentes, c'est à dire non observées.

Apprentissage de représentations causales robustes. Plus récemment, dans le cadre du projet ANR JCJC ClearDeep, mes travaux se sont concentrés sur l'étude de la robustesse des modèles causaux génératifs

face aux changements de distribution. Nous avons notamment introduit le concept de robustesse contre-factuelle, qui permet d'évaluer la sensibilité des modèles causaux aux interventions et aux variations de distribution.

Ces travaux montrent notamment que tous les modèles causaux génératifs contrefactuels (c'est à dire sous une intervention) ne présentent pas les mêmes propriétés de robustesse, et que certaines interventions sont intrinsèquement plus robustes que d'autres. Ils ouvrent ainsi de nouvelles perspectives pour la conception de modèles causaux plus fiables dans des environnements dynamiques.

Découverte causale à partir de données observationnelles. Enfin, mes travaux portent également sur la découverte automatique de structures causales. Dans ce cadre, j'ai contribué au développement d'un nouveau critère invariant d'échelle permettant d'identifier l'ordre causal dans les modèles à bruit additif. Cette contribution a conduit au développement d'un algorithme simple et efficace pour la découverte causale, ainsi qu'à la mise à disposition d'outils pour l'évaluation expérimentale de ces méthodes.

3.2 Publications et productions scientifiques

Dans les publications, les doctorants, postdoctorants et stagiaires que j'ai co-encadrés sont soulignés, et mon nom apparaît en **gras**. Les travaux sont également disponibles sur ma page web. Sauf mention contraire, l'ordre des auteurs suit l'usage : doctorant en premier auteur, directeur/directrice en dernier. En particulier, en informatique, **certaines conférences internationales sont aussi sélectives que des revues à fort Impact Factor (IF)**. Leur qualité est notamment évaluée via le classement **CORE**¹, largement utilisé dans la discipline : A* (top 7,47 %), A, B, C, les 3,11 % restantes étant non notées.

Revues internationales avec comité de lecture : 2

[J1] M. El Bouchattaoui, **M. Tami**, B. Lepetit, P. H. Cournède : Causal Dynamic Variational Autoencoder for Counterfactual Regression. Transactions on Machine Learning Research, 2025. (29p).

[J2] A. Barbieri, **M. Tami**², X. Bry, D. Azria, S. Gourgou, C. Bascoul-Mollevi, C. Lavergne : EM algorithm estimation of a structural equation model for the longitudinal study of the quality of life. Statistics in Medicine, 2017, vol. 37, no 6, p. 1031-1046. (Rang : Q1)

Articles de conférences internationales avec comité de lecture et actes publiés : 16

[C1] M. Benhamza, M. Clausel, **M. Tami** : Counterfactual Robustness : a framework to analyze the robustness of Causal Generative Models across interventions. Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham : Springer Nature Switzerland, 2025. p. 391-408. (CORE : A)

[C2] M. El Bouchattaoui, **M. Tami**, B. Lepetit, P. H. Cournède : Causal contrastive learning for counterfactual regression over time. Advances in Neural Information Processing Systems, 2024, vol. 37, p. 1333-1369. (CORE : A*)

[C3] M. L. Bouchouia, H. Khemissa, E. Gherbi, **M. Tami**³, D. Lopes, N. Alkhatib, M. Ayrault : Cybersecurity Metrics for AI-Based In-Vehicle Intrusion Detection Systems. 2024 IEEE Vehicular Networking Conference (VNC). IEEE, 2024. p. 269-270.

[C4] A. G. Reisach, **M. Tami**, C. Seiler, A. Chambaz, S. Weichwald : A Scale-Invariant Sorting Criterion to Find a Causal Order in Additive Noise Models. Advances in Neural Information Processing Systems, 2023, vol. 36, p. 785-807. (CORE : A*)

1. <http://www.core.edu.au/>
2. A. Barbieri et moi sommes co-premier auteurs.
3. J'ai assuré la co-supervision scientifique du postdoctorant, en encadrant les orientations de recherche lors de réunions mensuelles et en réalisant la relecture finale des travaux.

[C5] P. Colombo, V. Pellegrin, M. Boudiaf⁴, **M. Tami**⁵, V. Storchan, C. Hudelot, I. Ben Ayed, P. Piantanida : Transductive Learning for Textual Few-Shot Classification in API-based Embedding Models. Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing. 2023. p. 4214-4231. (CORE : A*)

[C6] M. Boudiaf⁶, E. Bennequin, **M. Tami**⁷, A. Toubhans, P. Piantanida, C. Hudelot, I. Ben Ayed : Open-Set Likelihood Maximization for Few-Shot Learning. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 24007-24016. (CORE : A*)

[C7] V. Bouvier, P. Very, C. Chastagnol, **M. Tami**⁸, C. Hudelot : Robust Domain Adaptation : Representations, Weights and Inductive Bias (Extended Abstract). Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21) Sister Conferences Best Papers Track, pp. 4750-4754. (CORE : A*)

[C8] V. Bouvier, P. Very, C. Chastagnol, **M. Tami**⁹, C. Hudelot : Stochastic Adversarial Gradient Embedding for Active Domain Adaptation. Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham : Springer International Publishing, 2020. p. 353-377. (CORE : A)

[C9] E. Bennequin, V. Bouvier, **M. Tami**, A. Toubhans, C. Hudelot : Bridging Few-Shot Learning and Adaptation : New Challenges of Support-Query Shift. Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham : Springer International Publishing, 2021. p. 554-569. (CORE : A)

[C10] Y. Ouali, C. Hudelot, **M. Tami**¹⁰ : Spatial contrastive learning for few-shot classification. Joint European conference on machine learning and knowledge discovery in databases. Cham : Springer International Publishing, 2021. p. 671-686. (CORE : A)

[C11] S. Alkhoury, E. Devijver, M. Clausel, **M. Tami**¹¹, E. Gaussier, G. Oppenheim : Smooth And Consistent Probabilistic Regression Trees. Advances in Neural Information Processing Systems, 2020, vol. 33, p. 11345- 11355. (CORE : A*)

[C12] V. Bouvier, P. Very, C. Chastagnol, **M. Tami**¹², C. Hudelot : Robust Domain Adaptation : Representations, Weights and Inductive Bias. Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham : Springer International Publishing, 2020. p. 353-377. (CORE : A)

[C13] Y. Ouali, C. Hudelot, **M. Tami**¹³ : Autoregressive Unsupervised Image Segmentation. European conference on computer vision. Cham : Springer International Publishing, 2020. p. 142-158. (CORE : A*)

4. M. Boudiaf n'est pas souligné car sa supervision n'était pas assurée par moi-même mais localement à l'ETS, Montréal Canada.

