



Forum JCJC
Jeunes Chercheuses Jeunes Chercheurs
Actes de la 14^{ème} édition

INFORSID 2026

Toulouse, 27 mai 2026

Organisé par
Manuele Kirsch Pinheiro

Liste d'articles

Préface	i
Manuele Kirsch Pinheiro	
A dual-level framework for promoting lesser-known Points of Interest in mountain regions	
Adriana Pompejano	1
Vers une réponse de cybersécurité explicable par la logique probabiliste décisionnelle	
Zequan Huang	5
Détection d'anomalies basée sur les ontologies et l'apprentissage automatique	
Noa Barbosa	9
Semantic and Multimodal Structuring of Historical Archives: Towards New Information Systems	
Mohamed Amine Lasheb	13
Optimization of AI Agent Systems via Process Mining : Modeling, Analysis and Improving	
Clément Adandé	17
Un Cadre Décisionnel Basé sur Apprentissage Automatique pour la Santé et la Sécurité au Travail dans le cadre du Transport Intelligent	
Cynthia Chedrawi	21
Collaboration Dans Les Méthodes Agiles: Comment l'IA impacte le processus agile ?	
Haoyue Liu	25
Approche multimodale pour la détection précoce des troubles cognitifs par apprentissage automatique	
Mohamed Hedi Djemaa	29
Un framework performant et explicable pour la prévision des ventes de tickets au cinéma	
Lorenzo Lefevre	33
From Digital Traces to Shared Understanding: An AI-based approach to group awareness in distributed teams	
Kaoutar Akhsass	37
Intelligence Artificielle Explicable pour la détection de cyberattaques et la reconnaissance de plans via des règles probabilistes entraînées de manière adverse	
Antoine Leblanc	41
Détection et qualification des dégradations émotionnelles et relationnelles au sein d'un groupe familial dans un contexte d'aidance	
Paul Cariou	45

Préface

Le forum Jeunes Chercheuses Jeunes Chercheurs (JCJC) est un événement traditionnel d'INFORSID. Désormais annuel, le forum JCJC permet aux étudiants en 1^{ère} ou 2^{ème} année de thèse de présenter leurs travaux à la communauté française en Système d'Information et d'obtenir ainsi de valeureux retours à propos de leurs recherches.

L'objectif du forum JCJC n'est nullement d'être sélectif ou compétitif. L'objectif ici est d'accompagner les jeunes chercheuses et chercheurs dans leur premiers pas au sein de la communauté. Il s'agit d'un moment particulièrement important, à la fois pour les doctorants, qui découvrent pour beaucoup la communauté Système d'Information, mais aussi pour la communauté elle-même, qui a l'opportunité d'accueillir des nouveaux talents sur lesquels repose le futur de la communauté.

L'absence de sélectivité ne veut pas dire absence de qualité. Chaque article a été évalué par trois évaluateurs distincts, qui ont donné leur avis et proposé des axes d'améliorations pour les articles et les travaux de thèse de manière générale. Il en résulte un panel d'articles particulièrement intéressants, qui démontre à lui seul le dynamisme et la variété de la recherche en Système d'Information aujourd'hui.

Par ailleurs, il est important de souligner que, tradition l'oblige, les présentations des articles au Forum JCJC ont suivi le format « Dragons & Chevalier ». Chaque doctorante et doctorant a eu 3 min pour présenter son propre article, mais elle/il a également dû endosser, tour à tour, le rôle de dragon pour un collègue, en critiquant, de manière constructive, le travail du collègue, et rôle de chevalier d'un autre collègue, en mettant en valeur les points forts du travail de ce collègue. Cet exercice, loin d'être anodin, permet aux doctorants de favoriser le développement de l'esprit critique, mais également le contact et les échanges entre doctorants. En outre, une fois la séance de présentations réalisée, les articles ont été affichés sous la forme de poster A3, sur lesquels le public a pu venir coller des « post-its » avec des commentaires et des questions, permettant d'accroître les échanges entre jeunes chercheuses / jeunes chercheurs et la communauté INFORSID. Il en résulte de tout ça, une séance riche et passionnante, permettant des nombreux échanges entre jeunes et moins jeunes chercheuses/chercheurs.

Tout ceci n'aurait pas été possible sans la collaboration de plusieurs personnes, à commencer par les jeunes eux-mêmes et leurs encadrants qui ont accepté le jeu.

Je souhaite également remercier Luiz Angelo Steffemel, mon acolyte de toujours, qui a largement contribué à la production des actes.

Je souhaite remercier chaleureusement les membres du comité de relecture, qui ont accepté de réviser les articles. J'aimerais les remercier par le temps qu'ils ont consacré à nos jeunes chercheuses et jeunes chercheurs et aux commentaires fournis, toujours très constructifs qui vont, sans nul doute, leur aider à progresser dans leur thèse.

- Lylia Abrouk, Université Bourgogne Europe
- Pierre-Emmanuel Arduin, Université Paris-Dauphine
- Nourhène Ben Rabah, Université Paris 1 Panthéon Sorbonne
- Khalid Benali, Université de Lorraine
- Hugo Boisaubert, Université de Toulouse - IRIT
- François Charoy, Université de Lorraine
- Yohann Chasseray, Centre Génie Industriel - IMT Mines Albi-Carmaux
- Cécile Favre, Université Lyon 2
- Alexis Guyot, Aix Marseille Université
- Lydia Khelifa, EFREI / Université Paris-Panthéon-Assas
- Elena Kornyshova, CNAM
- Sébastien Laborie, Université de Pau et des Pays de l'Adour
- Florent Nolot, University of Reims Champagne-Ardenne
- Claudia Roncancio, Grenoble University
- Luiz Angelo Steffenel, University of Reims Champagne-Ardenne
- Cyrille Suire, Université de La Rochelle

Enfin, j'aimerais tout particulièrement remercier le bureau d'INFORSID de m'avoir fait confiance pour organiser ce Forum.

Manuele Kirsch Pinheiro

Manuele.Kirsch-Pinheiro@univ-paris1.fr

Centre de Recherche en Informatique

Université Paris 1 Panthéon Sorbonne

Responsable du Forum JCJC 2026

A dual-level framework for promoting lesser-known Points of Interest in mountain regions

Adriana Pompejano*

Encadrants : Sébastien Laborie*, Maxime Masson*, Philippe Roose*, Christian Sallaberry, Bianca Tescaşiu*****

* E2S UPPA, LIUPPA, Université de Pau et des Pays de l'Adour
1 Allée du Parc Montaury, 64600 Anglet, France

** E2S UPPA, LIUPPA, Université de Pau et des Pays de l'Adour
Avenue de l'Université, 64000 Pau, France

*** Department of Marketing, Tourism and International Business, Transilvania University of Braşov

29 Eroilor Boulevard, 500036, Braşov, Romania

*,** *firstname.lastname@univ-pau.fr*, *** *bianca.tescasiu@unitbv.ro*

ABSTRACT: Overtourism increasingly threatens mountain regions, damaging ecosystems and compromising residents' quality of life. This phenomenon concentrates visitors at a few overcrowded Points of Interest (POIs) while leaving other sites, referred to as lesser-known POIs, undervalued. To address this, our study proposes a dual-level framework structured across: a macro-level analysis of the broader territory, and a micro-level view of specific POIs. This enables the development of indicators (e.g., carrying capacity and lesser-known scores) aiming to balance the promotion of undervalued sites with the preservation of their sustainability. The integration of these metrics supports a more responsible redistribution of tourist flows, fostering a sustainable spatial distribution of visitors throughout the territory.

KEYWORDS: Overtourism, Lesser-known POIs, Heterogeneous Data Integration, Multi-scale analysis, Recommender systems, Sustainable Tourism.

1. Introduction

In recent decades, overtourism has emerged as a critical challenge where visitor numbers exceed the carrying capacity of a destination or specific POIs (World Tourism Organization, 2018), causing environmental degradation, diminishing tourist experiences and tensions with local communities.

A strategy to mitigate overtourism is redistributing tourist flows from overcrowded sites to lesser-known POIs (Masson and Mine, 2025). However, in mountain regions, identifying and integrating them into territorial strategies remains challenging due to limited data, environmental complexity and considering the influence of tourist behavior. For this reason, our research investigates how to design a recommendation-oriented framework to identify lesser-known POIs while preventing the replication of overtourism. To address this challenge, we propose a holistic and generalizable framework that (1) aggregates and analyzes mountain POIs from heterogeneous data sources, (2) integrates data-driven techniques with responsible territorial marketing strategies, and (3) applies a macro–micro perspective to develop scoring systems considering regional and POI-specific needs. Our framework lays technical foundations for a future Recommender System to support Destination Management Organisations (DMOs) in identifying and promoting lesser-known POIs while respecting the carrying capacity.

2. State of the art

Overtourism is identified as a failure to manage a destination’s carrying capacity (Capocchi *et al.*, 2019), particularly in vulnerable mountain regions (Fidelus-Orzechowska *et al.*, 2021; Senetra *et al.*, 2020) where there is a need to balance ecological conservation and cultural identity (Baloch *et al.*, 2023; Senetra *et al.*, 2020). In response, slow tourism has emerged as a strategic alternative, emphasizing local engagement and sustainable mobility (Klarin *et al.*, 2023; Paul, 2014). Mountain regions provide an ideal setting for this approach, offering diverse slow-paced options, such as cycling or trekking.

Several studies have explored Recommender Systems in redirecting tourist flows, focusing on uncovering hidden gems (Buranasing *et al.*, 2021) by connecting tourists with local guides (Pradipta *et al.*, 2021). While these systems have been a core subject of computer science, this study focuses specifically on their application within destination management. Recent research highlights the identification of lesser-known POIs as a strategic alternative to overcrowded sites (Masson and Mine, 2025). In mountain environments, dominant applications such as AllTrails¹ or Komoot² rely on popularity-based algorithms that often reinforce visitation to already well-known sites. Consequently, territorial constraints are often overlooked, including site accessibility,

1. <https://www.alltrails.com>

2. <https://www.komoot.com>

infrastructure capacity, and local pressures, affecting both environmental sustainability and the quality of life of local communities.

3. Contributions for improving territorial management

To overcome the limitations of current approaches, our framework combines computational methods with responsible marketing strategies (e.g., slow tourism). It adopts both macro-level and micro-level perspectives to promote lesser-known POIs while ensuring that the redistribution of tourist flows respects the destination's carrying capacity and the needs of local communities. Our contribution addresses four main challenges: (I) collecting heterogeneous data from limited sources, (II) performing multi-level analysis to discover lesser-known POIs, (III) supporting sustainable marketing strategies, (IV) and enabling the recommendation of lesser-known POIs in mountain regions.

Our research method follows a pipeline consisting of four steps summarized below. A more detailed version, including an illustrative figure, is available online³.

1) **POIs Collection:** we aggregate heterogeneous data to construct a POIs knowledge base. This includes Geographic data (from OpenStreetMap sources), Digital Insights (from WikiData and Pageviews) and Local Stakeholder Insights (through semi-structured interviews).

2) **POIs Analysis:** combining Deep Learning analysis and selected indicators (such as online presence, data scarcity etc., chosen according to professionals' needs) we identify lesser-known POIs to create an enriched dataset.

3) **Sustainable Marketing Strategies:** we develop scoring systems to prioritize lesser-known POIs, balancing new tourism opportunities with carrying capacity and ensuring that promotional strategies respect the territory's environmental and social limits.

4) **POIs Application:** we integrate these scores into a Decision Support System to provide recommendations for DMOs to redirect tourist flows more responsibly.

