

Haut Conseil de l'Évaluation de la Recherche et
de l'Enseignement Supérieur



DOCUMENT D'AUTOÉVALUATION
Équipe LFI



Campagne d'évaluation 2023-2024 — Vague D

Table des matières

1	INFORMATIONS GÉNÉRALES SUR L'ÉQUIPE LFI	3
1.1	Les thématiques scientifiques et leurs enjeux	3
	Positionnement scientifique par rapport au contexte international	3
	Collaborations	4
	Animation scientifique	4
2	INTRODUCTION DU PORTFOLIO	5
3	AUTOÉVALUATION DU BILAN	6
3.1	Autoévaluation de l'équipe	6
	Domaine 2. Attractivité	6
	Domaine 3. Production scientifique	7
	Domaine 4. Inscription des activités de recherche dans la société	8
4	RÉFÉRENCES BIBLIOGRAPHIQUES SIGNIFICATIVES DE LFI	9
A	ANNEXE — MEMBRES PERMANENTS AU 31/12/2022	10

1 INFORMATIONS GÉNÉRALES SUR L'ÉQUIPE LFI

Nom de l'équipe : Learning, Fuzzy and Intelligent systems (LFI)

Responsable de l'équipe : Christophe Marsala

	2017	2018	2019	2020	2021	2022
PR	1	1	1	1	2	2
MCF HDR	1	1	1	1	1	1
MCF	1	1	1	2	2	2
DR	0	0	0	0	0	0
CR HDR	0	0	0	0	0	0
CR	0	0	0	0	0	0
Total permanents	3	3	3	4	5	5
Émérites	1	1	1	1	1	1
Doctorants	14	6	5	6	4	6
Ingénieurs CDD ou hors tutelles	0	0	0	0	0	0
Post-doc, ATER, etc.	0	1	0	1	0	0
Stagiaires	3	3	4	3	6	9
Total non permanents	17	10	9	10	10	15
Total avec émérites	21	14	13	15	16	21
Equivalent temps plein recherche	1.5	1.5	1.5	2.0	2.5	2.5

TABLE 1 – Personnels LFI sur la période 2017-2022 (au 1er juillet de chaque année)

Sur la période, l'équipe a été renforcée par le recrutement de Jean-Noël Vittaut (MCF SU) en septembre 2019. De plus, Isabelle Bloch, recrutée sur la Chaire Sorbonne Université/SCAI en Intelligence artificielle pour 4 ans (2020-2024) a rejoint notre équipe à l'arrivée sur sa chaire. Le nombre de doctorants a, de son côté, été stabilisé sur la période pour réduire le taux d'encadrement (cf. section ??).

1.1 Les thématiques scientifiques et leurs enjeux

Positionnement scientifique par rapport au contexte international

Les recherches de l'équipe *Learning Fuzzy and Intelligent systems* (LFI) ont pour cadre le développement théorique et pratique d'approches interprétables d'intelligence artificielle pour l'aide à la décision, la science des données et l'apprentissage automatique.

Systèmes flous, intelligence artificielle explicable, expressivité. Le besoin de systèmes d'intelligence artificielle facilement explicables à l'utilisateur final et fournissant des résultats qu'il comprenne clairement a été mis en évidence depuis plusieurs années. Tous les aspects de l'intelligence artificielle sont concernés, des ontologies aux systèmes à base de règles, de l'apprentissage à la classification ou l'aide à la décision. Les membres de l'équipe LFI ont toujours développé des approches tenant compte d'un tel objectif d'interprétation, en favorisant l'intégration de la sémantique à différents niveaux : la représentation des données, les connaissances qui en sont extraites, ou connues a priori, ainsi que les outils d'extraction et de raisonnement. Ils s'inscrivent en cela dans le contexte de l'eXplainable Artificial Intelligence (XAI).

Nos travaux exploitent le cadre formel de logiques non classiques, notamment pondérées (logiques floue, possibiliste ou multi-valente), ainsi que les systèmes flous qui sont des méthodes particulièrement adaptées pour rendre interprétables de tels systèmes étant donné leur proximité avec le raisonnement humain et leur capacité à manipuler des informations linguistiques.

IA hybride et explicabilité. L'équipe développe aussi ses travaux pour une IA hybride, c'est-à-dire impliquant différents formalismes d'IA (logique, flou, graphes, apprentissage), avec un fort ancrage dans les formalismes algébriques (treillis, dans différents formalismes mathématiques).

Dans le domaine de l'apprentissage automatique, nous travaillons aussi sur la recherche de modèles d'apprentissage par renforcement généraux capables de résoudre une grande quantité de problèmes décrits sous la forme de systèmes de transitions d'états modélisés sous la forme de règles logiques. Toujours dans ce domaine, nous développons des travaux sur l'utilisation et l'extension des outils d'apprentissage automatique interprétables comme, par exemple, l'algorithme d'apprentissage par arbres de décision, pour traiter des données imprécises, subjectives, et des émotions et qui peuvent être appliqués dans des environnements dynamiques.