5. Ma contribution a consisté à participer à une dizaine de réunions de travail, à accompagner la réflexion expérimentale notamment pour le débogage du code, et à assurer la relecture finale des travaux.

6. M. Boudiaf n'est pas souligné car sa supervision n'était pas assurée par moi-même mais localement à l'ETS, Montréal, Canada.

7. M. Boudiaf et E. Bennequin sont co-premier auteurs en tant que 2 doctorants dans le cadre de cette collaboration avec l'ETS.

8. J'ai participé aux réunions bi-hebdomadaires à partir de la deuxième année de thèse, cadré les aspects mathématiques, contribué aux idées méthodologiques dont la MMD avec les IPMs et assuré la relecture des développements et de l'article.

9. J'ai participé aux réunions bi-hebdomadaires à partir de la deuxième année de thèse, cadré les aspects mathématiques, contribué aux idées méthodologiques et assuré la relecture des développements et de l'article.

10. J'ai participé aux réunions bimensuelles, cadré les aspects mathématiques, échangé sur les idées méthodologiques et assuré la relecture des développements et de l'article.

11. J'ai transformé la problématique appliquée fournie par l'entreprise en une direction de recherche, construit la démarche scientifique validée par mes superviseurs, conçu le prédicteur et ses résultats mathématiques, et réalisé les premières expérimentations. Après mon recrutement, les expériences ont été poursuivies par un ingénieur S. Alkhoury et la preuve de consistance affinée avec É. Devijver. J'ai rédigé l'ensemble des premières versions de l'article.

12. J'ai participé aux réunions bi-hebdomadaires à partir de la deuxième année de thèse, cadré les aspects mathématiques, contribué aux idées méthodologiques dont la MMD avec les IPMs et assuré la relecture des développements et de l'article.

13. J'ai participé aux réunions bimensuelles, cadré les aspects mathématiques, échangé sur les idées méthodologiques et assuré la relecture des développements et de l'article.

[C14] Y. Ouali, C. Hudelot, **M. Tami**¹⁴ : Semi-Supervised Semantic Segmentation with Cross-Consistency Training. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020. p. 12674-12684. (CORE : A*)

[C15] **M. Tami**, X. Bry, C. Lavergne : EM estimation of a structural equation model. The 22nd International Conference on Computational Statistics (COMPSTAT 2016). (p.36)

[C16] **M. Tami**, X. Bry, C. Lavergne : EM estimation of a structural equation model. The 36th Conference on Applied Statistics in Ireland (CASI 2016) (2p).

Articles de conférences nationales avec comité de lecture et actes publiés : 9

[N1] M. Benhamza, M. Clausel, **M. Tami** : AdaptiveCausalWVAE : Certifiably Robust and Stable Causal Disentanglement. Conférence de la Société Française de Statistique, 56èmes Journées de la Statistique, 2025. p. 749-754.

[N2] M. Benhamza, M. Clausel, **M. Tami** : AdaptiveCausalWVAE : Certifiably Robust and Stable Causal Disentanglement with Lipschitz-Constrained Spline Neural Networks. The 9th edition of the Junior Conference on Data Science and Engineering (JDSE), Best poster, 2024. (6p).

[N3] M. El Bouchattaoui, **M. Tami**, B. Lepetit, P.H. Cournède : A non-linear mixed effect model for the estimation of causal effects through time. Conférence de la Société Française de Statistique, 54èmes Journées de Statistique, 2023. (10p.)

[N4] V. Pellegrain, **M. Tami**, M. Batteux, C. Hudelot : Apprentissage multimodal pour le diagnostic de fautes sur données séquentielles non alignées et arbitrairement longues. Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle (APIA 2022) - Event affiliated with PFIA 2022. (10p.)

[N5] V. Pellegrain, **M. Tami**, M. Batteux, C. Hudelot. : On the difficulties of training VAE-based models for anomaly detection in industrial systems. The 4th edition of the Junior Conference on Data Science and Engineering (JDSE), 2019. (6p).

[N6] **M. Tami**, M. Clausel, E. Devijver, E. Gaussier, J-M. Aubert, M. Chebre : Decision Tree for uncertain measures. Conférence de la Société Française de Statistique, 50èmes Journées de la Statistique, 2018. (6p).

[N7] **M. Tami**, A. Barbieri, X. Bry, C. Bascoul-Mollevi, C. Lavergne : Analyse longitudinale de la qualité de vie relative à la santé en cancérologie par équation structurelle et modèles mixtes. Conférence de la Société Française de Statistique, 48èmes Journées de la Statistique, 2016. (6p).

[N8] **M. Tami**, A. Barbieri, X. Bry, C. Bascoul-Mollevi, C. Lavergne : Analyse longitudinale de la qualité de vie en cancérologie par modèles à équations structurelles et à effets aléatoires. Colloque EPICLIN 9 des 22èmes Journées des statisticiens de Centre de Lutte Contre le Cancer (CLCC), 2015. (2p).