This four-phase pipeline is organized around a dual-level analysis⁴: a **macro-level** perspective, examining territorial information, and a **micro-level** perspective, focusing on POI-specific data within the same area. Both levels are analyzed across four dimensions, illustrated by the comparison between a municipality (macro-level) and a restaurant (micro-level): **General Information** (e.g., population density and surface area vs. restaurant GPS coordinates, average prices, and opening year); **Infrastructures and Accessibility** (e.g., city restaurant count vs. specific opening hours); **Tourism Demand** (e.g., city total visitor arrivals vs. daily restaurant clients); and **Digital Marketing** (e.g., municipal online reputation vs. individual restaurant reviews). These dimensions provide a comprehensive view of the destination and facilitate the development of scoring systems to respect the environmental and social integrity of localities.

3. <https://mmasson003.perso.univ-pau.fr/images/pipeline.jpg>

4. <https://mmasson003.perso.univ-pau.fr/images/views.png>

4. Future actions

Future efforts focus on three directions: (1) developing a prototype to demonstrate the integration of diverse micro and macro-level data within our four dimensions; (2) automating the score-calculation such as carrying capacity and lesser-known POIs; and (3) designing a recommendation strategy that balances territorial sustainability with new tourism opportunities in lesser-known POIs. This stage will involve collaboration with local stakeholders to validate our results. Ultimately, we aim to ensure our approach is generic enough to be applied across diverse mountainous contexts, such as the Pyrenees, Alps, and Carpathians, each offering unique challenges in terms of tourism maturity and environmental protection.

References

- Baloch Q. B., Shah S. N., Iqbal N., Sheeraz M., Asadullah M., Mahar S., Khan A. U., “Impact of tourism development upon environmental sustainability: a suggested framework for sustainable ecotourism”, *Environmental Science and Pollution Research*, vol. 30, n^o 3, p. 5917-5930, 2023.
- Buranasing W., Meeklai P., Pattarathananan P., “Recommendation system for lesser-known places to visit in Thailand”, *Proceedings of the 2021 3rd ICSED*, p. 24-28, 2021.
- Capocchi A., Vallone C., Pierotti M., Amaduzzi A., “Overtourism: A literature review to assess implications and future perspectives”, *Sustainability*, vol. 11, n^o 12, p. 3303, 2019.
- Fidelus-Orzechowska J., Gorczyca E., Bukowski M., Krzemień K., “Degradation of a protected mountain area by tourist traffic: case study of the Tatra National Park, Poland”, *Journal of Mountain Science*, vol. 18, n^o 10, p. 2503-2519, 2021.
- Klarin A., Park E., Xiao Q., Kim S., “Time to transform the way we travel?: A conceptual framework for slow tourism and travel research”, *Tourism Management Perspectives*, vol. 46, p. 101100, 2023.
- Masson M., Mine T., “Explainable Recommender System to Promote Lesser-Known POIs for Sustainable Tourism and Small Municipality Revitalization”, *Opening Ceremony of the 2025 Summer Program of the Japan Society for the Promotion of Science*, 2025.
- Paul B. D., “From slow food to slow tourism”, *Annals of the university of oradea, Economic Science Series*, vol. 23, n^o 2, p. 137-144, 2014.
- Pradipta I. A., Anggraini R. I. *et al.*, “Journey: digital tourism platform for hidden gems destination”, *Proc. BIEC 2020*, Atlantis Press, p. 125-131, 2021.
- Senetra A., Dynowski P., Cieślak I., Źróbek-Sokolnik A., “An evaluation of the impact of hiking tourism on the ecological status of alpine lakes—a case study of the valley of dolina Pięciu Stawów polskich in the Tatra Mountains”, *Sustainability*, vol. 12, n^o 7, p. 2963, 2020.
- World Tourism Organization, “‘Overtourism’? – Understanding and Managing Urban Tourism Growth beyond Perceptions, Executive Summary”, 2018.

Vers une réponse de cybersécurité explicable par la logique probabiliste décisionnelle

Zequan Huang* **

Encadrants : Bénédicte Le Grand*, **Jacques Robin****, **Nourhène Ben Rabah***

* *Centre de Recherche en Informatique (CRI), Université Paris 1 Panthéon-Sorbonne
90 rue de Tolbiac,
75013 Paris*

** *Learning Data & Robotics (LDR), esieaLab, ESIEA
4 All. Katherine Johnson,
94200 Ivry-sur-Seine*

* {zequan.huang, nourhene.ben-rabah, benedicte.le-grand}@univ-paris1.fr

** jacques.robin@esiea.fr

RESUME. Ce papier présente l'état d'avancement de ma première année de thèse autour de la réponse aux cyberattaques dans Centres Opérationnels de Sécurité. Mon objectif est de concevoir un système de réponse automatique à un incident de cybersécurité d'une manière explicable pour l'analyste et l'opérationnel en infrastructure réelle. L'approche proposée met en œuvre un socle ontologique, un moteur d'inférence en logique probabiliste décisionnelle et son adaptation adverse.

ABSTRACT. This paper presents the current progress of the first year of my PhD on cyberattack response in Security Operations Centers. My objective is to design an autonomic cybersecurity incident response system that remains explainable for analysts and operational in real infrastructures. The approach combines an ontological foundation, an inference engine based on decision-theoretic probabilistic logic, and its adaptation to adverse attacker behaviors.

MOTS-CLÉS : Réponse à incident de cybersécurité, programmation logique probabiliste décisionnelle, IA explicable, raisonnement adverse

KEYWORDS: Cybersecurity incident response, decision-theoretic probabilistic logic, explainable AI, reasoning under adverse behavior

1. Introduction

Les équipes des **Centres Opérationnels de Sécurité (SOC)** traitent aujourd’hui des volumes d’alertes élevés dans des environnements techniques hétérogènes (Vielberth *et al.*, 2020). Malgré les progrès de l’automatisation, la réponse à un incident de cybersécurité demeure largement humaine, avec une surcharge cognitive qui peut dégrader la rapidité et la qualité des décisions (Zidan *et al.*, 2024). Dans le cadre de ma thèse, menée dans le projet franco-luxembourgeois ANCILE, je m’intéresse à la réponse aux **attaques persistantes avancées (APT)**, avec un objectif clair : produire des recommandations à la fois efficaces, explicables et directement exploitables par les analystes.

2. Positionnement scientifique

2.1. *Limites des approches de réponse existantes*

Les travaux de réponse à des cyberattaques reposent classiquement sur la modélisation de menace, l’évaluation de risque et la sélection de contre-mesures (Nespoli *et al.*, 2018). Les modèles les plus utilisés sont les arbres d’attaque (Schneier, 1999) et les graphes d’attaque (Ou *et al.*, 2006). Ces approches ont permis des progrès notables, mais leur transfert vers un SOC opérationnel reste délicat.

Deux limites majeures persistent :

– **Faible interopérabilité sémantique.** Les concepts de menace, d’actif, de vulnérabilité et de contre-mesure sont souvent définis de manière locale au modèle, ce qui complique la réutilisation et la validation inter-outils.

– **Compromis explicabilité / passage à l’échelle.** Les méthodes performantes en contexte stochastique et adverse sont souvent difficiles à interpréter, alors que les SOC exigent des justifications traçables pour la prise de décision.

Mon hypothèse est qu’une meilleure robustesse opérationnelle est atteignable en combinant 3 axes : 1) un socle ontologique, 2) un moteur d’inférence déclaratif probabiliste orienté décision et 3) une adaptation continue face à des attaquants dynamiques.

3. Approche proposée

3.1. *Socle de connaissances unifié*

Le premier axe structure une ontologie de cybersécurité orientée réponse, à partir de travaux en ontologie de risque et de sécurité (Sales *et al.*, 2018; Oliveira *et al.*, 2022). Le moteur d’inférence manipule ainsi des concepts normalisés (actifs, compromission, objectifs, capacités défensives, coût/impact). Les contre-mesures sont exprimées d’abord comme intentions, puis raffinées en actions techniques selon le contexte d’exécution.

3.2. *Inférence déclarative décisionnelle*

Le deuxième axe développe un moteur d'inférence en logique probabiliste décisionnelle, inspiré de DTProbLog (Van den Broeck *et al.*, 2010). Ce moteur prend en entrée les alertes corrélées, l'état du système décrit par l'ontologie et un ensemble d'actions de réponse candidates. Il évalue ensuite l'utilité attendue de ces actions sous incertitude, en tenant compte d'un compromis multicritère (réduction du risque, impact métier, coût opérationnel).

L'intérêt de cette formalisation est double : (i) produire une décision calculée de manière explicite et (ii) fournir une justification directement exploitable par l'analyste SOC, sous forme de règles activées, d'hypothèses retenues et de critères ayant pesé dans le choix final.

3.3. *Raisonnement adverse*

Le troisième axe fait évoluer dynamiquement le moteur défini (3.2) face aux changements de tactiques adverses. Je retiens pour cela un cadre de jeu stochastique partiellement observable (Zonouz *et al.*, 2014; Hammar *et al.*, 2025), dans lequel le moteur d'inférence initial sert de politique défensive de départ.

Concrètement, l'adaptation porte sur les paramètres et priorités du moteur d'inférence (par exemple les poids de certaines règles probabilistes), à partir des retours d'exécution et des observations en cours d'attaque. L'enjeu est d'améliorer la performance défensive sans perdre l'explicabilité : les mises à jour doivent rester traçables et compatibles avec la structure déclarative construite en 3.2.

4. **Avancement et perspectives**

Les travaux menés durant ma première année de thèse ont abouti à deux papiers acceptés à la conférence EDOC 2025 (Huang *et al.*, 2026; Leblanc *et al.*, 2026), ce qui consolide la base scientifique du projet. Au-delà de ce bilan, la deuxième année est structurée autour de trois étapes :

- 1) Finalisation de l'ontologie et validation avec experts du domaine ;
- 2) Prototypage du moteur d'inférence (3.2) pour des scénarios APT représentatifs ;
- 3) Campagne de benchmarks (temps de calcul, qualité des plans, robustesse), puis intégration dans un environnement expérimental reproductible.

5. **Conclusion**

Cette première année a permis de stabiliser une ligne scientifique cohérente pour la réponse à incident en SOC : combiner un socle ontologique, un moteur d'inférence

décisionnel explicable et une adaptation adverse contrôlée. L'enjeu de la suite est de démontrer la valeur opérationnelle de ce cadre sur des scénarios APT réalistes. La deuxième année visera à établir, par prototypage et benchmarks, jusqu'où cette approche peut améliorer la qualité de réponse tout en conservant la traçabilité des décisions.

