



Institut MIAI@Grenoble Alpes
(Multidisciplinary Institute for Artificial Intelligence / Instituts 3 IA)
<https://miai.univ-grenoble-alpes.fr/>

Chaire « Ethique et IA »

Responsable de la chaire :

Thierry Ménissier, Prof. Univ. Grenoble Alpes, philosophie,
thierry.menissier@univ-grenoble-alpes.fr

La chaire "éthique et IA" est ancrée en philosophie, mais se veut pluridisciplinaire afin de faire dialoguer l'informatique et la robotique avec les sciences humaines et sociales. Son activité a pour objectif une meilleure compréhension des enjeux psycho-sociaux, moraux et politiques du déploiement de l'IA, ainsi que la détermination de règles d'organisation éthiques, durables et équitables, en vue de solutions IA compatibles avec les valeurs de la démocratie.

Sommaire

I/ Description de la chaire

Éthique philosophique, apports de l'IPhiG, membres de la chaire, dialogue SHS / STS

II/ Projet scientifique

3 axes directeurs, 5 défis scientifiques

III/ Gouvernance

Animation de la chaire et Comité d'orientation scientifique

IV/ Partenariats académiques internationaux

Chercheurs étrangers impliqués

V/ Partenariats extra-académiques

Sociétés en soutien de la chaire et en accompagnement

Recherches éthiques appliquées, proposition de « diagnostics éthiques »

VI/ Programme de formation à l'éthique

Outils pour l'évaluation éthique mis à disposition des chercheurs et des partenaires

VII/ Personnels recrutés dans le cadre de la chaire

Modalités de recrutement et doctorats, post-doctorats, stages de recherche

I/ Description de la chaire

Comprise parmi les chaires de l'institut MIAI@Grenoble Alpes, la chaire « Ethique et IA » oeuvre, pour une durée de 4 ans (2019-2023) à la connaissance de l'intelligence artificielle, par le biais de la recherche, de la formation et de la valorisation. Elle ancrée dans la discipline philosophique et reflète les activités de l'Institut de Philosophie de Grenoble (IPhiG), unité de recherche de l'Université Grenoble Alpes. Ces activités se déclinent en méta-éthique, éthique normative, éthique appliquée, philosophie politique, philosophie du droit et philosophie de la cognition. Cette variété d'angles de vue promet un potentiel élevé de la chaire en termes de traitement approfondi des sujets de recherche auquel donnent aujourd'hui lieu les relations entre éthique et IA. Le porteur du projet de chaire, pour sa part spécialisé en éthique publique et en philosophie de l'innovation, est garant de cette variété portée par son équipe de recherche.

La chaire contribue au pilier « éthique et société » de MIAI, et entretient des relations soutenues avec les recherches menées dans les autres piliers de l'institut, à commencer par celles développées dans le « cœur de l'IA ». Des chercheurs en informatique sont à ce titre invités à faire partie du Comité d'orientation scientifique (COS) qui veille à la gouvernance de la chaire.

Dans le cadre des projets qu'elle développe, la chaire associe des chercheurs d'autres disciplines issues du secteur santé (médecine, pharmacie) ou des sciences humaines et sociales (telles que : psychologie clinique et sociale, sciences de l'information et de la communication, sciences de gestion et marketing – liste non exhaustive).

Sont membres de la chaire les unités et les personnes suivantes :

- Laboratoire d'Informatique de Grenoble (UMR LIG : informatique et robotique) : Véronique Aubergé, Patrick Loiseau, Damien Peller, Sylvie Pesty
- INRIA (sciences du numérique) : Claude Castellucia, Cédric Lauradou
- Laboratoire Interuniversitaire de Psychologie (LIP-PCS2, UGA-USMB : psychologie clinique et sociale) : Emmanuel Montfort,
- Laboratoire de Psychologie et de Neurocognition (UMR LPNC : psychologie cognitive) : Aurélie Campagne
- Centre de recherche appliqué à la gestion (CERAG, UGA : management, marketing, finance, comptabilité, logistique, innovation) : Agnès Helme-Guizon, Karine Samuel, Emilie Hoareau
- Groupe de recherche sur les enjeux de la communication (Gresec, UGA : sciences de l'information et de la communication) : Fabienne Martin-Juchat.

II/ Projet scientifique

Construite autour de trois axes, la chaire « Ethique et IA » entend relever 5 défis scientifiques majeurs.