[N9] **M. Tami**, X. Bry, C. Lavergne : Estimation par algorithmes EM de modèles à facteurs et à équations structurelles. Conférence de la Société Française de Statistique, 46èmes Journées de la Statistique, 2014. (6p).

Articles de workshop internationaux avec comité de lecture et actes publiés : 3

[W1] M. Benhamza, M. Clausel, **M. Tami** : Towards Understanding When Causal Structure Improves Robustness : Evidence from Generative Models. AISTATS CauScale 2026 workshop. (5p). (CORE : A)

14. J'ai participé aux réunions bimensuelles, cadré les aspects mathématiques, échangé sur les idées méthodologiques et assuré la relecture des développements et de l'article.

[W2] M. El Bouchattaoui, M. Tami, B. Lepetit, P. H. Cournède : Toward a more transparent causal representation learning. 9th Causal Inference Workshop at UAI 2024. (3p). (CORE : A)

[W3] E. Bennequin, **M. Tami**, A. Toubhans, C. Hudelot : Few-Shot Image Classification Benchmarks are Too Far From Reality : Build Back Better with Semantic Task Sampling. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2022, pp. 4767-4776. (CORE : A*)

Articles de workshop nationaux avec comité de lecture : 1

[WN1] M. Benhamza, M. Clausel, **M. Tami** : Counterfactual Robustness : a framework to analyze the robustness of Causal Generative Models across interventions. Workshop Fondements Mathématiques de l'IA, 2025.

Séminaires invités : 15

La liste de mes communications séminaires depuis 2018 sont répertoriées en sous-section [3.8.4](#).

Autres : 7

[A1] M. Benhamza, M. Clausel, **M. Tami** : AdaptiveCausalWVAE : Certifiably Robust and Stable Causal Disentanglement. PEPR IA Days, 2025.

[A2] N. Sella, A. Asadi, **M. Tami**¹⁵, L. Verny : MIIC-SR : From Complex Data to Structural Causal Models, Conference EGC'25 - Strasbourg, January 2025. (13p).

[A3] M. El Bouchattaoui, **M. Tami**, B. Lepetit, P.H. Cournède : CDVAE : Estimating causal effect over time under unobserved adjustment variables. Quarter Causality - Fundamental Challenges in Causality, may 2023.

[A4] A. G. Reisach, **M. Tami**, C. Seiler, A. Chambaz, and S. Weichwald : All Parameters Matter When Simulating Data for Causal Discovery. Quarter Causality - Fundamental Challenges in Causality, may 2023.

[A5] M. El Bouchattaoui, **M. Tami**, B. Lepetit, P.H. Cournède : Estimating causal effect over time under unobserved adjustment variable. Quarter Causality - When Causal Inference meets Statistical Analysis, april 2023.

3.3 Principales collaborations académiques et industrielles

Principales collaborations académiques :

- Université de Lorraine (Marianne Clausel, Ricardo Borsoi), supervision de stagiaires depuis 2023 et de doctorants depuis 2023, organisation des workshops fondements mathématiques de l'IA.
- CEA (Aurore Lomet), supervision d'une thèse en collaboration avec l'Université de Lorraine (Marianne Clausel et Ricardo Borsoi)
- Ecole Polytechnique (El Mahdi El Mhamdi), supervision d'un stage en cours et financé par la BNP en cybersécurité et IA. La soumission d'un projet ANRT est prévu pour une poursuite en thèse du stagiaire.
- Ecole Polytechnique (El Mahdi El Mhamdi), participation à la chaire SRAI (Safe And Reliable AI).
- Institut du Cancer de Montpellier (C. Bascoul-Mollevi), collaboration académique où j'ai étendu une méthode d'estimation par algorithme EM d'un SCM à leur cas d'usage clinique sur la Qualité de Vie des patientes atteintes du cancer du sein (2016).
- l'ETS (Ismail Ben Ayed, Pablo Piantanida - Montréal, Canada), collaboration académique avec échange de chercheurs invités.

15. J'ai assuré la co-supervision du postdoctorant (réunions mensuelles), participé au cadrage scientifique du projet et réalisé la relecture finale.

- CVN (Emilie Chouzenoux), soumission pour le financement par Data IA d'un chercheur invité et le lancement d'une collaboration académique à l'intersection des domaines de l'optimisation, des statistiques bayésiennes, d'estimation par algorithme EM et de la causalité.
- BRGM Service géologique national (Konstantinos Trevlopoulos), un projet ANR en cours (soumission en phase 2 effectuée).

Collaborations industrielles :

- TotalEnergies (Antoine Bertoncello, Louis Verny, Emmanuel Leborgne, Nadir Sella), financement d'un postdoctorat (2023-2024), financement d'un stagiaire (2026), thèse Cifre en montage (2026), construction d'un projet pour l'évaluation de l'enseignement Apprentissage automatique de 2023.
- Rakuten (Parantapa Goswami, Aloïs de La Comble), échanges autour d'un datachallenge organisé pour l'enseignement EI de la ST4 Data@web.
- SAFT (Karrick Mergo Mbeya, BENJAMIN Sebastien), stage, 2023.
- Saint Gobain (Benoit Lepetit), thèse Cifre, 2022-2025.
- Sidetrade (Antoine Toubhans, Clément Chastagnol), thèse Cifre, 2018-2021.
- SICARA (Antoine Toubhans, Pierre-Henri Cumenge), thèse Cifre, 2020-2023.
- IRT SystemX (Michel Batteux, Elies Gherbi), stage (2019), thèse Cifre (2019-2023), postdoctorat (2023-2024), échanges avec Eliel Gherbi autour de l'enseignement et qui a été chargé de TD du cours Apprentissage automatique durant 2 ans.
- Pollen Métrology (Johann FOUCHER, Julien BADEROT) montage d'un projet BPI (avorté en 2026 et en reconstruction pour une soumission en 2027)
- SAFRAN (Clément Remacha, Julian Betancur), thèse Cifre en montage pour un début en juin 2026. Une collaboration en montage pour financer un CDD de 3 mois au candidat identifié afin de le stabiliser d'ici le début de la thèse.
- BNP (Mariam Barry), Stage en cours (Mars 2026 - Novembre 2026), montage de thèse Cifre en cours pour fin 2026 ou début 2027.