Apprentissage automatique. L'équipe LFI développe des approches d'apprentissage automatique interprétable (interprétabilité "par construction") pour construire des modèles interprétables capables de produire des décisions

explicables. Dans ce cadre, nous intégrons des techniques issues de l'intelligence computationnelle telles que la théorie des sous-ensembles flous pour représenter des données numériques à l'aide d'étiquettes linguistiques. Nous réalisons des travaux à la fois dans le domaine de l'apprentissage supervisé et dans le domaine de l'apprentissage non-supervisé. Les approches classiques sont utilisées pour produire des modèles explicables (par exemple, l'apprentissage par renforcement pour construire des bases de règles floues) ou sont étendues pour améliorer leur interprétabilité (par exemple, l'apprentissage par arbres de décision flous, le clustering, ou l'introduction de connaissances sémantiques dans les modèles d'apprentissage profond).

Fiabilité de la source, qualité des attributs, description du modèle, satisfaction de l'utilisateur. Un axe plus transverse de nos recherches a pour cadre la qualité des données et de l'information qui est, par exemple, un des problèmes majeurs en science des données. Les sources de données sont plus ou moins expertes sur le sujet traité, pertinentes, fiables, et elles peuvent fournir des données imparfaites. Les données obtenues peuvent être imprécises, incertaines, incomplètes, bruitées, contenant éventuellement des erreurs. modèle d'intelligence artificiel lui-même doit être compréhensible par

Collaborations

D'un point de vue collaboratif, notre équipe s'attache à profiter de notre environnement scientifique. Ainsi nous réalisons des travaux de recherche avec différentes équipes du LIP6, soit sous la forme de projets communs (projets LIP6 par exemple), soit par le co-encadrement de thèse. Ainsi, avec l'équipe ACASA, nous travaillons sur les liens entre argumentation et modèles causaux, entre argumentation et langages d'actions, explications contrastives dans un cadre flou, mais aussi en éthique computationnelle. Avec l'équipe SYEL, nous développons des approches pour l'intégration d'approches d'IA pour des systèmes embarqués et pour la reconnaissance de structures dans des images dans le domaine médical, ou pour la reconnaissance des émotions. Des travaux communs ont aussi été menés avec les équipes BD, RO et SMA dans le cadre de projets LIP6.

En dehors du LIP6, nous collaborons, au sein de Sorbonne Université avec l'ISIR et le LIMICS, et, à l'extérieur de Sorbonne Université, avec le département "Cognitions Humaine et Artificielle" de l'Université Paris 8, ainsi qu'avec l'université Paris Nord, Centrale-Supélec, l'Université d'Artois, Dauphine, Télécom Paris.

Sur l'explicabilité, des travaux sont menés en collaboration avec l'ISIR et avec l'équipe ACASA (liens entre argumentation et modèles causaux, entre argumentation et langages d'actions, explications contrastives dans un cadre flou), et d'autres travaux sont menés en abduction par morpho-logique, raisonnement spatial.

Des applications sont développées en imagerie médicale (avec l'ISIR, Télécom Paris, et plusieurs hôpitaux de l'AP-HP), sur les représentations symboliques de la musique (avec STMS/IRCAM), en histoire (sceaux byzantins avec la faculté de lettres, images historiques avec l'EHNE et l'université Paris Cité).

Animation scientifique

L'équipe s'investit dans de nombreuses activités permettant de faire connaître ou de développer ses domaines de recherche. Nous organisons régulièrement des séminaires scientifiques en invitant des chercheurs à présenter leurs travaux récents. Ainsi, nous avons organisé trois séminaires en 2020, trois en 2021, et trois en 2022. Cette organisation de séminaires se fait conjointement avec la section d'IEEE France. A cela s'ajoutent les séminaires organisés dans le cadre du laboratoire commun TRAIL (2 séminaires organisés en 2022, année de son lancement).

En matière d'organisation scientifique, nous avons par ailleurs pris en charge, en 2021, l'organisation de la conférence nationale sur la logique floue et ses applications (LFA) qui est la conférence majeure de notre domaine pour la communauté francophone. Nous avons également organisé la conférence Belief'2022 à Sorbonne Université, en collaboration avec l'Université Paris Saclay, Nos membres permanents sont aussi souvent sollicités pour participer à l'organisation (*scientific chair* par exemple) de conférences internationales de notre communauté (IPMU, Fuzz-IEEE, WCCI, etc.).

D'autre part, afin de pouvoir enrichir nos thématiques de recherches et initier des collaborations, nous avons accueilli des chercheurs d'universités étrangères (Gabriella Pasi (univ. de Milan), Davide Petturiti (univ. de Perugia), Marek Reformat (univ. d'Edmonton)) ou des collègues d'universités françaises lors de leur délégation CNRS au LIP6 (Cyril de Runz (univ. Reims), Fadi Badra (univ. Paris Nord)).