Bibliographie

- Hammar K., Li T., Stadler R., Zhu Q., « Adaptive security response strategies through conjectural online learning », *IEEE Transactions on Information Forensics and Security*, 2025.
- Huang Z., Robin J., Herbaut N., Ben Rabah N., Le Grand B., « Toward an Intent-Based and Ontology-Driven Autonomic Security Response in Security Orchestration Automation and Response », *Enterprise Design, Operations, and Computing*, vol. 16213 of *Lecture Notes in Computer Science*, Springer, p. 266-283, 2026.
- Leblanc A., Robin J., Ben Rabah N., Huang Z., Le Grand B., « Rethinking Cybersecurity Ontology Classification and Evaluation : Towards a Credibility-Centered Framework », *Enterprise Design, Operations, and Computing*, vol. 16213 of *Lecture Notes in Computer Science*, Springer, p. 284-299, 2026.
- Nespoli P., Papamartzivanos D., Mármol F. G., Kambourakis G., « Optimal countermeasures selection against cyber attacks : A comprehensive survey on reaction frameworks », *IEEE Communications Surveys & Tutorials*, vol. 20, n° 2, p. 1361-1396, 2018.
- Oliveira Í., Sales T. P., Baratella R., Fumagalli M., Guizzardi G., « An ontology of security from a risk treatment perspective », *International conference on conceptual modeling*, Springer, p. 365-379, 2022.
- Ou X., Boyer W. F., McQueen M. A., « A scalable approach to attack graph generation », *Proceedings of the 13th ACM conference on Computer and communications security*, p. 336-345, 2006.
- Sales T. P., Baião F., Guizzardi G., Almeida J. P. A., Guarino N., Mylopoulos J., « The common ontology of value and risk », *International conference on conceptual modeling*, Springer, p. 121-135, 2018.
- Schneier B., « Attack Trees », *Dr. Dobb's Journal*, vol. 24, n° 12, p. 21-29, 1999.
- Van den Broeck G., Thon I., Van Otterlo M., De Raedt L., « DTProbLog : A decision-theoretic probabilistic Prolog », *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 24, p. 1217-1222, 2010.
- Vielberth M., Böhm F., Fichtinger I., Pernul G., « Security operations center : A systematic study and open challenges », *IEEE Access*, vol. 8, p. 227756-227779, 2020.
- Zidan K., Alam A., Allison J., Al-sherbaz A., « Assessing the Challenges Faced by Security Operations Centres (SOC) », *Future of Information and Communication Conference*, Springer, p. 256-271, 2024.
- Zonouz S. A., Khurana H., Sanders W. H., Yardley T. M., « RRE : A game-theoretic intrusion response and recovery engine », *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, n° 2, p. 395-406, 2014.

Détection d'anomalies basée sur les ontologies et l'apprentissage automatique

Application aux données environnementales

Noa Barbosa*

Encadrants : Lyliabrouk*, Nadine Cullot*, Claire Bourgeois-République*, Alexis Guyot**

* *Laboratoire Informatique de Bourgogne (LIB)*
21078 Dijon Cedex

** *Aix Marseille Univ, CNRS, LIS, Marseille*

* *Prenom.Nom@ube.fr*, ** *alexis.guyot@lis-lab.fr*

RESUME. Cet article présente les travaux menés dans le cadre d'une thèse qui débute, portant sur le peuplement d'ontologies à partir de données textuelles dans le domaine environnemental. L'objectif est d'extraire automatiquement des entités et des relations relatives aux polluants afin de peupler une ontologie et de les exploiter dans une approche hybride combinant raisonnement sur les ontologies et méthodes d'apprentissage automatique pour la détection d'anomalies environnementales liées à la pollution.

ABSTRACT. This paper presents work carried out in the context of a just started PhD thesis, focusing on ontology population from textual data in the environmental domain. The objective is to automatically extract entities and relations related to pollutants in order to populate an ontology and exploit it with a hybrid approach combining ontological reasoning and machine learning methods for environmental pollution anomaly detection.

MOTS-CLÉS : Extraction de connaissances, population d'ontologie, apprentissage automatique, détection d'anomalies, environnement, PFAS.

KEYWORDS: Knowledge extraction, ontology population, machine learning, anomaly detection, environment, PFAS.

1. Contexte et motivation

L'analyse de données repose de plus en plus sur l'exploitation de sources d'information hétérogènes, combinant des données structurées et des documents textuels non structurés. Si les premières peuvent être traitées efficacement à l'aide de méthodes statistiques ou d'apprentissage automatique, l'extraction automatique de connaissances à partir de textes demeure plus complexe. Pourtant, ces documents contiennent souvent des informations riches et complémentaires qui peuvent contribuer à une meilleure compréhension des phénomènes étudiés.

Dans ce contexte, les ontologies constituent un outil pertinent pour représenter et structurer les connaissances d'un domaine. Elles permettent de formaliser les concepts et les relations qui les relient et offrent ainsi un cadre sémantique facilitant l'intégration de données provenant de sources variées. Le peuplement d'ontologies vise à extraire automatiquement des entités et des relations à partir de différentes sources afin d'instancier les concepts et relations définies (Khadir *et al.*, 2021).

Ces approches trouvent un intérêt particulier dans le domaine de la surveillance environnementale, où l'analyse des données porte à la fois sur des mesures quantitatives issues de campagnes d'échantillonnage, et sur des informations contenues dans des rapports scientifiques, techniques ou d'expertise. Ce travail s'intéresse à l'utilisation combinée de techniques d'extraction de connaissances à partir de textes, d'ontologies et de méthodes d'apprentissage automatique pour structurer et exploiter les connaissances relatives aux polluants, afin d'améliorer la détection d'anomalies dans le domaine environnemental, comme par exemple des dépassements anormaux de seuils de concentration de PFAS, des incohérences spatio-temporelles dans les mesures environnementales ou encore des situations de non-conformité vis-à-vis des normes réglementaires.

2. Problématique

La problématique principale du travail de thèse est la suivante : *Comment extraire automatiquement des connaissances à partir de textes et d'ontologies de sorte à structurer les données environnementales, et ainsi améliorer les résultats des méthodes de détection d'anomalies basées sur l'apprentissage automatique ?*

Cette problématique se décline en trois questions de recherche : **Q1.** *Comment effectuer de la reconnaissance d'entités nommées (NER) et de l'extraction de relations (RE) à partir de textes dans un domaine spécifique (p. ex. domaine environnemental) ?* **Q2.** *Comment peupler une ontologie avec les entités et relations extraites afin de structurer des connaissances spécifiques à un domaine ?* **Q3.** *Comment exploiter l'ontologie peuplée dans une approche hybride combinant raisonnement sur les ontologies et méthodes d'apprentissage automatique pour la détection d'anomalies dans un domaine spécifique ?*

3. Architecture générale

La Figure 1 illustre la démarche générale proposée dans cette thèse. Dans un premier temps, des connaissances sont extraites à partir de textes du domaine environnemental à l'aide de techniques de reconnaissance d'entités nommées et d'extraction de relations. Les entités et relations identifiées sont ensuite utilisées pour peupler l'ontologie OntoPFAS (Di Piero *et al.*, 2026). La qualité du peuplement est ensuite évaluée selon plusieurs critères, notamment la précision des entités et relations extraites, la couverture des connaissances du domaine ainsi que la cohérence logique de l'ontologie, vérifiée à l'aide de mécanismes de raisonnement. Si nécessaire, le processus peut être itéré à partir de nouvelles connaissances extraites. Enfin, l'ontologie peuplée est exploitée dans des approches hybrides combinant connaissances ontologiques et apprentissage automatique pour la détection d'anomalies environnementales.

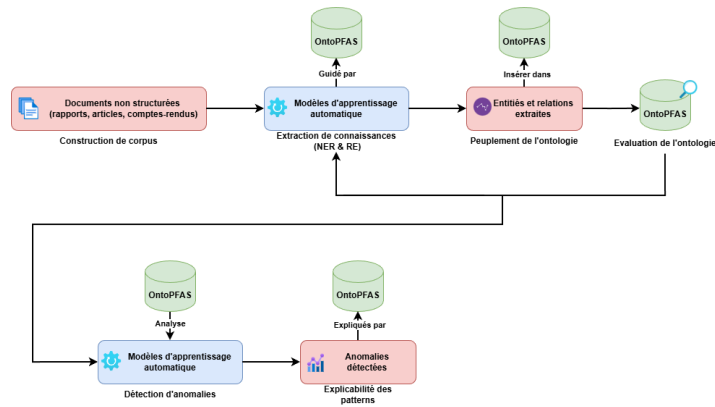


FIGURE 1. Architecture générale

4. Etat de l'art

L'extraction de connaissances à partir de textes repose sur deux tâches principales : la reconnaissance d'entités nommées (*NER*), qui permet d'identifier des instances de concepts, et ; l'extraction de relations (*RE*), qui vise à déterminer les relations entre ces instances. Les approches d'extraction de connaissances peuvent être regroupées en quatre grandes familles.

La première famille regroupe les méthodes à base de règles ou de patrons linguistiques. Par exemple, (Zhou *et al.*, 2025) proposent une méthode d'extraction de relations basée sur des règles définies à partir de connaissances expertes. Ces approches nécessitent une construction manuelle des règles et sont difficiles à généraliser à d'autres domaines.

La deuxième famille inclut les méthodes d'apprentissage automatique qui permettent d'identifier automatiquement des régularités dans les textes. Par exemple,

(Kordjamshidi and Moens, 2015) proposent une approche basée sur un modèle *Structured Support Vector Machines (SSVM)*. Ces approches améliorent la capacité de généralisation, mais leur performance reste dépendante de la qualité des données annotées disponibles.

La troisième famille, celle des approches *deep learning*, a montré de bonnes performances pour l'extraction d'information. Par exemple, (Liu and Cheng, 2024) proposent une approche basée sur un modèle *BERT-BiLSTM-CRF* qui capture les dépendances contextuelles entre les entités dans les textes. Toutefois, ces méthodes nécessitent souvent de grands volumes de données annotées et restent sensibles aux variations de domaine.

La quatrième famille est composée des méthodes à base de grands modèles de langage (*LLM*). Par exemple, (Hu *et al.*, 2025) proposent une approche basée sur du *few-shot prompting* guidé par une ontologie et reposant sur un système multi-agents. Ces approches permettent d'exploiter les connaissances contenues dans les modèles tout en réduisant l'effort d'annotation. Cependant, elles intègrent encore assez peu les connaissances externes issues d'ontologies.

5. Discussion et perspectives

Afin de répondre aux questions Q1 et Q2, nous envisageons une approche hybride combinant LLM et ontologies pour la population d'OntoPFAS. Pour répondre à Q3, nous exploiterons OntoPFAS peuplée pour proposer une approche hybride de détection d'anomalies basée sur l'apprentissage automatique et les ontologies.

Bibliographie

- Di Pierro D., Abrouk L., Guyot A., Symeonidou D., Labadie P., Lysaniuk B., « OntoPFAS : An Ontology for the Forever Chemicals », *ESWC 2026*, 2026. Accepted paper.
- Hu Y., Liu D., Wang Q., Yu C., Xu C., Zheng Q., Ji H., Xiong J., « Automating intervention discovery from scientific literature : a progressive ontology prompting and dual-LLM framework », *IJCAI '25*, Montreal, Canada, p. 9701-9709, 2025.
- Khadir A. C., Aliane H., Guessoum A., « Ontology learning : Grand tour and challenges », *Computer Science Review*, vol. 39, p. 100339, February, 2021.
- Kordjamshidi P., Moens M.-F., « Global machine learning for spatial ontology population », *Journal of Web Semantics*, vol. 30, p. 3-21, January, 2015.
- Liu D., Cheng L., « MAKG : A maritime accident knowledge graph for intelligent accident analysis and management », *Ocean Engineering*, vol. 312, p. 119280, November, 2024.
- Zhou Z., Yu X., Magoua J. J., Cui J., Luan H., Lin D., « Integrating machine learning and a large language model to construct a domain knowledge graph for reducing the risk of fall-from-height accidents », *Accident Analysis & Prevention*, vol. 215, p. 108009, June, 2025.