3.4 Encadrements de thèses

Co-encadrements de thèses ¹⁶ :

Je compte 5 thèses co-encadrées soutenues, dont 4 en financement Cifre, et 3 sont en cours.

- Déc 2021 : Victor Bouvier, je n'ai pas officiellement co-encadré la thèse de Victor Bouvier (taux d'encadrement 0 %). En revanche, j'y ai participé ¹⁷ volontairement à partir de l'année 2019 jusqu'en 2021. Il s'agissait d'une thèse CIFRE avec Sidetrade. La directrice de thèse était Céline Hudelot.
 - Towards Adaptive Learning with Invariant Representations
 - Publications : [C7][C8][C9][C12]. L'une de ces publications a d'ailleurs reçu un prix [C7].
 - Devenir : Après sa soutenance de thèse Bouvier a hésité avec le souhait de faire une carrière académique. Mais, il a fini par s'orienter vers le secteur privé et a été recruté comme chercheur en IA chez Dataiku où il a exercé pendant un an et demi. Ensuite, il a décidé de rejoindre Suez en tant que "Data Scientist".
- Avr 2023 : Yassine Ouali, encadrement à 30 % avec Céline Hudelot (Université Paris Saclay, CentraleSupélec, MICS), financement MICS.
 - Label efficient Deep Learning methods for image segmentation
 - Publications : [C10][C13][C14].
 - Devenir : L'expérience de recherche réalisée chez Samsung AI pendant sa thèse a abouti à l'embauche de Ouali en tant que chercheur en IA et qui y est aujourd'hui très épanoui.
- Juil 2023 : Victor Pellegrain, encadrement à 60 % avec Céline Hudelot (directrice de thèse, taux d'encadrement de 20 %, MICS) et Michel Batteux (co-encadrant, taux d'encadrement de 20 %, IRT

16. En bas de page une référence sourçant ma contribution de co-encadrante, pour les cas où la thèse n'apparaît pas officiellement dans ma liste ADUM

17. Durant cette expérience, j'ai participé à quatre de ses publications dans des conférences internationales de haut rang et dont je suis co-auteur. L'une de ces publications a d'ailleurs reçu un prix. Je me suis investie dans le cadrage mathématique des travaux de thèse et de nombreuses relectures et aide à la rédaction. J'étais présente de façon hebdomadaire pour des points d'avancement et je l'ai soutenu durant les répétitions de ses communications orales. Je l'ai aussi soutenu durant ses démarches administratives pour d'éventuelles candidatures académiques.

SystemX). Il s’agissait d’une thèse CIFRE avec l’IRTSystemX. La directrice de thèse était Céline Hudelot.

- Deep Learning methods for monitoring and diagnosis from heterogeneous and complex data
- Publications : [C5][N4][N5].
- Devenir : Victor Pellegrin a pris quelques mois de vacances et il est aujourd’hui « Lead Data Scientist » chez ILLUIN Technology.
- Sept 2023 : Etienne Bennequin, encadrement à 50 % avec Céline Hudelot (directrice de thèse, taux d’encadrement de 50 %, MICS). Il s’agissait d’une thèse CIFRE avec Sicara. La directrice de thèse était Céline Hudelot.
 - Towards Interpretable Meta-Learning : Applications to Few-Shot Object Detection
 - Publications : [C6][C9][W3].
 - Devenir : À la suite de sa thèse, Etienne Bennequin a décidé de prendre une année pour se consacrer à une de ses passions : le théâtre. Plus précisément, son projet actuel est l’écriture d’une pièce de théâtre.
- Nov 2025 : Mouad El Bouchattaoui, encadrement à 50 % avec Paul-Henry Cournède (directeur de thèse, taux d’encadrement de 50 %, MICS). Il s’agissait d’une thèse CIFRE avec Saint Gobain. Le directeur de thèse était Paul-Henry Cournède.
 - Learning causality with machine learning for the analysis and prediction of customer behavior
 - Publications : [J1][C2][W2][N3][A3][A5]
 - Devenir : À la suite de sa thèse, Mouad El Bouchattaoui a décidé de poursuivre chez Saint Gobain en tant que Senior Data Scientist.
- Début en oct 2024 : Manal Benhamza, encadrement à 75% en tant que directrice de thèse avec Marianne Clausel (taux d’encadrement de 25%, Université de Lorraine, CRAN), financement ANR JCJC ClearDeep.
 - Apprentissage de représentations causales structurées pour plus de robustesse et de transparence
 - Publications : [C1][N1][N2][W1][WN1][A1]
- Début en jan 2026 : Adrien Letelier, encadrement à 25 % en tant que directrice de thèse avec Marianne Clausel (taux d’encadrement de 25%, Université de Lorraine, CRAN), Aurore Lomet (taux d’encadrement de 25%, CEA), Riccardo Borsoi (taux d’encadrement de 25%, Université de Lorraine, CNRS, CRAN,), financement CEA.
 - Apprentissage profond et apprentissage de représentations causales pour la détection d’anomalies – Application à la détection d’anomalies sur la ligne de lumière du synchrotron
- Début en avr 2026 : Guangjie Li, encadrement à 50 % en tant que directrice de thèse avec Marianne Clausel (Université de Lorraine, CRAN), financement ANR JCJC ClearDeep.
 - IA causale pour la décision séquentielle : apprentissage de politiques robustes sous contraintes causales