2 INTRODUCTION DU PORTFOLIO

- ▶ **Élément 1 (publication)** : “Recherches en IA explicable dans l’équipe LFI du LIP6” (C. Marsala, I. Bloch, M.-J. Lesot, S. Tollari & J.-N. Vittaut, Bulletin de l’AFIA, 2022). Cet article présente nos recherches en intelligence artificielle explicable et offre une vision transverse à l’équipe de nos travaux dans ce domaine.
- ▶ **Élément 2 (vidéo)** : “Contextualization and exploration of local feature importance as explanations to improve understanding and satisfaction of non-expert users” (C. Bove, J. Aigrain, M.-J. Lesot, C. Tijus & M. Detyniecki, conférence IUI 2022). Cette vidéo présente un axe de recherche en IA explicable, mettant en évidence notre collaboration avec AXA et l’équipe de C. Tijus en psychologie cognitive.
- ▶ **Élément 3 (projet ou collaboration)** : AXA / Sorbonne Université / LIP6 Trustworthy and Responsible Artificial Intelligence Laboratory (TRAIL). Le AXA/SU/LIP6 TRAIL est un laboratoire de recherche conjoint de Sorbonne Université et d’AXA. Son objectif est d’étudier et de relever les défis liés à l’utilisation généralisée des systèmes d’IA, en particulier pour une utilisation responsable et digne de confiance de l’IA. Il regroupe des chercheurs du LIP6 et de l’équipe de recherche en IA du groupe AXA.
- ▶ **Élément 4 (publication)** : “Fuzzy Approaches for Soft Computing and Approximate Reasoning : Theories and Applications” (M.-J. Lesot, C. Marsala). Ce livre co-édité rassemble 24 articles dans le domaine de l’intelligence computationnelle, présentés par des spécialistes renommés dans ce domaine.

3 AUTOÉVALUATION DU BILAN

3.1 Autoévaluation de l'équipe

Domaine 2. Attractivité

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

Les membres de l'équipe participent activement à l'activité scientifique de leur communauté en étant membres de comités de programme des conférences de leur domaine : pour les conférences internationales majeures on peut citer IJCAI, AAI, Fuzz-IEEE, IPMU, et pour les conférences nationales, la conférence majeure de notre communauté : LFA. D'autre part, tous les membres de l'équipe sont relecteurs pour les grandes revues de notre communauté (*Fuzzy Sets and Systems* (Q1), *International Journal of Approximate Reasoning* (Q1), *Information Sciences* (Q1), *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* (Q3), etc.). Bernadette Bouchon-Meunier est éditrice en chef de la revue *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, et Isabelle Bloch est *area chair* pour *Fuzzy Sets and Systems*, ainsi que pour la revue *Medical Image Analysis* (Q1), Marie-Jeanne Lesot est *area editor* pour *Fuzzy Sets and Systems*.

En 2021, l'équipe LFI a organisé les 30e rencontres francophones sur la logique floue et ses applications (LFA'21), organisée à Paris, dans les locaux de Sorbonne Université. Cette conférence nationale est la conférence francophone majeure de notre communauté. En 2022, elle a organisée la conférence internationale Belief 2022, en collaboration avec l'université Paris Saclay.

La visibilité nationale et internationale de nos domaines de recherche, bien identifiés, nous a permis d'attirer, pour des séjours plus ou moins long, des chercheurs d'universités étrangères (Gabriella Pasi (univ. de Milan), Davide Petturiti (univ. de Perugia), Marek Reformat (univ. d'Edmonton)) ou des collègues d'universités françaises lors de leur délégation CNRS au LIP6 (Cyril de Runz (univ. Reims), Fadi Badra (univ. Paris Nord)).

D'autre part, l'équipe LFI bénéficie d'une visibilité très importante dans la communauté IEEE car Bernadette Bouchon-Meunier a été pendant deux ans présidente de l'IEEE *Computational Intelligence Society* qui recouvre l'ensemble de la communauté IEEE internationale travaillant dans notre domaine (environ 8000 membres).

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

L'équipe LFI est très attachée à développer un esprit de communauté et attentive à maintenir une vie d'équipe incluant tous ses membres : permanents, doctorants ainsi que chercheurs de passage. Pour cela, des réunions scientifiques d'équipe régulières sont organisées (toutes les 3 à 4 semaines), toujours suivies par une rencontre informelle réunissant tous les membres de l'équipe.

L'importance que nous portons à maintenir un lien d'équipe s'est aussi manifestée durant la période du confinement pendant laquelle, outre les réunions scientifiques qui étaient maintenues (à distance pour cause sanitaire), deux points informels hebdomadaires (lundi et vendredi) d'une demi-heure environ avaient été instaurés et qui ont permis de garder un lien entre tous les membres de l'équipe. Ces activités ont eu un impact très important car elles ont permis de garder un contact étroit entre les permanents, les doctorants et les stagiaires et se sont révélées très positives.