Semantic and Multimodal Structuring of Historical Archives: Towards New Information Systems

Mohamed Amine Lasheb*

Encadrants : Olivier Pons*

**Laboratoire CEDRIC, Conservatoire National des Arts et Métiers (Cnam)
292 rue Saint-Martin
75003 Paris
France
mohamed-amine.lasheb@lecnam.net, olivier.pons@lecnam.net*

RESUME. Les bibliothèques numériques historiques nécessitent aujourd'hui de dépasser la simple recherche textuelle pour s'orienter vers des Systèmes d'Information sémantiques. Cet article présente un axe récemment ouvert au cours de ma deuxième année de thèse, visant à transformer des archives bruitées en Graphes de Connaissances interopérables. Nous exposons les défis liés à la post-correction par Modèles de Langage (LLMs), les limites actuelles du traitement textuel, et proposons une architecture future basée sur la multimodalité et le standard IIIF pour garantir la traçabilité des factoides extraits.

ABSTRACT. Historical digital libraries are evolving beyond simple full-text search toward semantic Information Systems. This paper outlines a recently opened track during my second year of PhD research, which focuses on transforming noisy archives into interoperable Knowledge Graphs. We detail the challenges of OCR post-correction using Large Language Models (LLMs), discuss the current limits of text-only processing, and propose a future architecture relying on native multimodality and the IIIF standard to ensure the traceability of extracted factoids.

MOTS-CLÉS : Systèmes d'Information, Humanités Numériques, Graphes de Connaissances, Extraction d'Information, Multimodalité, IIIF.

KEYWORDS: Information Systems, Digital Humanities, Knowledge Graphs, Information Extraction, Multimodality, IIIF.

1. Introduction

Since the 1990s, institutions such as the Conservatoire Numérique des Arts et Métiers (Cnum)¹ have conducted large-scale digitization campaigns (Hohnsbein *et al.*, 2025). While harvesting protocols like OAI-PMH programmatically retrieve thousands of these scientific documents, heritage Information Systems (IS) face a structural challenge : storing raw data is no longer sufficient. Researchers in the Digital Humanities increasingly need tools to query these collections semantically.

Building a structured IS from a raw archive involves two main obstacles. First, Optical Character Recognition (OCR) on 18th and 19th-century typography produces noisy text, significantly reducing the performance of standard information extraction tools (Quaresma and Finatto, 2020). Second, traditional digitization processes often decouple the text from its visual layout (illustrations, spatial arrangement), leading to a loss of contextual information.

Part of my research focuses on bridging this semantic gap. The objective is to design an automated, sovereign pipeline that directly ingests Cnum archives via OAI-PMH, corrects OCR errors, and extracts structured information ("factoids") while maintaining a persistent link to the source image. This translates into the following four-step architecture :

Proposed Architecture : 1. Ingestion (OAI-PMH) → 2. Processing (VLM: OCR+Information Extraction) → 3. Structuring (Graph DB) → 4. Federation (IIIF ↔ External KBs)

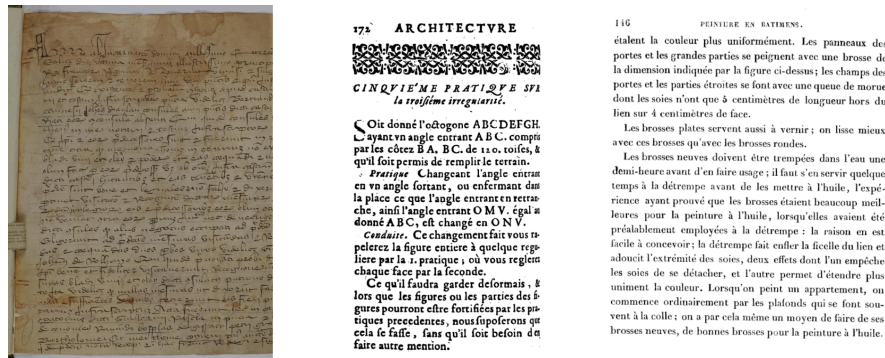


FIGURE 1. *Typographical and layout complexities in Cnum archives : Délibération des consuls de Brusque (1417), L'architecture militaire (1655), and Traité complet de la peinture (1836).*

2. State of the Art and Problem Statement

Transforming historical archives into Knowledge Graphs typically relies on sequential Natural Language Processing (NLP) pipelines. However, these systems yield low

1. <https://cnum.cnam.fr>

precision when applied to noisy OCR outputs (Quaresma and Finatto, 2020). Recent studies have explored Large Language Models (LLMs) for post-correction to address this issue (Boros *et al.*, 2024). Although effective, LLMs introduce a risk of hyper-correction. This phenomenon rewrites archaic syntax into modern prose and alters the historical integrity of the document.

Generative AI facilitates joint extraction directly from text, but relying on commercial APIs raises data sovereignty and cost concerns. On the other hand, local open-source models often face a "context wall" that makes it difficult to process long documents without generating hallucinations.

Text-centric approaches inherently discard the original visual layout (Figure 1). Standard Document Layout Analysis using computer vision (e.g., YOLO) often fails to generalize across centuries of typographical variations (Lang *et al.*, 2024).

This leads to the core research question for this track : *How can we design a sovereign, multimodal information extraction architecture that simultaneously resolves OCR noise, bypasses traditional NLP pipelines, and leverages spatial layouts to populate verifiable Knowledge Graphs ?*

3. Textual Approach : Initial Findings

During the initial phase of this work, we investigated whether Large Language Models could act as joint OCR correctors and information extractors.

Our experiments revealed a clear chronological divide². For OCR post-correction, local models like Qwen2.5 (Bai *et al.*, 2023) effectively processed modern 20th-century data (51.8% correction rate). Conversely, the commercial Gemini-2.5-Flash API (Google Team, 2023) outperformed them on 18th-century materials (62.7% correction). For downstream information extraction tasks such as Named Entity Recognition (NER), we relied exclusively on sovereign local models to ensure data privacy. This pipeline produces structured JSON factoids with high precision (>80%) and an F1 score close to 60% on 2k-token windows. To assess practical applicability, we integrated the JSON outputs into a Neo4j (Neo4j Inc., 2024) graph database, instantiating a prototype Knowledge Graph that connects entities across historical documents.

This text-only approach quickly revealed several limitations. In particular, aligning extracted elements with their actual presence in the source documents remains challenging. Preliminary experiments on keyword extraction and localization within historical pages show frequent partial matches and a large number of false positives, highlighting grounding and contextual consistency issues. Beyond a 4k-token window, local models tend to lose factual accuracy and produce hallucinations, while archaic typographic features such as the 18th-century "long s" are often misinterpreted.

2. A comprehensive evaluation of these models is the subject of a separate INFORSID 2026 submission.

4. Future Work : Towards a Multimodal and Traceable IS

To overcome these constraints, the next steps in my research will involve adapting the extraction pipeline around native multimodality and international interoperability standards.

Instead of trying to correct text that has already been degraded by an external OCR engine, we will evaluate Vision-Language Models (VLMs). We hypothesize that processing the visual facsimile directly will better resolve typographic ambiguities. It should also preserve the document's topology, particularly the semantic links between text paragraphs and historical illustrations (Lang *et al.*, 2024).

Automated extraction also raises a critical verifiability issue. For our Knowledge Graph to be a reliable tool for historians, every extracted entity requires a spatially verifiable source. We plan to interface our extraction pipeline with the existing IIF (International Image Interoperability Framework) (Snydman *et al.*, 2015) infrastructure of the Cnum to achieve this. The model will output spatial bounding boxes alongside the extracted factoids. Within our Neo4j interface, querying a specific entity will therefore not only map its semantic network but also redirect the user to the Cnum's viewer. This will visually highlight the exact region on the digitized page where the information was originally printed.

5. Conclusion

Building historical Knowledge Graphs requires moving beyond sequential extraction pipelines. By adopting local multimodal architectures and anchoring extracted data using the IIF visual standard, this work aims to support the Digital Humanities with transparent and sovereign Information Systems. In the long term, these structured databases will serve as a foundation for specialized historical models.

Bibliographie

- Bai J. *et al.*, « Qwen Technical Report », *arXiv :2309.16609*, 2023.
- Boros E. *et al.*, « Post-Correction of Historical Text Transcripts with Large Language Models », *Proc. of LaTeCH-CLjL*, 2024.
- Google Team, « Gemini : A Family of Highly Capable Multimodal Models », *arXiv :2312.11805*, 2023.
- Hohnsbein A. *et al.*, « Aux origines du Conservatoire numérique des arts et métiers (Cnum) », *Cahiers d'histoire du Cnam*, 2025.
- Lang S. A. *et al.*, « Toward a Computational Historiography of Alchemy », *CEUR Workshop Proceedings*, 2024.
- Neo4j Inc., « Neo4j Graph Database », 2024.
- Quaresma P., Finatto M. J. B., « Information Extraction from Historical Texts : a Case Study », *Proc. of the 12th Int. Conf. on Computational Processing of the Portuguese Language*, 2020.
- Snydman S. *et al.*, « The International Image Interoperability Framework (IIF) », *Archiving Conference*, 2015.

Optimization of AI Agent Systems via Process Mining: Modeling, Analysis and Improvement

Clément Adandé*

Encadrants : Ronan Champagnat*, Cyrille Suire*, Antoine Vidon**, Gabriel Jouffrai**

* *Laboratoire Informatique, Image, Interaction
La Rochelle Université, Institut LUDI, Bâtiment Pascal
17042 La Rochelle*

** *Alouette AI
12 avenue Jacques Bujault
79000 Niort*

* *Prenom.Nomauteur@univ-lr.fr*; ** *Prenom.Nomauteur@alouette.ai*

RESUME. Bien que son émergence soit récente, l'IA générative a évolué, passant de systèmes isolés de génération de texte à des systèmes agentiques complexes de plus en plus déployés en entreprise. Ces systèmes produisent des traces d'exécution structurées qui peuvent bénéficier des techniques de fouille de processus pour améliorer leurs comportements. Cependant, l'exploitation de la fouille de processus pour comprendre les comportements complexes et non déterministes de ces systèmes reste peu explorée. Cette thèse étudie comment les techniques de fouille de processus peuvent être mobilisées pour modéliser, analyser et améliorer le comportement d'exécution des systèmes d'IA agentiques.

ABSTRACT. Despite its recent emergence, generative AI has evolved from isolated text-generation systems to complex agentic systems more and more deployed across enterprises. Those systems produce structured execution traces that can benefit from process mining techniques to improve their behaviors. However, making sense of the complex and non-deterministic behaviors of these systems through process mining remains less explored. This thesis investigates how process mining techniques can be leveraged to model, analyze, and improve the execution behavior of agentic AI systems.

MOTS-CLÉS : Systèmes d'IA agentique; fouille de processus; observabilité

KEYWORDS: Agentic AI systems; process mining; observability

1. Context and Motivation

The deployment of Large Language Model (LLM)-based agentic systems is accelerating rapidly across enterprises. An agentic AI system is a software system in which one or more LLM-based agents autonomously pursue goals by planning, invoking tools, maintaining context, and adapting their behavior based on intermediate results. At the same time, process mining has developed over three decades as a mature set of techniques to extract, analyze and improve processes from event logs (van der Aalst, 2011). While agentic AI systems are recent and their behaviors remain difficult to explain, process mining provides a well-established methodological foundation for making execution behaviors analyzable. Despite early work pointing toward this convergence, the field remains under explored. And this research aims to bridge this gap by investigating how process mining can be applied to efficiently understand and improve the behavior of agentic AI systems.