3.5 Encadrements de postdoctorants

J’ai co-supervisé deux postdoctorants, l’un en collaboration avec Elies Gherbi à l’IRT SystemX et l’autre en collaboration avec Louis Verny et Nadir Sella à TotalEnergies (Mai 2024 - Avril 2025). J’ai également supervisé un stagiaire ingénieur de recherche durant 6 mois en 2022 qui a soutenu une thèse à Paris Cité sous la direction d’A Chambaz et est en postdoctorat au Laboratoire de mathématiques d’Orsay sous la supervision de G Blanchard. Voici les informations principales liées à ces activités d’encadrement :

- Arefe Asadi, postdoctorante que j’ai encadré à 33 % avec Louis Verny (taux d’encadrement de 33 %)et Nadir Sella (taux d’encadrement de 33 %) de TotalEnergies (Mai 2024 - Avril 2025) dans le contexte de la problématique industrielle suivante : « Inférence causale pour un apprentissage de représentation et l’intervention causale sur la prédiction de la qualité du soudage d’éoliennes ». Les résultats ont abouti à une communication dans une conférence nationale à EGC’25 [A2]. Ce travail a été préliminaire à un travail récemment publié à NeurIPS 2026 par l’équipe TotalEnergies.
- Mohammed Lamine Bouchouia, postdoctorant co-encadré à 50 % avec Elies Gherbi à l’IRT SystemX (juin 2023 - mai 2024) et le sujet est celui de l’« Apprentissage causal pour la détection fiable des cyberattaques pour la voiture connectée et autonome ». Les résultats ont abouti à une publication conférence IEEE VNC 2024 (Japon) [C3].

- Alexander Gilbert Reisach, stagiaire de recherche de niveau ingénieur financé par le MICS. Ce stage a duré d'avril 2022 à septembre 2022 et portait sur une Étude « benchmark » sur l'impact des choix des paramètres sur la découverte causale. Pendant la durée du stage j'ai été l'encadrante principale. Puis à partir d'octobre 2022, Alexander a démarré une thèse de doctorat sous la supervision d'Antoine Chambaz. Ceci a permis d'initier une collaboration avec lui ainsi qu'une collaboration internationale avec d'anciens superviseurs d'Alexander Reisach. Ce travail a donné lieu à une publication à NeurIPS 2023 [C4].
- Publications associées : [A2][C3][C4].

3.6 Encadrements de stagiaires de M1-M2 ou assimilés

A ce jour, dans le cadre de l'UE Lab Research Projets (niveaux M2), j'ai supervisé les élèves et sujets suivants :

- 2023-2024 : Manal Benhamza et Abdelkarim El Ouafi, sujet : Generative modeling for Causal Representation learning, 2023-2024.
- 2021-2022 : Jean Loubet et Paul Loubet, titre : Causal learning for multi-sources data.
- 2020-2021 : Paul Dampierre et Sedki Nouredine, titre : Semi-Supervised Probabilistic Classification Trees.

A ce jour, j'ai supervisé au laboratoire MICS les stages de recherche de fin d'étude (M2) des élèves et sujets suivants :

- 2026 : Mohamed-Jalal Baim, avec EM El Mhamdi (Ecole Polytechnique, CMAP) et M Barry (BNP), titre : Causality Based Indirect Poisoning : Attacks and Defenses, niveau M2, source de financement : BNP
- 2026 : Yahya El Fataoui, titre : Unsupervised Anomaly Detection in Photovoltaic Time Series Using Generative Models, niveau M2, source de financement : TotalEnergies
- 2024 : Manal Benhamza, avec M Clausel (Université de Lorraine, CRAN), titre : Apprentissage de représentation latente causale structurée, niveau M2, source de financement : ANR JCJC CLearDeep
- 2023 : Ahmed Nait Slimen, titre : Estimation des intervalles de confiance d'un modèle de vieillissement de batterie Li-ion, niveau Master 1 à l'ENSTA, source de financement : collaboration avec SAFT.
- 2022 : Alexander Gilbert Reisach, titre : Un critère de tri invariant à l'échelle pour trouver un ordre causal dans les modèles de bruit additif, niveau Master 2, source de financement : MICS. Ce travail a donné lieu à une publication NeurIPS 2023 et une poursuite en thèse.
- 2022 : Thomas Dujardin, avec M Clausel (Université de Lorraine, CRAN), titre : Apprentissage multi-modal pour la prédiction d'émotions, niveau M2, source de financement : IECL (ancien laboratoire du co-encadrant de stage).
- 2021 : Maxime de Richaud, titre : Extension d'arbres de régression probabilistes "smooth" pour une tâche de classification, niveau M2, source de financement : MICS.
- 2019 : Victor Pellegrain, avec C Hudelot (MICS) et M Batteux (IRT SystemX), titre : Deep Learning pour les données de séries temporelles, niveau Master 2 MVA, source de financement : collaboration IRT SystemX. Ce stage a donné lieu à une poursuite en thèse avec la même équipe encadrante.
- Publications associées : [C4][A4][N5][N3].

J'ai aussi supervisé en tant que tutrice académique les stages de fin d'étude (proposés par des industriels ou des laboratoires de recherche) des étudiants et sujets suivants :

- 2025 : Manon Lagarde, titre : Intégration de l'IA pour accélérer la simulation d'un jumeau numérique de capteur ESM, entreprise : Thalès.
- 2025 : Henri Langlois, titre : Pragmatic Rewards : A Framework for Augmenting Multi-Agent Reinforcement Learning with Legible Behavior, laboratoire : ILLS, ETS (Montréal, Canada).
- 2025 : Maria Kontaratou, titre : Cross-Scanner Breast Adenocarcinoma Image Segmentation with Deep Learning Algorithms, laboratoire : Archimedes Unit (Grèce).
- 2024 : Abdelkarim El Ouafi, titre : Causal Representation Learning, entreprise : TotalEnergies.