D'un point de vue matériel, une place de bureau est réservée pour chaque doctorant, stagiaire, ou contractuel, dans les locaux que nous octroie le LIP6. Cela permet d'assurer une présence régulière (au moins un jour par semaine pour les doctorants en entreprise) dans nos locaux et favorise les interactions avec les autres membres de l'équipe.

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

L'équipe LFI est sollicitée pour répondre à des appels à projets financés et nous participons dans la mesure de nos ressources. Sur la période, divers projets ont pu ainsi être financés, nous pouvons citer des financements ANR (BHA1, projet ANR JCJC, projet ANR JCJC BHA1, projet ANR pHCP, LabCom OT4D), de type FUI BPI (projet

Faucon), ou financé par le ministère de l'enseignement supérieur, de la recherche et de l'innovation (projet de recherche sur les données Parcoursup réalisé avec le LEMMA / Université de Paris Panthéon-Sorbonne).

De plus, une grande partie de nos travaux sont financés par des contrats industriels (contrats directs ou financement de thèses de type CIFRE) essentiellement avec des grands groupes (Airbus, AXA, Thales, Total), mais aussi avec des PME (Ekimetrics, Robeauté). Ce type de collaborations nous offre des moyens de pouvoir travailler sur des problèmes concrets sur lesquels nous pouvons développer nos approches et nos modèles théoriques.

Tous ces financements nous permettent de recruter des doctorants sur des sujets de thèse financés. Il est à noter que les conditions de collaborations et de propriété intellectuelle, bien gérées et négociées en cela par nos contacts de la DRV de Sorbonne Université, nous laissent une entière liberté de direction de nos recherches et de diffusion de nos travaux.

Domaine 3. Production scientifique

LFI, Évolution des publications (2017–2022)

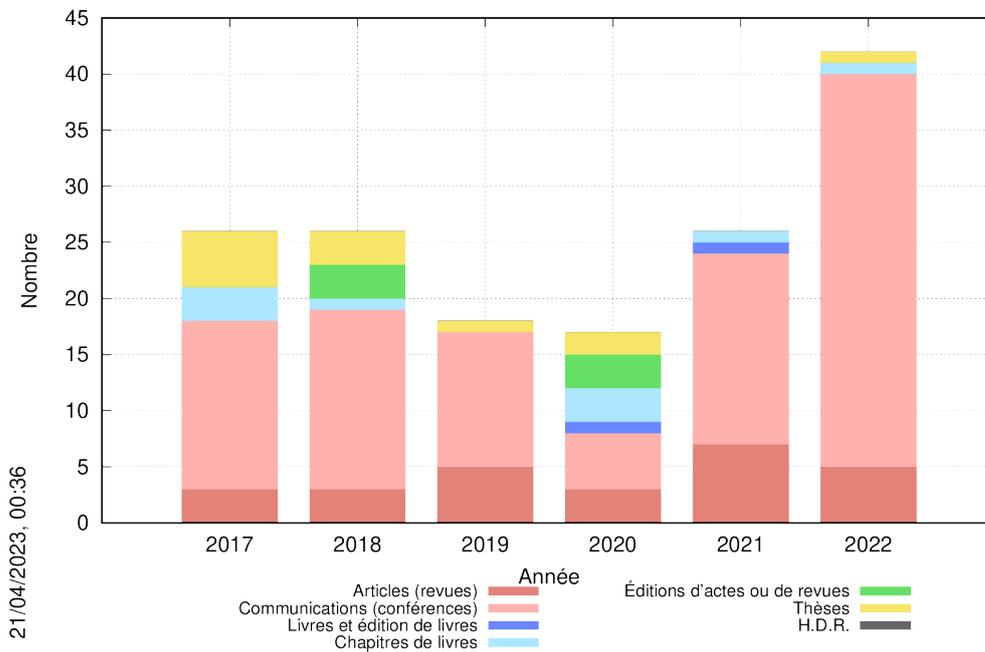


FIGURE 1 – Évolution des publications entre 2017 et 2022

	2017	2018	2019	2020	2021	2022
Articles (revues)	2.00	2.00	3.33	1.50	2.80	2.00
Communications (conférences)	10.00	10.66	8.00	2.50	6.80	14.00

TABLE 2 – Publications par ETPR par an entre 2017 et 2022

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

Dans ses travaux, l'équipe LFI met en avant des approches théoriques et méthodologiques à la pointe de la théorie. Nous publions dans les conférences les plus importantes de notre communauté (IJCAI, Fuzz-IEEE, IPMU). Une reconnaissance de l'importance de nos travaux est les *best paper award* ou *best student paper award* qui ont récompensé des articles de recherche produits par nos chercheurs (Fuzz-IEEE 2018, FQAS 2021, ICPRAI 2021, WBIR 2022) ainsi que la première place à la compétition IEEE CoG 2021.