2. State of the Art

Agentic systems produce multi-level execution traces, structured sequences of timestamped reasoning steps, actions, observations, and tool interactions, from the moment a user query is received to the delivery of a final answer. And many observability tools exist to capture such traces for debugging and monitoring purposes, but they do not support process-level analysis of agent execution behavior. Nevertheless, several works have begun to apply process mining to agent traces. Fournier *et al.* (2025) used process and causal discovery techniques to identify behavioral variability and execution dependencies in agent trajectories, while Dorsch *et al.* (2025) applied process mining for better prompt engineering of LLM agents. Moshkovich *et al.* (2025) proposed taxonomies for observability and analytics of agentic systems, along with a benchmark based on execution flow analysis rather than black-box evaluation. In parallel, recent studies suggest that mined process knowledge can support agent improvement, for instance Redis *et al.* (2025) showed that process mining can extract reusable skills from LLM-generated plans and feed them back into agent workflows. Agent evaluation is also expanding beyond task completion to include criteria such as consistency, efficiency, and quality Mohammadi *et al.* (2025), though process-mining-based analysis of execution behavior remains largely unexplored. Despite this progress, a clear gap remains : process mining is mature and agentic systems generate exploitable traces, but very few works systematically connect the two to analyze, explain, and improve agent execution behavior.

3. Research Objectives

Despite recent progress, the literature remains fragmented along three dimensions. First, no established framework currently guides the choice of activity granularity and semantics in agent event log construction. While case-centric process mining provides

a natural starting point, the multi-object nature of agent execution suggests the potential relevance of object-centric approaches, which remain largely unexplored in this context. Second, complex agentic tasks produce highly variable execution traces that require semantic-aware preprocessing to reduce noise and variability before process mining techniques can be effectively applied to such dynamic and stochastic workflows. Third, as agent evaluation is expanding beyond task completion, an emerging opportunity arises to correlate mined execution patterns with outcome quality, and to leverage such knowledge for targeted intervention on agent behavior. These observations lead to our central research question : how can agent execution traces be modeled, structured, and exploited to understand and improve agentic AI system behavior ? We address it through three objectives : (1) defining a semantic framework for trace segmentation with an opening toward object-centric representations, (2) identifying suited process representations and developing adapted methods to structure behavioral variability and detect deviations, and (3) enabling process-aware evaluation by correlating mined patterns with outcome quality, and designing intervention mechanisms to optimize agent behavior at design time and runtime.

4. Preliminary Work

Our preliminary work is grounded in a concrete use case : a single agent, implemented with Google ADK (Agent Development Kit), that reads raw text chunks from a technical report and iteratively calls tools to build a hierarchical tree structure (document as root, sections as internal nodes, text passages as leaves). This use case, though involving a single agent, exhibits the behavioral variability typical of agentic systems : diverse execution strategies, nested loops, and varying intensities across runs, as illustrated by the process tree in Figure 1.

Regarding the first objective, we instrumented agent runs, collected their execution records, and preprocessed them into timestamped execution events. On top of this, we have proposed a semantic convention that guides trace segmentation into process events. Concretely, each agent session maps to a case, events correspond to observable actions within a session such as tool invocations, agent-user interactions. Activities label the action performed, and each event is enriched with contextual attributes such as tool parameters, output status, and token counts.

Toward the second objective, we are developing a trace clustering approach that adapts Byte-Pair Encoding (BPE) to learn a recursive vocabulary of behavioral fragments from agent traces. These fragments are combined with numeric execution features and embedded into vector representations suitable for clustering. Clustering allows to investigate and identify different regimes of behaviors (e.g., baseline runs, heavy iterative strategies). This serve as a basis for performance profiling and targeted intervention.

5. Conclusion and Future Work

This research investigates how process mining can transform raw agent execution traces into actionable process knowledge, addressing challenges of trace modeling, behavioral structuring, and process-aware evaluation. Our next steps include validating

the proposed semantic framework on multi-agent use cases beyond a single agent and framework, extending trace clustering to support runtime classification of partial traces, and developing process-aware evaluation methods that correlate execution patterns with outcome quality.

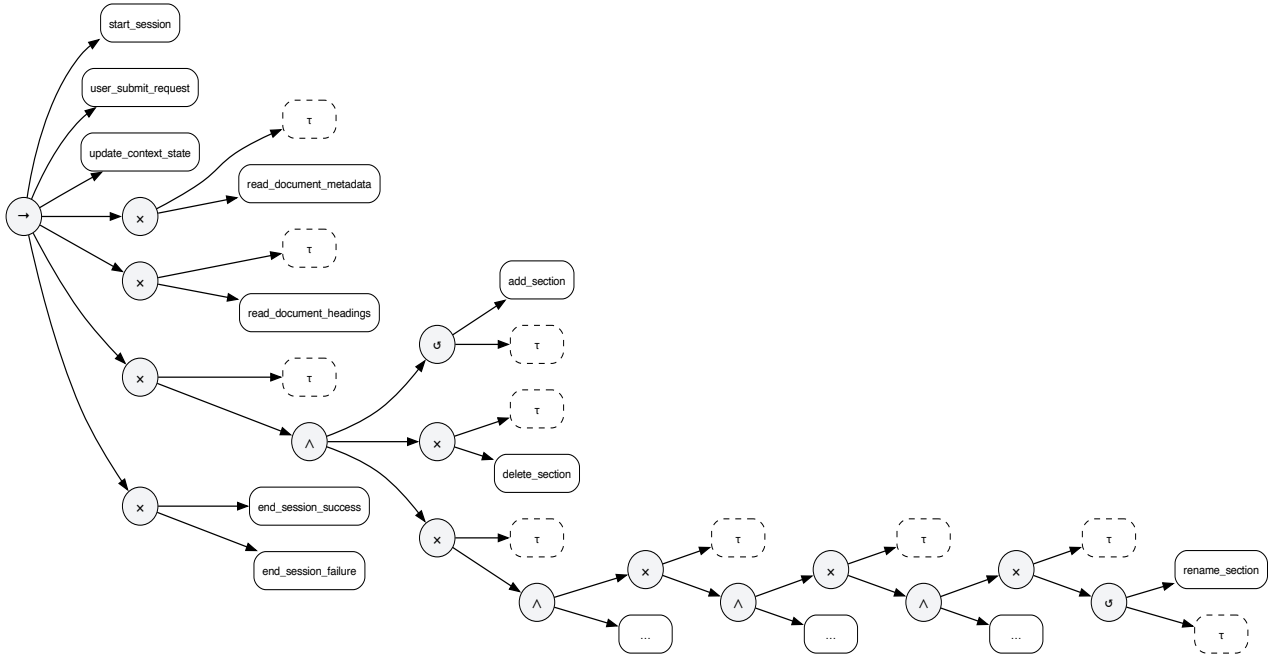


FIGURE 1. Process tree discovered from 56 agent sessions using the Inductive Miner algorithm. Each leaf represents an action issued by the agent; τ denotes a silent transition (i.e., a path that may be skipped). Operator nodes denote control-flow relations (\rightarrow sequence, \wedge parallelism, \times exclusive choice, \circlearrowleft loop). Deep subtrees are collapsed as \dots for space.

Bibliographic

- Dorsch R., Henselmann D., Harth A., « COMPASS : A Process Mining-based Methodology for Prompt Analysis of Large Language Model Agents », *CEUR Workshop Proceedings*, 2025.
- Fournier F., Limonad L., David Y., « Agentic AI Process Observability : Discovering Behavioral Variability », 2025.
- Mohammadi M., Li Y., Lo J., Yip W., « Evaluation and Benchmarking of LLM Agents : A Survey », 2025.
- Moshkovich D., Mulian H., Zeltyn S. *et al.*, « Beyond Black-Box Benchmarking : Observability, Analytics, and Optimization of Agentic Systems », 2025.
- Redis A., Sani M., Zarrin B., Burattin A., « Skill Learning Using Process Mining for Large Language Model Plan Generation », p. 650-662, March, 2025.
- van der Aalst W. M. P., *Process Mining : Discovery, Conformance and Enhancement of Business Processes*, Springer, 2011.

Un Cadre Décisionnel Basé sur Apprentissage Automatique pour la Santé et la Sécurité au Travail dans le cadre du Transport Intelligent

Cynthia Chedrawi^{*,}**

Encadrants : Elena Kornyshova^{*}, Laurent Boutal^{*,}**

** Conservatoire National des Arts et Métiers (CNAM), 2 Rue Conté, Paris*

*** Régie Autonome des Transports Parisiens (RATP), 54 Quai de la Rapée, Paris*

cynthia.chedrawi@lecnam.net

RESUME. Les approches de Santé et la Sécurité au Travail (SST) visent à prévenir les accidents et à protéger les travailleurs. De nombreux secteurs utilisent aujourd'hui l'Internet des Objets (IoT) et l'Apprentissage Automatique (Machine Learning-ML) pour mieux surveiller les risques, détecter les dangers et adopter des stratégies de prévention les plus proactives. Toutefois, leur utilisation structurée dans le domaine du transport reste limitée. Cette thèse vise à développer un cadre méthodologique permettant de préconiser les algorithmes de ML les plus adaptés aux cas d'usage les plus représentatifs de SST en fonction du contexte, à savoir des objectifs associés à la SST, des caractéristiques des dispositifs IoT utilisés, du type de données, etc. L'approche sera développée au sein de la Régie Autonome des Transports Parisiens (RATP).

ABSTRACT. Occupational Health and Safety seeks to prevent accidents and protect workers. Many sectors now use IoT technologies and ML methods to monitor risks, detect hazards earlier, and support more proactive safety management. However, their structured use in the transportation sector remains limited. This thesis aims to develop a methodological framework to recommend the most suitable ML algorithms for the most representative occupational health and safety use cases, depending on the context, namely the SST-related objectives, the characteristics of the IoT devices used, the type of data, etc. The approach will be developed within the Paris Autonomous Transport Authority (RATP).

MOTS-CLÉS : IoT, Apprentissage Automatique, Transport Intelligent, Santé et Sécurité au Travail

KEYWORDS: IoT, Machine Learning, Smart Transportation, Occupational Health and Safety

1. Introduction

L'intégration de l'IoT et du ML a profondément transformé la SST dans plusieurs secteurs à haut risque. Les dispositifs IoT génèrent des données sur l'état et les mouvements des travailleurs, qui, combinées au ML, permettent de détecter les risques et de mettre en place des stratégies de prévention plus proactives. Dans le secteur du transport, IoT et ML sont principalement utilisées pour des applications telles que la gestion du trafic et la conduite autonome (Azad *et al.*, 2024), tandis que leur usage pour la SST reste encore limitée. Pourtant, les risques sont nombreux : chutes lors des opérations de maintenance, exposition à des environnements confinés ou fatigue des conducteurs. La prévention peut s'appuyer sur des capteurs IoT variés (environnementaux, de mouvement, systèmes embarqués). Dans ce contexte, ce papier présente les travaux menés dans le cadre d'une thèse CIFRE débutée en 2025 au sein de la Régie Autonome des Transports Parisiens (RATP), visant à développer une approche méthodologique exploitant l'IoT et le ML pour améliorer la SST dans le transport.

2. Problématique

Le déploiement des IoT et des méthodes de ML dans le transport intelligent soulève plusieurs défis, notamment l'identification des situations à risque, la qualité et la fiabilité des données, les contraintes de latence et de robustesse des modèles, l'interprétabilité des résultats ainsi que l'acceptabilité par les travailleurs. Ces limitations mettent en évidence l'absence d'un cadre structuré permettant d'identifier et de recommander les méthodes de ML les plus adaptées aux différentes situations de SST. Cela conduit à la question de recherche suivante : **Comment l'IoT et le ML peuvent-ils améliorer la santé et la sécurité au travail dans les systèmes de transport ?** L'objectif de la thèse est de proposer un cadre permettant de recommander les algorithmes de ML les plus adaptés pour exploiter efficacement les dispositifs IoT au service de la SST dans le secteur du transport.