- 2023 : Amandine Allmang, titre : Développement de méthode de régression Bayésienne et d'Inférence Causale pour le marketing de traitement pour des maladies rares, entreprise : Sanofi.
- 2023 : Marie Bouvard, Etude et prédiction des incertitudes d'ECG, entreprise : Cardiologs.
- 2022 : William Binquet, Prise en compte des modèles issus du "Machine Learning" pour la modification d'une plateforme, entreprise : InUse.
- 2021 : Adil Dinia, titre : Développement d'un modèle de Machine Learning pour identifier les patients HPV avec mauvais pronostic, entreprise : IFSBM Gustave Roussy.
- 2021 : Jean-Baptiste Gourlet, titre : Développement d'un modèle de calibration basé sur une population, entreprise : Novadiscovery.
- 2021 : Laurène Vaugante, titre : Définition et expérimentation d'un référentiel de règles et de bonnes pratiques IA Responsable, entreprise : EDF.
- 2021 : Maryam Cherradi, titre : Étude d'un outil pour l'analyse de la pureté des échantillons d'ADN produits, entreprise : DNA Script.
- 2021 : Quentin Mondon-Cancel, titre : automatisation du contrôle de lubrification sur les montres, entreprise : manufacture horlogère ValFleurier.

3.7 Projets financés

Je suis porteuse du **projet ANR JCJC (2024–2027, env. 340 k€) nommé CLearDeep (A CausaL Explanation-oRiented DEEP Structured framework adaptable to temporal data for trustability and robustness)**, qui a été sélectionné pour financement en juillet 2023. Ce projet vise à intégrer la causalité dans les modèles profonds pour renforcer explicabilité, robustesse et confiance. Il finance un stage, deux thèses, une décharge d'enseignement, des missions, du matériel et un workshop. J'en assure la direction scientifique et l'encadrement doctoral. Avant cette ANR, une sous-partie du projet avait obtenu un financement **Digicosme 2023** pour une thèse (méthodes causales explicables pour deep learning). Le financement a cependant été perdu suite au désistement du candidat recruté.

Je suis également membre du projet **IMPT 2023 – Institut des Mathématiques pour la Planète Terre** (98 k€), finançant un postdoctorant. Mon rôle est d'apporter une expertise en apprentissage statistique et causalité et de suivre les avancées scientifiques. Les porteurs sont M. Clausel (CRAN) et M. Cuntz (INRAE). J'ai par ailleurs dirigé ou co-dirigé plusieurs contrats industriels :

- **SAFT (2022–2023, 10 k€)** : vieillissement de batteries Li-ion; recrutement et supervision d'un stagiaire.
- **TotalEnergies (2022, 119 k€)** : inférence causale pour la qualité de soudure d'éoliennes; recrutement et co-supervision d'un postdoctorant.
- **IRT SystemX (2023, 15 k€)** : apprentissage causal pour la détection fiable de cyberattaques; co-supervision d'un postdoctorant.
- **Saint-Gobain (2021, 60 k€)** : inférence causale pour l'analyse de comportements clients; co-supervision d'un doctorant.
- **TotalEnergies (2026, 3 k€)** : apprentissage de représentations latentes et modèles de diffusion pour la détection d'anomalies photovoltaïques; co-supervision d'un stagiaire.
- **SAFRAN (2026, 60 k€)** : méthode d'accélération des analyses 3D des mesures Tomographiques des aubes de turbines; co-direction d'une thèse avec un début prévu pour juin 2026.

3.8 Responsabilités collectives en recherche et animation scientifique :

3.8.1 Activités éditoriales

Je suis invitée comme membre de comités de programmes de conférences internationales, workshops nationaux ou journaux de recherches. Parmi eux sur les 4 dernières années :

- UAI 2026 : Senior Area Chair
- Workshops nationaux « Fondements mathématiques de l'IA » (depuis 2023) : reviewer
- CVPR 2023 : reviewer
- NeurIPS 2021, NeurIPS 2022, NeurIPS 2023 : reviewer

- IJCAI 2019, IJCAI 2020, IJCAI 2021, IJCAI 2022, IJCAI 2023 : reviewer
- EURO Journal on Decision Processes (2021) : reviewer

3.8.2 Participation à des comités de programmes

J'ai été ou suis membre des comités de programme des événements scientifiques :

- Workshops nationaux « Fondements mathématiques de l'IA » (Paris, depuis 2023 et 2024)
- Quarter on Causality (Paris, 2022-2023)
- École d'été en IA de CentraleSupélec (Paris, été 2024)
- École d'été en statistiques et sciences des données pour les jeunes chercheurs d'Afrique francophone (M'Bour, Sénégal, 2019)

3.8.3 Participation à un réseau de recherche et consortium

J'ai participé au consortium interdisciplinaire MIMS INRAE sur l'analyse multi-sources pour des prédictions explicables, contribuant aux réflexions ayant conduit à un ANR piloté par M. Hanafi (ONIRIS). J'y ai présenté mes travaux sur les modèles à équations structurelles, leurs liens bayésiens et mes recherches en causalité.

3.8.4 Invitations séminaires et séjours de recherche

J'ai réalisé un séjour de recherche d'un mois à l'École de Technologie Supérieure (ETS) au Canada. Dans le contexte du co-encadrement de la thèse d'E. Bennequin, une collaboration internationale est née avec un autre doctorant de l'ETS, M. Boudiaf (supervisé localement à Montréal, Canada). Ce travail a donné lieu à une publication conférence de haut rang.