Outre ces récompenses, on peut aussi mettre en avant les publications réalisées dans les revues internationales de premier plan de notre communauté : *Fuzzy Sets and Systems* (Q1), *International Journal of Approximate Reasoning* (Q1), *IEEE Access* (Q1), *Progress in AI* (Q2), *Soft Computing* (Q2), ainsi que dans les conférences renommées (IJCAI, ICML, Fuzz-IEEE, ECML-PKDD, etc.).

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

Les publications de l'équipe sont réparties sur l'ensemble des personnels. Les doctorants sont encouragés à commencer à publier dès la fin de leur première année afin de préparer au mieux leur comité de suivi de thèse. Par la suite, ils sont incités à rédiger des articles afin de préparer au mieux leur continuation après la thèse. Les doctorants préparent leurs articles, aidés par leur directeur de thèse et leurs encadrants qui participent à l'élaboration de l'article afin de maximiser les chances d'acceptation. En cela, ils peuvent s'appuyer sur l'expérience et l'expertise des membres permanents de l'équipe. Aidés par les directrices et directeurs de thèse, ils sont incités à publier aussi bien dans les grandes conférences internationales que dans les revues du domaine.

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

L'équipe LFI attache une grande importance au suivi de ses publications. Ainsi, dans le respect de notre intégrité scientifique, nous ne publions pas dans des revues pour lesquelles il est nécessaire de financer la publication des articles. De même, dans le respect d'une science ouverte et accessible, nous nous attachons à mettre en ligne (sur le site HAL, ou ArXiv, par exemple) les publications que nous réalisons.

Dans cette optique, nous nous attachons, grâce au support des juristes de la DRV de l'université, à demander systématiquement à nos partenaires industriels la possibilité de publier les travaux que nous réalisons dans le cadre de nos travaux communs.

Domaine 4. Inscription des activités de recherche dans la société

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

Mise en place et création du laboratoire commun AXA/SU TRAIL (*Trustworthy and Responsible Artificial Intelligence Lab*). Le AXA/SU TRAIL a été créé en décembre 2021 et a donné lieu à une convention de partenariat entre le groupe d'assurances AXA et Sorbonne Université sur une durée de 6 ans. Ce laboratoire commun regroupe des chercheurs du groupe AXA et les membres de l'équipe LFI. Il a pour but de développer des recherches communes sur l'Intelligence artificielle explicable et équitable (XAI). Ce partenariat scientifique s'accompagne d'un financement conséquent de la part d'AXA qui permet le recrutement de doctorants pour la réalisation de thèses co-encadrées par les deux parties.

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

En matière de valorisation, un brevet a été déposé avec la société Guerbet. La même personne en prépare un autre sur ses projets avec l'hôpital Necker Enfants Malades (avec lequel une start-up est aussi en cours de montage), en collaboration avec Télécom Paris.

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

Plusieurs actions de médiation scientifique sont menées, en particulier, Isabelle Bloch est intervenue dans le cadre de "Maths en Jeans"¹ et elle a donné plusieurs conférences pour des collégiens et lycéens, et pour des collégiennes et lycéennes.

Nos membres participent aussi à des conférences invitées et des tables rondes en conférences ou à destination de publics non-informaticiens (par exemple, Journées Francophones de Radiologie, Cité des Sciences).