Le **premier axe** vise à **questionner le sens de l'intelligence de l'IA**, selon les termes du Rapport Villani (2018), où le mot « éthique » est entendu en fonction de la compréhension ou de la réflexivité nécessaires pour saisir les apports de l'IA à la société contemporaine. Ce qui implique de déterminer quels sont les modèles dont on se sert pour qualifier l'intelligence dans l'expression « IA ». Ce point constitue un problème majeur pour la philosophie d'aujourd'hui, et porte notamment sur le statut épistémologique et ontologique de la calculabilité, de l'incomplétude et de l'incertitude. Tel est le **défi scientifique n°1**, transversal à l'ensemble de MIAI : voué à répondre à la question « *qu'est-ce que l'intelligence dans et pour l'intelligence artificielle ?* », il engage un dialogue avec les sciences de l'informatique et de la cognition, et tout particulièrement avec les chaires MIAI « Parole » (Pascal Perrier) et « Langage » (Laurent Besacier).

Le **deuxième axe** concerne les projets développés par **l'explicitation des finalités poursuivies par le développement de l'IA, ainsi que par la détermination des valeurs qui sous-tendent ce dernier**. En quoi le ou les paradigmes d'intelligence IA donnent un nouveau sens ou font déjà sens pour la société ? Comment l'IA fait-elle évoluer les pratiques, usages & représentations

humains ? Tel est le *défi scientifique n°2 : quel type de société cherche-t-on à promouvoir avec l'IA ?* Ces interrogations regardent des domaines d'étude et de compréhension très variés, qui donneront lieu à des partenariats avec des partenaires extra-académiques (industriels et les collectivités territoriales) déclinés sous des formes variées (thèses CIFRE, études particulières, workshops réguliers). Voici quelques-uns des secteurs concernés et des thèmes qui seront abordés du fait de leur évolution sous l'effet de l'affirmation de la puissance algorithmique : production de la valeur et économie-finance (cas du trading haute fréquence) ; autonomie humaine de mobilité et mobilité automatiquement déterminée (comment dépasser les classiques et macabres dilemmes socio-moraux puisque pour ceux il n'existe pas de solution universelle, comme l'a montré Maxmen 2018 ?) ; détermination des valeurs d'usage, acceptabilité, émotions et marketing (*nudging*, incitations & motivations) ; domaines touchés par l'émergence de la blockchain : finance, contrats, tiers de confiance (peut-on instituer technologiquement la confiance ? Ménessier 2018) ; institution et politique : « gouvernementalité algorithmique » (Rouvroy et Berns 2013), contrôle automatique des comportements, sécurité vs. libertés publiques et privées ; émergence de l'IA dans les domaine(s) de l'art : musique, arts graphiques et plastiques, architecture & urbanisme ; médecine prédictive et génomique ; assurances et mutuelles, calculs des risques et modifications des conduites humaines. Sur plusieurs de ces thèmes, le travail de la chaire « Ethique et IA » se fera en collaboration d'une part avec la chaire MIAI « The Social Fabric of AI » (Gilles Bastin), de l'autre avec la chaire « Deep care » (Philippe Cinqin).

En annexe de ces interrogations se situent les questions liées aux nouvelles relations entre l'humain et la machine, dans le champ de la robotique sociale. Elles seront travaillées en partenariat avec Véronique Aubergé, responsable à Grenoble-INP de la chaire « Robots'Ethics ».

Le **troisième axe** regroupe la recherche à propos des normes et de la régulation de l'IA. **Contribuant à déterminer et à formaliser les règles d'éthique valables pour l'IA**, la chaire se donne pour objectif de **constituer le centre français d'éthique de/pour l'IA**.