3. Méthodologie de Recherche

Le travail de cette thèse est structuré en quatre phases complémentaires :

Phase 1 : État de l'art. Cette phase consiste à réaliser une revue systématique de la littérature selon la méthodologie de Kitchenham *et al.* (Kitchenham *et al.*, 2009). L'objectif est d'analyser comment les technologies IoT et les méthodes de ML sont utilisées pour traiter les problématiques de SST dans différents secteurs. L'objectif est d'identifier les solutions existantes, les types de données utilisés, les dispositifs IoT déployés et les algorithmes employés.

Phase 2 : Étude empirique des situations SST à la RATP. Cette phase consiste en une étude empirique (qui sera réalisée selon l'approche de (Runeson and Höst, 2009)) des pratiques et projets actuels de SST à la RATP. Elle vise à identifier les besoins réels, à confronter les situations observées à la littérature et à évaluer la faisabilité des

solutions envisagées.

Phase 3 : Développement d'un cadre d'aide à la décision pour la gestion des situations de SST. Cette phase vise à définir un cadre décisionnel d'aide à la décision, basé sur des patterns identifiés dans les deux premières phases. Ce cadre correspond à un ensemble structuré de connaissances permettant de relier : une situation de SST, son contexte, les dispositifs IoT disponibles, et les familles de méthodes de ML adaptées. Le cadre permettra de guider le choix des solutions techniques, notamment de solutions ML, en tenant compte de la qualité des données et en fonction des caractéristiques d'une situation donnée.

Phase 4 : Validation sur des cas réels. Cette phase inclut le test du cadre sur plusieurs cas d'usage du réseau RATP. Les performances des méthodes ML recommandées seront évaluées et les retours des équipes permettront d'affiner le cadre final.

4. Actions réalisées et futures

Nous avons débuté une revue systématique de la littérature (selon (Kitchenham *et al.*, 2009)) afin d'identifier et d'analyser les approches IoT et ML appliquées à la SST. Comme notre objectif porte sur les problématiques de SST, notre recherche se concentre principalement sur les travaux traitant de ce thème, tous secteurs confondus. Nous examinons ainsi des contributions issues de divers domaines à risques afin de repérer les avancées méthodologiques pertinentes et d'évaluer leur potentiel d'adaptation au contexte du transport. Nous avons d'abord défini plusieurs questions de recherche pour guider l'extraction des informations :

- RQ1 : Quelle est l'évolution des travaux intégrant IoT et ML pour la SST ?
- RQ2 : Dans quels secteurs et pour quels objectifs ces technologies sont-elles mobilisées ?
- RQ3 : Quels types de dispositifs IoT, de données et d'infrastructures sont utilisés ?
- RQ4 : Comment les méthodes de ML exploitent-elles les données IoT pour améliorer la SST ?

Une stratégie de mots-clés a ensuite été établie autour des trois concepts principaux (IoT, ML, SST), complétés par leurs synonymes, et combinés à l'aide d'opérateurs booléens, comme présenté ci-dessous :

```
('IoT' OR 'Internet of Things') AND ('Artificial Intelligence' OR 'AI' OR "Machine Learning" OR "ML" OR "unsupervised learning" OR "supervised learning" OR "reinforcement learning" OR "deep learning") AND ("Occupational safety and health" OR "OSH" OR "Occupational health and safety" OR "OHS" OR "Workplace safety" OR "Workplace health" OR "WHS" OR "Industrial safety" OR "workplace well-being" OR "workplace wellbeing").
```

Cette requête de recherche a été appliquée aux titres, résumés et mots-clés des articles.

La recherche bibliographique a été ensuite réalisée via l'API Scopus, donnant accès à des bases principales telles qu'Elsevier, Springer, IEEE, etc. Cette procédure

permet d’avoir un panorama complet des développements technologiques liés à la SST. Notre requête a abouti à 163 articles, auxquels seront appliqués des critères d’inclusion/exclusion afin de retenir les études les plus pertinentes. À l’issue de cette analyse, nous proposerons un cadre comparatif synthétisant les principales caractéristiques identifiées, ce qui permettra de clarifier le rôle concret des approches IoT et ML pour la SST et d’identifier les opportunités d’application à la RATP.

Nous mènerons ensuite une étude empirique à la RATP afin d’identifier les situations typiques de SST, servant à la construction d’un cadre décisionnel pour sélectionner les techniques de ML adaptées. Des patterns seront définis pour relier chaque situation à des caractéristiques telles que les types de données, les dispositifs IoT et les contraintes opérationnelles. Sur cette base, le cadre proposera les algorithmes de ML les plus appropriés. Ce processus est illustré dans la Figure 1, qui synthétise le passage de l’analyse des situations de SST à la recommandation de techniques de ML.

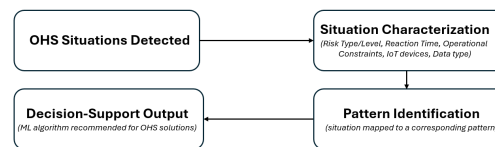


FIGURE 1. *Cadre décisionnel pour les situations de SST*

5. Conclusion

Cette thèse vise à explorer le potentiel de l’IoT et du ML pour améliorer la SST dans le domaine du transport. En combinant une revue systématique de la littérature et l’analyse de situations réelles à la RATP, elle propose un cadre décisionnel reliant les besoins de sécurité aux solutions IoT et ML les plus adaptées. Ce cadre a pour objectif de soutenir la prise de décision et sera validé à travers des cas d’usage réels, contribuant ainsi à renforcer la sécurité des travailleurs. Bien que cette approche soit appliquée ici au contexte de la SST dans le transport, elle peut être généralisée à d’autres domaines caractérisés par des données IoT hétérogènes et des problématiques de décision intelligente.

Bibliographie

- Azad A. K., Atkison T., Shah A. F. M., « A review on machine learning in intelligent transportation systems applications », *The Open Transportation Journal*, 2024.
- Kitchenham B., Brereton O. P., Budgen D., Turner M., Bailey J., Linkman S., « Systematic Literature Reviews in Software Engineering – A Systematic Literature Review », *Information and Software Technology*, vol. 51, p. 7-15, 2009.
- Runeson P., Höst M., « Guidelines for Conducting and Reporting Case Study Research in Software Engineering », *Empirical Software Engineering*, vol. 14, p. 131-164, 2009.

Collaboration Dans Les Méthodes Agiles: Comment l'IA impacte le processus agile ?

Haoyue Liu¹⁻²

Encadrants : Rebecca Deneckère¹, François Charoy²

1. Centre de Recherche en Informatique, Université Paris 1 Panthéon-Sorbonne
90 rue Tolbiac, Paris, France

haoyue.liu@inria.fr, rebecca.deneckere@univ-paris1.fr

2. INRIA, Université de Lorraine

615 Rue du Jardin-Botanique, 54600 Villers-lès-Nancy

francois.charoy@univ-lorraine.fr

RÉSUMÉ. L'intelligence artificielle occupe une place croissante au sein des équipes agiles, transformant les modes de collaboration dans les environnements de développement logiciel. Cette thèse examine les conditions dans lesquelles des équipes agiles hybrides, composées de membres humains et d'outils d'IA, peuvent collaborer efficacement. Elle s'appuie sur l'analyse des traces numériques issues des outils collaboratifs afin d'identifier des schémas d'interaction, des pratiques émergentes et des points de tension organisationnels. Une attention particulière est portée aux dimensions techniques, organisationnelles et éthiques liées à l'intégration d'agents intelligents.

ABSTRACT. Artificial intelligence is playing an increasingly prominent role within agile teams, transforming modes of collaboration in software development environments. This PhD thesis examines the conditions under which hybrid agile teams—composed of human members and AI-based tools—can collaborate effectively. It relies on the analysis of digital traces generated by collaborative tools to identify interaction patterns, emerging practices, and organizational tensions. Particular attention is given to the technical, organizational, and ethical dimensions related to the integration of intelligent agents.

MOTS-CLES : Intelligence artificielle, Agilité, développement logiciel, collaboration humain-IA, collaboration à long-terme, Fouille de processus

KEYWORDS: Artificial intelligence, agility, software development, human-AI collaboration, Long-term collaboration, Process Mining

1. Introduction

Depuis le Manifeste Agile (*K.Beck et al., 2001*), les méthodologies agiles ont profondément transformé le développement logiciel en mettant l'accent sur les cycles itératifs, la collaboration externe/interne et l'adaptation au changement. Les travaux

de Sharp et Robinson soulignent l'importance des trois dimensions centrales de l'agilité : collaboration, coordination et communication (H.Sharp, H.Robinson, 2010), la collaboration humaine jouant un rôle central. Des outils collaboratifs tels que Jira, Trello, Slack, Microsoft Teams ou Confluence facilitent les échanges en temps réel, le partage des connaissances et la visibilité du travail, en particulier dans des contextes distribués. Parallèlement, l'intelligence artificielle connaît une intégration rapide dans les environnements de développement logiciel (T. Herda et al., 2025). Les outils fondés sur le NLP améliorent la qualité des exigences et des user stories (E. Babulak et al., 2025), tandis que les systèmes d'IA générative, tels que ChatGPT et Copilot, assistent la production de contenu technique et la prise de décision (V. Stray et al., 2025). Ces évolutions modifient les pratiques de travail ainsi que les dynamiques de collaboration au sein des équipes de développement .

Bien que des études récentes (T. Herda et al., 2025 ; E. Babulak et al., 2025) aient montré l'intérêt de l'IA pour le développement agile, notre étude répond au besoin d'une évaluation longitudinale. Cette lacune nécessite d'analyser à la fois les perceptions des acteurs et les changements observables dans les pratiques collaboratives et de développement. Dans cette recherche, nous proposons d'utiliser la fouille de processus comme cadre méthodologique pour étudier ces évolutions (W. Van Der Aalst, 2012). L'analyse des journaux d'événements issus des outils collaboratifs et de développement permet d'observer l'évolution des flux de travail agiles, l'introduction d'activités assistées par l'IA et leurs effets sur la coordination, la communication et la réalisation des tâches. Cette analyse sera complétée par une investigation qualitative afin de mieux comprendre les usages et les tensions liés à la collaboration humain-IA. L'originalité de cette recherche réside ainsi dans l'analyse longitudinale des traces numériques et dans la construction progressive d'un support d'aide à la collaboration humain-IA adapté aux pratiques observées. Dans ce contexte, la principale question de recherche de cette étude est la suivante : **Comment l'IA impacte-t-elle le processus de développement agile ?**

Cette étude se concentre sur : **i)** Comprendre l'évolution des pratiques agiles dans des environnements fortement outillés numériquement et identifier les transformations organisationnelles associées. **ii)** Identifier des bonnes pratiques de collaboration hybride humain-IA adaptées aux contextes agiles. **iii)** Analyser l'évolution de ces pratiques dans le temps, selon une approche longitudinale, afin de comprendre leurs dynamiques de stabilisation et de transformation. Cette étude est structurée comme suit : la section 2 présente l'approche proposée et la section 3 décrit les actions réalisées et les travaux futurs.