1. <https://www.mathenjeans.fr/>

4 RÉFÉRENCES BIBLIOGRAPHIQUES SIGNIFICATIVES DE LFI

- [Aiguier and Bloch, 2022] Aiguier, M. and Bloch, I. (2022). Abstract categorical logic. *Logica Universalis*.
- [Badra et al., 2022] Badra, F., Lesot, M.-J., Barakat, A., and Marsala, C. (2022). Theoretical and Experimental Study of a Complexity Measure for Analogical Transfer. In *Int. Conf. on Case-based Reasoning (ICCBR)*, volume 13405 of *Lect. Notes in Comp. Science*, pages 175–189, Nancy, France. Springer.
- [Bloch, 2021] Bloch, I. (2021). Modeling imprecise and bipolar algebraic and topological relations using morphological dilations. *Mathematical Morphology - Theory and Applications*, 5(1) :1–20.
- [Bloch and Lesot, 2022] Bloch, I. and Lesot, M.-J. (2022). Towards a formulation of fuzzy contrastive explanations. In *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–8.
- [Bloch and Ralescu, 2023] Bloch, I. and Ralescu, A. (2023). *Fuzzy Sets Methods in Image Processing and Understanding : Medical Imaging Applications*. Springer Nature.
- [Bouchon-Meunier et al., 2022] Bouchon-Meunier, B., Laurent, A., and Lesot, M.-J. (2022). XAI : a Natural Application Domain for Fuzzy Set Theory. In Smith, A. E., editor, *Women in Computational Intelligence, Key Advances and Perspectives on Emerging Topics*, Women in Engineering and Science (WES), pages 23–49. Springer.
- [Bouchon-Meunier and Marsala, 2020] Bouchon-Meunier, B. and Marsala, C. (2020). Entropy and monotonicity in artificial intelligence. *International Journal of Approximate Reasoning*, 124 :111–122.
- [Bove et al., 2022] Bove, C., Aigrain, J., Lesot, M.-J., Tijus, C., and Detyniecki, M. (2022). Contextualization and exploration of local feature importance as explanations to improve understanding and satisfaction of non-expert users. In *Proc. of the Int. Conf. on Intelligent User Interfaces IUI'22*.
- [Chuquimia et al., 2018] Chuquimia, O., Pinna, A., Marsala, C., Dray, X., and Granado, B. (2018). FPGA-based Real Time Embedded Hough Transform Architecture for circles detection. In *DASIP 2018 - Conference on Design and Architectures for Signal and Image Processing*, Porto, Portugal.
- [Coletti and Bouchon-Meunier, 2019] Coletti, G. and Bouchon-Meunier, B. (2019). Fuzzy Similarity Measures and Measurement Theory. In *IEEE International Conference on Fuzzy Systems 2019 (FUZZ-IEEE 2019)*, New Orleans, United States.
- [Franchi et al., 2022] Franchi, G., Bursuc, A., Aldea, E., Dubuisson, S., and Bloch, I. (2022). One Versus all for deep Neural Network Incertitude (OVNNI) quantification. *IEEE Access*, 10 :7300–7312.
- [Israël-Jost et al., 2021] Israël-Jost, V., Weil-Dubuc, P., Adamsbaum, C., Bloch, I., et al. (2021). L'éthique en radiologie : quand, comment ? premiers éléments. *Journal d'Imagerie Diagnostique et Interventionnelle*, 4 :238–240.
- [Janiak et al., 2022] Janiak, V., Jousselin, V., Tallec, G., Marsala, C., Saleem, U., Dres, M., and Pinna, A. (2022). Prediction of the extubation outcome through Electrical Impedance Tomography measurements. In *IEEE Bio-medical Circuits and Systems (BIOCAS)*, Tapei, Taiwan.
- [Laugel et al., 2019] Laugel, T., Lesot, M.-J., Marsala, C., Renard, X., and Detyniecki, M. (2019). The dangers of post-hoc interpretability : unjustified counterfactual explanations. In *Proc. of the 28th Int. Joint Conf. on Artificial Intelligence, IJCAI*, pages 2801–2807.
- [Lefort et al., 2017] Lefort, S., Lesot, M.-J., Zibetti, E., Tijus, C., and Detyniecki, M. (2017). Interpretation of approximate numerical expressions : Computational model and empirical study. *Int. J. of Approximate Reasoning IJAR*, 82 :193–209.
- [Legastelois et al., 2017] Legastelois, B., Lesot, M.-J., and d'Allonnes, A. R. (2017). Typology of axioms for a weighted modal logic. *Int. J. of Approximate Reasoning IJAR*, 90 :341–358.
- [Marsala et al., 2022] Marsala, C., Bloch, I., Lesot, M.-J., Tollari, S., and Vittaut, J.-N. (2022). Recherches en IA explicable dans l'équipe LFI du LIP6. *Bulletin de l'AFIA*, (116) :9–13.
- [Marsala and Bouchon-Meunier, 2022] Marsala, C. and Bouchon-Meunier, B. (2022). Explainable Fuzzy Interpretative Reasoning. In *IEEE World Congress on Comput. Intelligence*, Padova, Italy.
- [Martin et al., 2022] Martin, G., El-Madafri, S., Becq, A., Szweczyk, J., and Bloch, I. (2022). Instruments Segmentation in X-ray Fluoroscopic Images for Endoscopic Retrograde Cholangio Pancreatography. *Studies in Health Technology and Informatics*, 294 :133–134.
- [Munro et al., 2022] Munro, Y., Bloch, I., Chetouani, M., Lesot, M.-J., and Pelachaud, C. (2022). Argumentation and Causal Models in Human-Machine Interaction : A Round Trip. In *8th International Workshop on Artificial Intelligence and Cognition*, Örebro, Sweden.

A ANNEXE — MEMBRES PERMANENTS AU 31/12/2022

La table ci dessous liste les membres permanents de l'équipe LFI.

NOM	Prénom	Corps	Employeur
BLOCH	Isabelle	PR	Sorbonne Université
LESOT	Marie-Jeanne	MCF (HDR)	Sorbonne Université
MARSALA	Christophe	PR	Sorbonne Université
TOLLARI	Sabrina	MCF	Sorbonne Université
VITTAUT	Jean-Noël	MCF	Sorbonne Université

ÉLÉMENT DE PORTFOLIO 01



Publication

1 DÉFINITION DE CET ÉLÉMENT

Titre de l'élément : Recherches en IA explicable dans l'équipe LFI du LIP6

URL de l'élément : <https://hal.science/hal-04013819>

Fichier de élément : 2022-AFIA-Bulletin-116_avr22-LFI.pdf

2 MOTIVATIONS DU CHOIX DE CET ÉLÉMENT

Cet article a été rédigé par les membres permanents de l'équipe afin de mettre en avant nos recherches en intelligence artificielle explicable. Il offre ainsi une vision transverse de nos travaux dans ce domaine. Il a été publié dans le bulletin de l'Association Française en Intelligence Artificielle (AFIA) qui est l'organisme fédérateur des recherches en IA en France.