Relevant de cet axe, le *défi scientifique n°3* peut se formuler ainsi : *Quelles sont les procédures, les règles et les valeurs les plus pertinentes tant pour la conception que pour les pratiques et les usages de l'IA ?* La dimension de l'éthique apparaît ici comme une valeur importante dans l'affirmation de l'IA. Par exemple, quelles sont les « théories de la Justice » (au sens de Rawls 1971) pour le monde de l'IA ou en régime généralisé d'IA ? Comment, dès la phase de conception des algorithmes, se pose le problème philosophique de l'allocation de moyens ou justice distributive ? Ainsi le *défi scientifique n°4*, appelé par le précédent, se formule ainsi : *quelle est la bonne perspective pour déterminer l'IA de manière démocratique ?* Les défis n°3 et 4 seront travaillés en étroite collaboration avec la chaire MIAI « Legal and Regulatory Implications of Artificial Intelligence » (Théodore Christakis). Par exemple, des workshops réguliers inter-chaîres auront lieu pour l'examen croisé de ces questions, portant notamment sur les aspects philosophiques et juridiques de la régulation de l'IA (Thomas Boccon-Gibod, maître de conférences de philosophie du droit dans l'IPhiG sera plus précisément en charge de ces aspects).

Enfin, quel type d'éthique sera le mieux adapté pour l'IA ? Penser philosophiquement l'IA permet (ou risque) donc de faire bouger l'état de l'art en théorie morale. De nouvelles questions et de nouveaux concepts sont déjà apparus, comme la *fairness* des algorithmes. D'où le *défi scientifique 5 : en quoi l'IA fait-elle évoluer (l'état de l'art de) l'éthique ?* Cette question sera abordée en discussion avec la chaire MIAI « Explainable and Responsible AI » portée par Marie-Christine Rousset et Patrick Loiseau.

III/ Gouvernance

Le responsable de la chaire gère cette dernière en s'appuyant sur l'administration de l'institut MIAI. Il rend des comptes de son action au Comité d'orientation scientifique (COS) de la chaire, selon la fréquence d'une réunion mensuelle. Le COS a les attributions suivantes :

- il conseille le responsable de la chaire en matière de stratégie de recherche, ses avis étant consignés par écrit,
- il vote à la majorité des membres présents sur l'attribution des moyens de la chaire en fonction des propositions d'actions de recherche que lui soumet le responsable de la chaire.

Outre du responsable de la chaire, le Bureau est composé des membres suivants :

- le directeur de l'équipe IPhiG (Pr. Denis Perrin) ou de sa/son représentant.e,
- du directeur de l'institut MIAI (Pr. Eric Gaussier) ou sa/son représentant.e,
- d'un.e chercheur/se volontaire issu d'une des unités de recherche en informatique du site de Grenoble ou de l'INRIA (volontaire pour l'année en cours : Pr. Sylvie Pesty, LIG),
- d'un.e chercheur/se volontaire issu d'une des unités de recherche en sciences humaines et sociales de l'Université de Grenoble volontaire pour l'année en cours : Pr. Agnès-Helme Guizon, CERAG).
- d'un.e représentant.e des partenaires extra-académiques (volontaire pour l'année en cours : un.e représentant.e de STMicroelectronics).

IV/ Partenariats académiques internationaux

La chaire entend développer des partenariats internationaux et nationaux. Les premiers se comprennent en fonction de la volonté de s'inscrire dans la durée. Les premiers collègues philosophes ayant accepté de collaborer sous plusieurs formes (séjour moyenne et longue durées, PhD en cotutelle, workshops croisés, etc.) sont les suivants :

- Pr. Alessandro Arienzo, Université Federico II, Naples, Dipartimento di Studi Umanistici, Italie,
- Pr. Thomas Berns, Université Libre de Bruxelles, Centre Perelman de philosophie du Droit, Belgique,
- Pr. Marc-Antoine Dilhac, Université de Montréal, Chaire de Recherche du Canada en éthique publique (CRSH), Département de philosophie, Canada,
- Pr. Paul Dumouchel, Université Ritsumeikan, Kyoto, Graduate School of Core Ethics and Frontier Sciences, Japon,
- Pr. Alain Létourneau, Université de Sherbrooke, Département de philosophie et d'éthique appliquée, Canada.

V/ Partenaires extra-académiques :

Conformément à l'esprit de l'appel à projets ANR Instituts 3IA, la chaire s'appuie sur des partenariats avec des acteurs industriels et entrepreneuriaux d'une part, et avec des acteurs publics de l'autre. La recherche sur les relations entre éthique et IA, si elle veut être pertinente, doit être socialement utile en portant sur les cas concrets et sur les difficultés rencontrées par de tels acteurs en cette période de transition technologique et sociale profonde. Il ne peut en être autrement pour une démarche qui sera orchestrée par le porteur de la chaire, à savoir, un philosophe ayant depuis de nombreuses années mis à son agenda scientifique la thématique de l'innovation, justement en lien avec de tels partenaires.