J'ai également été invitée aux présentations séminaires suivantes (seuls les communications séminaires depuis 2018 sont répertoriées ci-dessous) :

- Causal Inference from Longitudinal Data : Latent Adjustment Variables and Counterfactual Prediction, Séminaire de l'équipe A&O du LISN, 2026.
- Causal Inference from Longitudinal Data : Latent Adjustment Variables and Counterfactual Prediction, Séminaire de l'équipe Probabilités et Statistiques de l'IMAG, Université de Montpellier, 2026.
- Causal Inference from Longitudinal Data : Latent Adjustment Variables and Counterfactual Prediction, Séminaire de l'équipe INRIA Premedical, Montpellier, 2026.
- Introduction to Causality, Centre of Sustainable Infrastructure and Systems (CSIS), Mahindra University, École Centrale, Inde, 2025.
- Connections to ML from two variables case to multivariate case, présentation des chapitres 5 et 8 du livre de Peters et al.' (Elements of CI : foundations and learning algo., Peters et al., 2017) pour le groupe Causal-Tau, Gif-sur-Yvette, 2023.
- Weakly supervised learning for image data, invitée par le comité d'organisation d'ISSAI'23 (International Workshop on Smart Systems and Artificial Intelligence) pour une conférence plénière, Maroc, 2023.
- Structural equation models with latent variables, workshop MIMS (invitée par le consortium interdisciplinaire de l'INRAE), Paris, 2022.
- Uncertain Trees : Dealing with Uncertain Inputs in Regression Trees, séminaire MICS, Paris, 2019.
- Learning welding prediction (and classification) models from both heterogeneous variables and uncertain measures, Aix-Marseille Sciences Economiques (AMSE), Marseille, 2018.
- Learning welding prediction by random forest for uncertain measures and heterogeneous variables, invitée à MATHIAS2018, conférence internationale annuelle organisée par TotalEnergies, Paris, 2018.
- From structural equation modeling to latent causal representation learning for more trustable ML, séminaire LIVIA, Montréal, Canada, 2022.
- Estimation simultanée d'un modèle à une équation structurelle et de ses facteurs latents par algorithme EM, quatre séminaires invitée à INSERM Bordeaux, IECL Nancy, INRA Avignon, AgroParisTech Paris, 2018.

3.8.5 Expertise académique : comités et jurys

Participation à des comités de sélection pour les concours de Maîtres et Professeurs Assistants

J'ai été membre des jurys de recrutement suivants :

- Concours MCF en Géosciences Numériques et Data Science à l'Ecole Nationale Supérieure de Géologie de l'Université de Lorraine, année : 2026.
- Concours Professeur(e) Assistant(e) en Apprentissage Automatique et Sciences des données, au département de Mathématiques Appliquées de Polytechnique, année : 2024.
- Concours Maître Assistant(e) spécialité Data, Image, Multimédia à l'Institut Mines-Télécom Lille Douai, année : 2021.

Participation à des jurys de thèse et comités de suivi de thèse

A ce jour j'ai examiné deux thèses et dans les semaines qui arrivent j'ai une thèse à rapporter ainsi qu'une autre à examiner. Plus précisément, j'ai été ou serai membre des jurys de thèse des docteurs suivants :

- Ali Aghababaei Harandi, titre de la thèse : "Bridging Neural Compression and Semi-Supervised Learning : Optimal Rank Selection and Geometric Representations", Université Grenoble Alpes, date de soutenance prévue : juin 2026, mon rôle : rapportrice.
- Lucas Schott, titre de la thèse : "Diverse and adaptive adversarial training for robust reinforcement learning", Sorbonne Université, date de soutenance : 14 avril 2026, mon rôle : examinatrice.
- Benjamin Larvaron, titre de la thèse : "Modeling battery health degradation with uncertainty quantification", Université de Lorraine, date de soutenance : 21 mai 2024, mon rôle : examinatrice. Jérémie Assael, titre de la thèse : « Machine learning for ESG data in the financial industry », Université Paris Saclay, date de soutenance : 19 juin 2023, mon rôle : examinatrice
- Seyid Abdellahi Ebnou Abdem, titre de la thèse : « Contributions To Structural Equation Modeling With Applications In Educational Studies », Faculté des sciences de l'Université Mohammed V à Rabat, Maroc, date de soutenance : 13 juin 2023, mon rôle : examinatrice.
- Victor Bouvier, titre de la thèse : « Towards Adaptive Learning with Invariant Representations », Université Paris-Saclay, date de soutenance : 13 décembre 2021, mon rôle : invitée en tant que participante au co-encadrement. La raison est que j'ai été recrutée après le début de thèse.

A ce jour j'ai été membre des comités de suivi individuel de thèse des docteurs suivants :

- Coline Druart, titre de la thèse : « Méthodes d'apprentissage statistique pour l'identification de la tumeur d'origine dans les cancers de primitif inconnus », Oncologie computationnelle, Institut Curie, date : 2026.
- Malek Ben Salah, titre de la thèse : « Analyse de survie et régression contrefactuelle pour optimiser la stratégie séquentielle de traitement dans le cancer du sein métastatique », CentraleSupélec, MICS, date : 2025.
- Nikita Lagrange, titre de la thèse : « Apprentissage de graphes ancestraux par une approche search-and-score », Physique des cellules et cancer, Institut Curie, date : 2024.

Expertise et accompagnement de projets d'innovation

Membre du comité d'évaluation – experte IA, Paris Biotech Santé : évaluation de projets de startups en santé intégrant l'IA (25 dossiers/an) : revue scientifique préalable des dossiers, participation aux jurys de sélection (5/an) et analyse des approches algorithmiques et des stratégies liées aux données.

J'accompagne également ponctuellement certains porteurs de projets sur les aspects IA (validation méthodologique, stratégie technologique et valorisation scientifique).