3 PRÉSENTATION DE CET ÉLÉMENT

L'équipe Learning Fuzzy and Intelligent systems (LFI) du laboratoire LIP6 de Sorbonne Université développe des recherches en interprétabilité des méthodes d'intelligence artificielle dans les domaines de l'aide à la décision, la science des données et l'apprentissage automatique. Les objectifs scientifiques et applicatifs sont de concevoir et de proposer des approches explicables à la fois durant leur construction et lors de leur utilisation. Cet article présente brièvement ces contributions développées dans le cadre de multiples collaborations, en évoquant les tâches d'apprentissage automatique, dans des approches by design (dès la conception) ou post-hoc, la caractérisation de données par résumés linguistiques, l'interprétation d'images ou la formulation d'explications dans un cadre logique.

4 RÉFÉRENCES BIBLIOGRAPHIQUES

[1] Christophe Marsala, Isabelle Bloch, Marie-Jeanne Lesot, Sabrina Tollari, and Jean-Noël Vittaut. Recherches en IA explicable dans l'équipe LFI du LIP6. *Bulletin de l'AFIA*, (116) :9–13, April 2022.

ÉLÉMENT DE PORTFOLIO 02



Vidéo

1 DÉFINITION DE CET ÉLÉMENT

Titre de l'élément : Contextualization and exploration of local feature importance as explanations to improve understanding and satisfaction of non-expert users

URL de l'élément : <https://programs.sigchi.org/iui/2022/program/content/79950>

2 MOTIVATIONS DU CHOIX DE CET ÉLÉMENT

Cette vidéo présente un axe de nos recherches en IA explicable, plus précisément sur les *eXplanation User Interfaces*, dont l'objectif est l'amélioration de la présentation des explications fournies par des méthodes existantes, en les rendant plus adaptées aux utilisateurs. Dans ce travail, caractéristique des travaux que nous menons dans ce domaine, l'interface propose d'intégrer des outils de contextualisation et d'exploration, notamment à destination d'utilisateurs non experts, qui ne possèdent pas de connaissances spécifiques, ni en IA ni sur le domaine considéré. Ces travaux sont représentatifs des travaux de notre équipe, liant travail théorique et applications, et dont l'évaluation peut être faite par une expérimentation menée avec un panel d'utilisateurs. De plus, le travail réalisé met en évidence le succès de notre collaboration avec l'équipe de recherche d'AXA, ainsi qu'avec les membres de l'équipe de psychologie cognitive de C. Tijus à Paris 8.

3 PRÉSENTATION DE CET ÉLÉMENT

Cet article propose des principes génériques de design d'interfaces XAI pour contextualiser et permettre l'exploration d'explications exprimées sous la forme de vecteurs d'importance locale, afin d'améliorer l'accessibilité et l'utilité de ces explications. Pour évaluer la pertinence de ces principes génériques, nous les avons implémentés dans le contexte d'un scénario lié au choix d'une assurance et nous avons conduit une expérience utilisateurs avec 80 participants pour comparer 4 versions différentes de l'interface, selon l'intégration ou non de chacun des deux principes proposés. L'évaluation porte à la fois sur la compréhension objective et sur la satisfaction subjective des participants. Les résultats montrent que les principes d'exploration proposés améliorent de façon significative la satisfaction subjective et atteignent presque le seuil de signification pour la compréhension objective. Ils montrent également que les principes d'exploration permettent d'améliorer la satisfaction des utilisateurs.

4 RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] C. Bove, J. Aigrain, M.-J. Lesot, C. Tijus, and M. Detyniecki. Contextualization and exploration of local feature importance as explanations to improve understanding and satisfaction of non-expert users. In *Proc. of the Int. Conf. on Intelligent User Interfaces IUI'22, 2022*.

ÉLÉMENT DE PORTFOLIO 03



Projet ou collaboration

1 DÉFINITION DE CET ÉLÉMENT

Titre de l'élément : Laboratoire commun Sorbonne Université/LIP6/AXA Trustworthy and Responsible Artificial Intelligence Lab (TRAIL).

URL de l'élément : <https://trail.lip6.fr/>

2 MOTIVATIONS DU CHOIX DE CET ÉLÉMENT

Cet élément met en évidence une collaboration industrielle très importante pour l'équipe LFI, avec un partenaire industriel, le groupe AXA et son équipe de recherche en IA dirigée par Marcin Detyniecki. Cette collaboration, entamée en 2015 par un dépôt de projet auprès du Fonds AXA pour la recherche, a permis le co-encadrement de thèses, soutenues financièrement par AXA, et a offert des opportunités de recrutement pour les docteurs de l'équipe. En décembre 2021, pour pérenniser notre collaboration et la rendre plus formelle, le laboratoire commun Sorbonne Université/LIP6/AXA *Trustworthy and Responsible Artificial Intelligence Lab* (TRAIL) a été créé, pour une durée initiale de cinq ans, avec une dotation d'AXA pour le financement de thèses et l'embauche de post-doctorants.