Les sociétés suivantes soutiennent la chaire :

- STMicroelectronics (https://www.st.com/content/st_com/en.html)
- COMONGO (<https://comongo.fr/launch/>)

Elle est également accompagnée par :

- Sylha (<https://sylha.io/>)

La participation aux activités de la chaire est possible pour les entreprises, les collectivités et les associations. Elle sera formulée par un courrier motivant la demande, adressé au porteur de la chaire, qui instruira la demande et la soumettra au COS afin que celui-ci rende un avis.

Les formes de collaboration possibles sont diverses : participation aux séminaires mensuels de la chaire et accès aux informations qu'ils génèrent, contrats doctoraux en financement CIFRE, workshops dédiés sur des thématiques co-construites avec les chercheurs et les acteurs industriels et publics, etc.

Une option originale consiste à être en mesure de proposer des évaluations diagnostiques des projets d'IA sur les différents plans possibles concernés et impliqués par la démarche éthique : détermination de principes, de procédures et de valeurs axiologiques permettant de juger du caractère

éthique des projets ou programme d'IA ; compréhension des enjeux induits par la démarche éthique ; observation des effets de transformation produits par elle. Cette démarche est originale du fait qu'elle est différente de celle qu'on entend généralement par « acceptabilité » des technologies ; elle converge avec le vœu, exprimé plus haut, de militer pour une IA non seulement socialement acceptable, mais juste. Les domaines de l'IA en santé et pour l'environnement, les deux domaines d'application de MIAI, sont d'ailleurs tout-à-fait propices à une telle démarche, puisqu'ils concernent des biens communs ou des ressources précieuses de l'humanité.

Le porteur de la chaire se trouve déjà en capacité de délivrer de telles évaluations à propos des innovations technologiques et sociales. Les bénéficiaires de ces évaluations pourraient être les projets de recherche en IA, les start-up incubées dans MIAI, et bien entendu les partenaires de la chaire. Cette démarche pourrait enfin être transférée comme savoir-faire méthodologique en direction des doctorants en IA et en SHS de MIAI, par exemple dans le cadre d'un accompagnement par la SATT Linksium.

VII/ Programme de formation à l'éthique

La chaire entend également développer la formation à l'éthique, **orienté en vue d'une recherche en IA nourrie par les SHS, et d'une conception de l'IA efficiente du point de vue éthique**. Ce programme est destiné en priorité :

- aux chercheuses et chercheurs titulaires impliqué.e.s dans l'institut MIAI,
- aux jeunes chercheuses et chercheurs doctorant.e.s financé.e.s par lui et d'autres doctorant.e.s du site grenoblois (doctorant.e.s en informatique, en philosophie et dans d'autres SHS),
- aux conceptrices et concepteurs d'IA (ingénieur.e.s en informatique, cadres d'entreprises et de collectivités territoriales),
- aux étudiant.es en informatique et robotique intéressé.es (niveau master).

Ce programme, lié aux activités de la chaire, est distribué dans l'ensemble de l'institut MIAI, ouvert à l'ensemble des contributrices et contributeurs à ce dernier. Il vise à aider à la conception d'une IA qui, parce qu'elle est déterminée aussi par la réflexion éthique, est rationnellement désirable et plus juste.

VIII/ Personnels recrutés dans le cadre de la chaire

Sur l'ensemble des quatre années de sa durée, la chaire recrute des personnels de recherche sur contrat temporaire, selon les modalités en vigueur dans la recherche française : contrats doctoraux et post-doctoraux dans le cadre de l'UGA sur les fonds alloués à MIAI, contrats doctoraux sous convention CIFRE, contrats de stage de recherche de courte durée rattachés aux unités de recherche engagés dans la chaire.

Pour l'année 2019-2020, les personnels recrutés sont les suivants (état au 29/09/19) ;

- Juin-juillet 2019 : Eva Pasini, étudiante de master 2 de philosophie, ENS-Lyon (stage de recherche).
- 1^{er} octobre 2019-30 septembre 2020 : Dr. Tyler Reigeluth, Université Libre de Bruxelles (post-doctorat).
- 1^{er} octobre 2019-30 septembre 2022 : Seddik Benlaksira, doctorat en co-direction Gresec-IPhiG.