2. Approche proposée

L'approche proposée vise à comprendre l'évolution des pratiques décisionnelles et opérationnelles des équipes agiles intégrant des outils d'IA, afin de concevoir un support d'aide à la collaboration hybride. Elle est structurée en deux processus complémentaires. Cette décomposition permet de distinguer, d'une part, la formulation de recommandations à partir d'un contexte initial et, d'autre part, leur ajustement progressif à partir des pratiques effectivement observées. L'approche

adopte ainsi une logique itérative cohérente avec les principes agiles : observer, recommander, expérimenter, évaluer, puis ajuster. (Voir Figure 1).

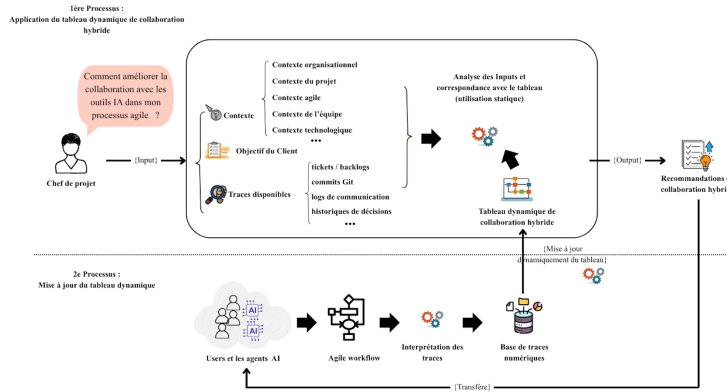


Figure 1. Approche de recherche pour une collaboration hybride adaptative

Application du tableau dynamique de collaboration hybride (premier processus). Le premier processus présente l'utilisation du tableau dynamique de collaboration hybride afin de formuler des recommandations adaptées au contexte agile étudié. Ce processus mobilise un ensemble d'entrées comprenant les différents contextes: contexte organisationnel et projet, objectifs, ainsi que traces numériques issues des outils collaboratifs (backlogs, tickets, commits, logs de communication). Ces données sont analysées pour produire des recommandations contextualisées de collaboration humain-IA.

Mise à jour du tableau dynamique (deuxième processus). Le 2e processus permet assurer l'évolution continue du tableau dynamique de collaboration hybride. Les interactions réelles entre acteurs humains et outils d'IA génèrent des traces numériques analysées par des techniques de fouille de processus. Cette analyse permet d'identifier des schémas récurrents et des écarts collaboratifs. Lorsque les résultats sont pertinents, le tableau est ajusté afin de rester aligné avec les pratiques observées.

3. Actions réalisées et travaux futurs

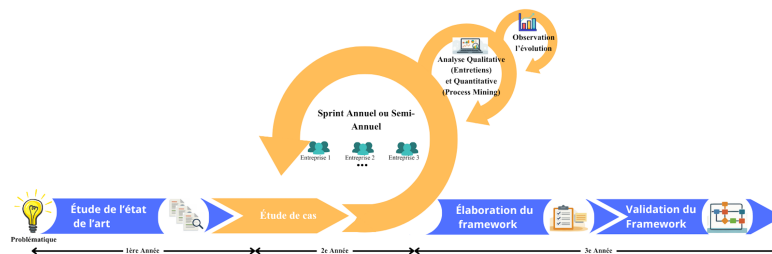


Figure 2. Feuille de route de la recherche doctorale (3 ans)

Au cours des premiers mois de la recherche (*voir Figure 2*), nous avons conduit une revue de littérature afin d'identifier l'état de l'art en matière de collaboration agile et d'intégration des outils d'IA. Nous avons porté une attention particulière aux techniques collaboratives fondées sur l'IA, aux défis liés à leur adoption et aux limites des propositions existantes.

En parallèle, nous sommes en contact avec des entreprises disposant d'équipes agiles utilisant des outils d'IA, dans le cadre d'entretiens semi-structurés menés de façon régulière. Des entretiens ont déjà été réalisés au sein de trois entreprises différentes. Afin de constituer une base de traces numériques diversifiée, nous cherchons à multiplier les études de cas. À l'issue de chaque entretien, et lorsque les conditions le permettent, des traces numériques issues des outils collaboratifs sont collectées : backlogs, user stories, tickets, historiques de commits ou logs de communication. Ces données sont exploitées à l'aide de techniques de fouille de processus pour identifier des schémas de collaboration, des variations de pratiques et des écarts organisationnels. Ainsi, la validation possible reposera à la fois sur l'analyse sur plusieurs études de cas et le retour qualitatif des équipes agiles sur l'utilité du support de collaboration hybride.

Bibliographie

- Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R.C., Mellor, S., Schwaber, K., Sutherland, J., Thomas, D. (2001). *Manifesto for Agile Software Development*. Retrieved from <https://agilemanifesto.org>
- Sharp, H., & Robinson, H. (2010). « Three 'C's of agile practice: collaboration, co-ordination and communication ». In *Agile software development: current research and future directions*(pp. 61-85). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Herda, T., Pichler, V., Zhang, Z., Abrahamsson, P., & Hanssen, G. K. (2025, June). AI and Agile Software Development: From Frustration to Success XP2025 Workshop Summary. In International Conference on Agile Software Development (pp. 3-13). Cham: Springer Nature Switzerland.
- Babulak, E., Kumar, A., Guru, D., Kudithipudi, K., Maddini, S., Gonaygunta, H., & Nadella, S. (2026). « Human-Centric AI Tools Into Agile Methodologies for Optimized Software Development ». In *AI-Driven Approaches for Fully Automated Smart Engineering*(pp. 41-84). IGI Global Scientific Publishing.
- Stray, V., Barbala, A., & Wivestad, V. T. (2025, June). « Human-AI Collaboration in Software Development: A Mixed-Methods Study of Developers' Use of GitHub Copilot and ChatGPT ». In Proceedings of the 33rd ACM International Conference on the Foundations of Software Engineering (pp. 1325-1332).
- Van Der Aalst, W. (2012). Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems (TMIS)*, 3(2), 1-17.

Approche multimodale pour la détection précoce des troubles cognitifs par apprentissage automatique

Mohamed Hedi DJEMAA*

Encadrants : Yohann Chasseray, Rafika Thabet, Farah Jemili, Elyes Lamine

* Université de Toulouse, INU Champollion, Castres, France
mohamed_hedi.djemaa@univ-jfc.fr;

RESUME. La détection précoce des troubles cognitifs à partir de données multimodales hétérogènes fait face à des défis majeurs : rareté des données, désalignement temporel et déséquilibre des classes. Les contributions de ce travail sont les suivantes : (1) une analyse comparative de deux stratégies d'intégration multimodale couplées à l'augmentation (Embed-Then-Balance et Balance-Then-Embed), mettant en évidence l'impact de l'ordre des opérations sur la cohérence sémantique inter-modale ; (2) une validation expérimentale sur des données réelles démontrant la supériorité de l'approche ETB dans un contexte de déséquilibre de classes ; (3) la proposition d'un cadre unifié articulant apprentissage multimodal et ingénierie des connaissances pour guider la collecte de données.

MOTS-CLÉS : Données Multimodales, Détection précoce, Fusion, Gestion des connaissances

1. Introduction

Les troubles cognitifs et neurodégénératifs se caractérisent par une perte progressive et irréversible de neurones dans des zones spécifiques du cerveau, entraînant une altération des fonctions exécutives, motrices et comportementales. Aujourd'hui, plus de 55 millions de personnes sont touchées, et ce chiffre pourrait atteindre 152 millions d'ici 2050 (OMS, Organisation mondiale de la santé, 2025). Dans la pratique clinique, le diagnostic repose encore majoritairement sur des évaluations statiques et ponctuelles,

souvent réalisées à un stade avancé de la maladie, dépendantes de la subjectivité de l'observateur et ne capturant qu'un instantané de l'état du patient. Or, des modifications subtiles de la prosodie vocale, de l'amplitude d'écriture ou de la démarche peuvent constituer des signes précoces. L'essor des technologies numériques et de l'intelligence artificielle (IA) ouvre ainsi la voie à une détection précoce et à une surveillance continue via des capteurs (smartphones, montres connectées, caméras), permettant de collecter des biomarqueurs dans l'environnement naturel du patient. Cependant, ces données sont massivement hétérogènes, fragmentées, déséquilibrées et temporellement désalignées, ce qui constitue un verrou majeur pour leur exploitation par des modèles d'apprentissage automatique. Ce travail propose un cadre d'analyse des stratégies de fusion et d'augmentation multimodales pour améliorer la robustesse des modèles de détection précoce. La maladie de Parkinson est retenue comme cas d'étude initial, en raison de la disponibilité de données multimodales structurées, les méthodes proposées étant conçues pour être transposables à d'autres troubles neurodégénératifs.

2. État de l'art

L'évolution des systèmes d'information de santé vers la détection précoce des troubles cognitifs suppose la capacité à traiter une complexité croissante de données. L'analyse de la littérature montre que l'efficacité des modèles de détection repose à la fois sur les algorithmes exploités et sur la maîtrise des connaissances expertes nécessaires à la construction et à la fusion des jeux de données multimodaux. La problématique multimodale dépasse le cadre des troubles cognitifs : des travaux fondateurs ont structuré le domaine autour des défis transverses de représentation, alignement et fusion (Baltrušaitis *et al.*, 2018; Ramachandram and Taylor, 2017). Sur cette base, la modélisation multimodale des troubles cognitifs s'articule autour de trois piliers : représentation, fusion et augmentation. Des traitements algorithmiques spécifiques aux modalités sont nécessaires avant combinaison (Shokrpour *et al.*, 2025; Zhang *et al.*, 2024), et la fusion précoce s'avère efficace pour des modalités temporellement synchronisées (tracés, écriture, échelles cliniques) (Benredjem *et al.*, 2024). Face au déséquilibre des classes, l'augmentation par réseaux adversaires génératifs (GAN) a été validée dans d'autres contextes biomédicaux, comme l'imagerie pulmonaire en tomodensitométrie (Wang *et al.*, 2019), mais reste rarement couplée à une fusion précoce dans le cadre des troubles cognitifs. La sélection préalable des modalités pertinentes est un enjeu pour tout système de détection précoce des troubles cognitifs. Certains travaux montrent l'intérêt d'avoir à disposition une base de connaissance pour augmenter les modèles multimodaux de détection (Lahiri and Hu, 2025; Li *et al.*, 2024). Nian *et al.* (2022) proposent une méthode d'extraction automatique de triplets depuis la littérature biomédicale permettant de cartographier les relations entre modalités diagnostiques, biomarqueurs et stades de la maladie d'Alzheimer. (Yang *et al.*, 2025) prolongent cette approche en utilisant une base de connaissances fédérant imageries, données cliniques et génomiques pour guider la collecte multimodale. Néanmoins ces approches ne s'appuient pas nécessairement sur une structure de connaissances pré-existante à l'extraction, ou se limitent à l'extraction de certaines relations.

3. Problématique

Ce travail vise à exploiter des signaux multimodaux hétérogènes pour la détection précoce des troubles cognitifs. Trois verrous majeurs sont identifiés : (i) la rareté de données multimodales appariées dans les bases publiques, (ii) le désalignement temporel et la non-correspondance entre modalités (sessions asynchrones, dispositifs hétérogènes, données manquantes), et (iii) l'absence de formalisation des connaissances expertes pour guider la collecte. Ces verrous sont adressés conjointement par des stratégies d'apprentissage multimodal robustes couplées à une représentation structurée de la connaissance experte.