3 PRÉSENTATION DE CET ÉLÉMENT

Le TRAIL est un laboratoire de recherche commun à Sorbonne Université, au LIP6 et à AXA. Son objectif est d'étudier et de relever les défis liés à l'utilisation des systèmes d'IA, en particulier pour une utilisation responsable et digne de confiance. TRAIL regroupe des chercheurs du LIP6 et des chercheurs en IA d'AXA. La création de ce laboratoire a été l'aboutissement de plusieurs années de collaboration entre le LIP6 et AXA dans le domaine de l'IA explicable.

Les thèmes de recherche de TRAIL couvrent l'évaluation de l'équité des systèmes d'IA et l'atténuation des biais, l'explicabilité et l'interprétabilité de ces systèmes (*eXplainable Artificial Intelligence – XAI*), leur robustesse et leur sécurité, ainsi que l'interface homme-système d'IA. TRAIL étudie également la gouvernance de l'IA dans les organisations.

D'un point de vue scientifique, TRAIL vise à fournir de nouvelles approches explicables et interprétables pour un déploiement fiable et responsable de l'IA. Plus concrètement, les travaux communs LIP6/AXA se font par le co-encadrement de thèses dans le domaine de l'IA explicable (XAI), financées par une dotation d'AXA. A terme, nous envisageons de déposer des demandes de projets conjointement dans ce domaine.

Il est à noter que TRAIL a une approche ouverte en diffusant tous les travaux qui y sont réalisés : la publication de nos recherches, la promotion des connaissances dans notre domaine et la mise à disposition des méthodes et outils développés. également la gouvernance de l'IA dans les organisations.

4 RÉFÉRENCES BIBLIOGRAPHIQUES

Une liste des publications réalisées par des membres de TRAIL est disponible sur le site web :

<https://trail.lip6.fr/publications/>

ÉLÉMENT DE PORTFOLIO 04



Publication

1 DÉFINITION DE CET ÉLÉMENT

Titre de l'élément : Fuzzy Approaches for Soft Computing and Approximate Reasoning : Theories and Applications

URL de l'élément : <https://link.springer.com/book/10.1007/978-3-030-54341-9>

2 MOTIVATIONS DU CHOIX DE CET ÉLÉMENT

Cet ouvrage co-édité par deux membres de l'équipe LFI met en évidence les liens de notre équipe avec notre communauté scientifique, tant au niveau national qu'au niveau international. De nombreux collègues avec qui nous collaborons par ailleurs, dans l'organisation de conférences, dans la gestion de notre communauté, que nous rencontrons lors de nos conférences et avec qui nous avons tissé des liens scientifiques, même si ceux-ci ne donnent pas lieu à des publications, ont répondu à l'appel à participation que nous avons lancé pour cet ouvrage.

3 PRÉSENTATION DE CET ÉLÉMENT

Ce livre, co-édité par deux membres de l'équipe LFI, rassemble 24 articles dans le domaine de l'intelligence computationnelle, présentés par des spécialistes renommés dans ce domaine. Il offre une large palette de travaux couvrant les grandes tendances de notre communauté scientifique et présente une vue d'ensemble des différents sujets liés à *soft computing* et des approches basées sur le raisonnement approximatif, à la fois d'un point de vue théorique et d'un point de vue appliqué.

De nombreux sujets sont couverts : les aspects fondamentaux de la théorie des ensembles flous, les approches de raisonnement (interpolatif, analogique, basé sur la similarité), la théorie de la décision et de l'optimisation, les bases de données floues, l'apprentissage automatique, la construction de résumés linguistiques, l'interprétabilité et l'IA explicable (XAI). Plusieurs articles décrivent des applications, par exemple, pour le traitement d'images, le web sémantique et les systèmes de tutorat intelligents.

Ce livre a été pensé par les éditeurs comme un moyen d'offrir à des lecteurs, non nécessairement spécialistes de ce domaine, une introduction et un éventail de travaux actuels. De plus, les éditeurs et tous les auteurs participants à ce livre l'ont dédié à Bernadette Bouchon-Meunier en remerciement de tous ses apports dans le domaine de l'intelligence computationnelle, tant du point de vue de ses travaux scientifiques, que de son investissement pour la communauté internationale.

4 RÉFÉRENCES BIBLIOGRAPHIQUES

[1] Marie-Jeanne Lesot and Christophe Marsala, editors. *Fuzzy Approaches for Soft Computing and Approximate Reasoning : Theories and Applications*. Number 394 in Studies in Fuzziness and Soft Computing. Springer Cham, 2021.