3.8.6 Diffusion logicielle et science ouverte

Les implémentations des méthodes présentées dans mes publications sont disponibles en libre accès sur :

- [C2] <https://github.com/moad-lihoconf/causal-cpc>
- [W2] https://github.com/moad-lihoconf/crl_transparency
- [J1] <https://github.com/moad-lihoconf/cdvae>
- [C4] <https://causaldisco.github.io/CausalDisco/>
- [C6] <https://github.com/ebennequin/few-shot-open-set>

- [W3] <https://github.com/sicara/semantic-task-sampling>
- [C7] et [C12] <https://github.com/vbouvier/ruda>
- [C9] <https://github.com/ebennequin/meta-domain-shift>
- [C10] <https://github.com/yassouali/SCL>
- [C11] <https://github.com/myriamtami/scikit-learn> et <https://gitlab.com/sami.courie/pr-tree>
- [C13] https://yassouali.github.io/autoreg_seg/
- [C14] <https://github.com/yassouali/CCT>

3.8.7 Diffusion du savoir et médiation scientifique : responsabilités et activités au sein de sociétés savantes ou associations

Engagement dans les sociétés savantes

Depuis 2013, je suis engagée dans différentes actions de diffusion scientifique et au sein de sociétés savantes. Je suis membre de la SFdS depuis 2013, où j'ai été élue au Groupe Jeunes (2018–2021) et trésorière (2019–2021). Dans ce cadre, j'ai notamment co-organisé une conférence internationale de type ISBIS ainsi qu'une école d'été en statistique et science des données au Sénégal.

Je suis également membre de la SMAI, de l'AFIA et de la SSFAM.

Organisation d'événements scientifiques et groupes de travail

Depuis 2023, je contribue ponctuellement au groupe Datacraft (<https://datacraft.paris/>), où j'ai co-construit et animé un tutoriel consacré à l'inférence causale.

La même année, j'ai été élue au groupe MALIA de la SFdS, au sein duquel je co-organise, en tant que membre du comité scientifique, les workshops *Fondements mathématiques de l'IA*. Je suis également membre du comité scientifique et contribue ponctuellement à l'organisation de ces workshops, financés par DataIA, l'UPS et SCAI.

Entre 2022 et 2023, j'ai participé au groupe Causal-Tau et contribué à la préparation du trimestre thématique international *A Thematic Quarter on Causality*, financé à hauteur de 150 k€ par AISSAI. Cet événement comprenait écoles thématiques, conférences et un data challenge que j'ai co-organisé avec mon ancien doctorant M. el Bouchattaoui sur un cas d'usage de Saint-Gobain.

Plus récemment, j'ai rejoint le Groupe de Travail ECLAIRE (*Causality, Learning, Analysis, Inference, Reasoning & Explainability*) au sein du GDR RADIA, dont l'objectif est d'éclairer les méthodes d'apprentissage automatique et d'IA par le prisme de l'inférence causale.

Mentorat et accompagnement de jeunes chercheurs

J'ai été mentore d'une doctorante en mathématiques entre 2021 et 2022 dans le cadre du programme de mentorat Femmes & Sciences pour jeunes chercheuses de l'Université Paris-Saclay.

En 2019, j'ai également participé aux Journées de la Statistique (JdS) afin de contribuer à l'orientation des jeunes doctorants, en présentant différentes trajectoires professionnelles ainsi que le processus de candidature pour l'obtention d'un poste d'enseignant-chercheur.

Participation à des actions de diffusion et médiation scientifique

Podcast ResponsAIble (Karine Majdalani, Anjuli Manche) : discussion autour de l'IA responsable, notamment sur les questions de fiabilité, d'interprétabilité et de confiance dans les systèmes d'IA. Cet échange a également porté sur la place des femmes en IA, le mentorat, les carrières en recherche et la conception de systèmes d'IA robustes et explicables.

3.8.8 Activités institutionnelles pour mon laboratoire de recherche MICS

Membre du comité de pilotage du Hub IA de CentraleSupélec

Depuis sept. 2023 je représente le laboratoire MICS, et suis suppléante de sa directrice au sein du Copil, je contribue à la coordination, la visibilité et la promotion des actions en IA à CentraleSupélec. Mes activités incluent la participation aux orientations stratégiques, le soutien aux initiatives étudiantes, la mise à disposition de ressources de calcul et l'organisation d'événements scientifiques. J'ai également accompagné

le coordinateur du Hub IA, favorisé le lien avec les doctorants et chercheurs du MICS, et conçu un projet de séminaire mensuel labellisé « Hub IA » visant à renforcer les collaborations académiques et industrielles et la rencontre des étudiants.

Missions transverses au MICS : Industrie 4.0 et Développement durable

Entre 2019 et 2022, j'ai recensé les activités de recherche et d'enseignement du laboratoire sur les thématiques Industrie du futur et Développement durable. Mes actions ont été les suivantes : contribution aux cartographies institutionnelles de CentraleSupélec ; participation aux réunions pilotées par la direction ; sensibilisation des équipes et promotion des outils Labo1point5 pour l'évaluation de l'impact environnemental ; réalisation et présentation de cartographies thématiques en conseil de laboratoire.

3.9 Récapitulatif des publications scientifiques

La liste complète de mes publications se trouve en section 3.2. J'en présente ici une synthèse avec les indicateurs usuels.

	IF/CORE ranking	Nombre	Nom des conférences ou revues
Revue internationale	pas encore d'IF pour tmlr ; 1.8 (Q1) pour Statistics in Medicine	2	Transactions on Machine Learning Research ; Statistics in Medicine
Conférences internationales	A*	3	NeurIPS
	A*	1	EMNLP
	A*	3	CVPR
	A*	1	IJCAI
	A*	1	ICLR
	A*	1	ECCV
	A	1	UAI
	A	5	ECML
	A	1	AISTATS
	Aucun	1	COMPSTAT
	Aucun	1	CASI
	Aucun	1	VNC
Conférences nationales		5	JdS
		2	Quarter Causality, Grenoble
		2	Quarter Causality, Paris
		1	APIA affilié à PFIA
		1	EPICLIN 9
ArXiv e-Print	aucun	1	Nombre de citations : 625. Volonté du doctorant de laisser cet article en libre accès sur arXiv.
	Total	34	

Myriam Tami
Université Paris-Saclay
Ecole CentraleSupélec 91190 Gif-sur-Yvette, France