4. Actions réalisées

Les travaux effectués jusqu'ici ont consisté à la définition générique des approches ETB et BTE. Ces deux méthodes ont été instanciées et évaluées sur le jeu de données PADS (*Parkinson's Disease Smartwatch*), relatif à la maladie de Parkinson, qui regroupe (i) des séries temporelles de mouvement issues de capteurs (accéléromètre et gyroscope), (ii) 30 indicateurs binaires sur les symptômes non-moteurs de l'auto-questionnaire PDNMS (*Parkinson's Disease Non-Motor Symptoms*), ainsi que (iii) des métadonnées cliniques et démographiques (âge, antécédents, latéralité, etc.). Pour traiter ces données, la fusion précoce repose sur des autoencodeurs dédiés à chaque modalité puis fusionnés par concaténation, réduisant l'espace de représentation à 90 dimensions. La stratégie ETB effectue l'augmentation par GAN dans cet espace latent compact à partir d'un unique générateur, tandis que la stratégie BTE fait appel à 13 GANs indépendants directement dans l'espace brut avant l'application des autoencodeurs. Ces deux approches sont évaluées sur une tâche aval de classification. L'approche ETB surpasse BTE en classification binaire (AUC , *Area Under the ROC Curve*; $AUC_{bin} = 0.92$ contre 0.80) et multi-classe ($AUC_{mult} = 0.95$ contre 0.79). Ce résultat suggère que l'augmentation dans un espace latent compact favorise un apprentissage plus stable et préserve la cohérence inter-modale, contrairement à l'approche BTE opérant dans un espace de plus grande dimension. Le choix de l'espace de représentation apparaît ainsi déterminant pour les stratégies d'augmentation multimodale en contexte de données rares et déséquilibrées.

5. Actions futures

Les travaux en cours visent à définir une méthodologie d'agrégation de la connaissance contenue dans la littérature scientifique relative à l'exploitation des données liées aux troubles cognitifs. Pour cela, il est prévu de construire un ensemble de questions de compétences, validées à l'aide d'un expert du domaine qui guideront la construction d'une structure ontologique qui guidera l'extraction d'informations sous la forme de triplets RDF. Une perspective future consiste alors à utiliser les graphes de connaissances instanciés au sein d'un système d'aide à la décision pour la récolte de données en fonc-

tion du contexte du patient et pour orienter les modèles de fusion définis précédemment en fonction de la prévalence des modalités dans certains contextes.

6. Conclusion

Ce travail propose une approche intégrée associant fusion multimodale et ingénierie des connaissances pour la détection précoce des troubles cognitifs. La stratégie ETB démontre l'intérêt d'opérer dans un espace latent compact pour pallier le déséquilibre des classes. Les prochaines étapes enrichiront ces modèles par extraction structurée de la connaissance issue de la littérature.

Bibliographie

- Baltrušaitis T., Ahuja C., Morency L.-P., « Multimodal machine learning : A survey and taxonomy », *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, n° 2, p. 423-443, 2018.
- Benredjem S., Mekhaznia T., Abdulghafor R., Turaev S., Bennour A., Sofiane B., Aborujilah A., Al Sarem M., « Parkinson's disease prediction : An attention-based multimodal fusion framework using handwriting and clinical data », *Diagnostics*, vol. 15, n° 1, p. 4, 2024.
- Lahiri A. K., Hu Q. V., « AlzheimerRAG : Multimodal retrieval-augmented generation for clinical use cases », *Machine Learning and Knowledge Extraction*, vol. 7, n° 3, p. 89, 2025.
- Li D., Yang S., Tan Z., Baik J. Y., Yun S., Lee J., Chacko A., Hou B., Duong-Tran D., Ding Y. *et al.*, « Dalk : Dynamic co-augmentation of llms and kg to answer alzheimer's disease questions with scientific literature », *Findings of the Association for Computational Linguistics : EMNLP 2024*, p. 2187-2205, 2024.
- Nian Y., Hu X., Zhang R., Feng J., Du J., Li F., Bu L., Zhang Y., Chen Y., Tao C., « Mining on Alzheimer's diseases related knowledge graph to identify potential AD-related semantic triples for drug repurposing », *BMC bioinformatics*, vol. 23, n° 6, p. 407, 2022.
- OMS, Organisation mondiale de la santé, « Dementia », , <https://www.who.int/news-room/fact-sheets/detail/dementia>, 2025. 31 mars 2025, consulté le 9 octobre 2025.
- Ramachandram D., Taylor G. W., « Deep multimodal learning : A survey on recent advances and trends », *IEEE signal processing magazine*, vol. 34, n° 6, p. 96-108, 2017.
- Shokrpour S. *et al.*, « Machine learning for Parkinson's disease : a comprehensive review of datasets, algorithms, and challenges », *npj Parkinson's Disease*, vol. 11, n° 1, p. 187, 2025.
- Wang Q., Zhou X., Wang C., Liu Z., Huang J., Zhou Y., Li C., Zhuang H., Cheng J.-Z., « WGAN-based synthetic minority over-sampling technique : Improving semantic fine-grained classification for lung nodules in CT images », *IEEE Access*, vol. 7, p. 18450-18463, 2019.
- Yang Y., Yu K., Gao S., Yu S., Xiong D., Qin C., Chen H., Tang J., Tang N., Zhu H., « Alzheimer's disease knowledge graph enhances knowledge discovery and disease prediction », *Computers in Biology and Medicine*, vol. 192, p. 110285, 2025.
- Zhang W., Mao K., Chen J., « A multimodal approach for detection and assessment of depression using text, audio and video », *Phenomics*, vol. 4, n° 3, p. 234, 2024.

Une approche performante et explicable pour la prévision des ventes de tickets au cinéma

Lorenzo Lefevé¹

Encadrants : Bénédicte Le Grand¹, Lionel Prevost², Nourhène Ben Rabah¹

1. Centre de Recherche en Informatique, Université Paris 1 Panthéon-Sorbonne
90 rue Tolbiac, 75013 Paris, France.

lorenzo.lefeve@etu.univ-paris1.fr, benedicte.le-grand@univ-paris1.fr,
nourhene.ben-rabah@univ-paris1.fr

2. Learning Data and Robotics, ESIEA,
4 All. Katherine Johnson, 94200 Ivry-sur-Seine, France
lionel.prevost@esiea.fr

RESUME. L'industrie du cinéma fait face à de nouveaux défis : baisse de fréquentation post-COVID, essor des plateformes de streaming et évolution des usages. Dans ce contexte, la prévision des entrées devient plus incertaine, tandis que la nécessité de produire des résultats explicables reste forte en contexte métier. Cet article étudie une approche de machine learning pour prédire les ventes de tickets au cinéma à partir de variables liées au film, au marketing et à l'opinion du public. Les résultats montrent que cette combinaison améliore la qualité des prévisions et permet d'atteindre un niveau comparable, voire légèrement supérieur, à celui d'experts du secteur sur le jeu de test.

ABSTRACT. The film industry is facing new challenges, including post-COVID attendance decline, the rise of streaming platforms, and changing audience habits. In this context, ticket-sales prediction has become more uncertain, while explainability remains essential for professional use. This paper investigates a machine-learning approach for predicting cinema ticket sales using movie-related, marketing, and audience-opinion features. The results show that combining these features improves prediction quality and reaches a level comparable to, or slightly better than, that of industry experts on the test set.

MOTS-CLÉS : Films, explicabilité, performance, Machine Learning, Feature Engineering, Prévision des entrées.

KEYWORDS : Movies, explainability, performance, Machine Learning, Feature Engineering, Ticket sales prediction.

1. Introduction

La prévision des entrées d'un film avant sa sortie constitue un enjeu économique important pour les professionnels du secteur. Dans un contexte marqué par l'incertitude sur la fréquentation en salle, l'essor des plateformes de streaming et l'évolution des usages, estimer les entrées reste difficile. Les principaux verrous concernent l'hétérogénéité des données, la prédiction directe d'une variable continue plutôt qu'une simple classe de succès, et la nécessité de produire des résultats explicables en contexte métier.

2. Travaux connexes

Les travaux récents sur la prévision du succès des films reposent majoritairement sur des formulations en classification, qu'elle soit binaire ou multiclasse, comme dans Al Fahoum et Ghobon (2023), Kumar et Pande (2024), de Souza et al. (2023), Sahu et al. (2023), Shahid et Islam (2023), Choi et Lee (2022) et Sharma et al. (2021). Ces approches visent principalement à répartir les films en classes de succès, plutôt qu'à prédire directement une variable continue comme le nombre d'entrées. Sur le plan méthodologique, les modèles les plus mobilisés sont surtout des méthodes d'ensemble, telles que Random Forest, XGBoost, Gradient Boosting, AdaBoost, le vote ou le stacking, parfois comparées à des réseaux de neurones ou à des modèles plus classiques. Des architectures plus complexes, comme les Transformers, apparaissent plus récemment, notamment dans un cadre de régression (Chao et al., 2023). Toutefois, même lorsque la régression est envisagée, comme dans Rahul et al. (2023) ou Chao et al. (2023), la prévision directe des entrées demeure moins fréquente que les approches de classification. L'explicabilité reste par ailleurs peu intégrée dans ces travaux. La plupart des études n'emploient pas d'outils dédiés, comme c'est le cas chez Al Fahoum et Ghobon (2023), Sahu et al. (2023), Rahul et al. (2023), Chao et al. (2023) ou Sharma et al. (2021). Quelques exceptions existent néanmoins, avec l'usage de SHAP et LIME chez Kumar et Pande (2024), ou de l'importance des variables chez de Souza et al. (2023), Shahid et Islam (2023) et Choi et Lee (2022).

3. Notre solution proposée : MovieForecast

Notre approche repose sur un pipeline de prévision utilisant des données Vertigo et Allociné, combinant 100 variables réparties en trois familles : variables film, variables marketing et variables issues d'études spectateurs. Ces variables couvrent notamment le réalisateur, le distributeur, les acteurs principaux, la typologie du film, le nombre d'écrans, ainsi que des indicateurs d'opinion et de ciblage comme la recommandation, le désir d'aller voir le film, la notoriété ou certains profils sociodémographiques. Afin de reproduire un cas d'usage réaliste, nous adoptons un split chronologique : les modèles sont entraînés sur des films plus anciens, puis évalués sur des sorties plus récentes. Certaines variables qualitatives, comme les acteurs, réalisateurs et distributeurs, sont encodées à partir d'historiques d'entrées résumés par des

statistiques descriptives, tandis que d'autres variables catégorielles sont traitées par one-hot encoding. La tâche est formulée comme un problème de régression, et plusieurs modèles sont comparés : Random Forest, XGBoost, LightGBM, MLP, Bayesian Ridge et régression linéaire. Les prévisions sont ensuite confrontées à celles de trois analystes, experts du marché cinéma chez Vertigo produisant leurs prévisions à partir de leur connaissance du marché et d'outils statistiques. Enfin, l'explicabilité est étudiée à l'aide de méthodes spécifiques aux modèles et de méthodes agnostiques, comme SHAP et LIME.

4. Expérimentations

Nous analysons trois points : la performance prédictive des modèles, la comparaison aux prévisions d'experts et l'explicabilité. Les résultats obtenus sur les jeux d'entraînement et de validation sont présentés dans le Tableau 1. Globalement, les méthodes d'ensemble obtiennent les meilleures performances, avec Random Forest en tête sur l'entraînement et LightGBM sur la validation, ce qui confirme la pertinence du pipeline proposé pour la prévision des entrées. La comparaison avec les prévisions des trois experts sur le jeu de test montre par ailleurs que Random Forest atteint un niveau proche de l'expertise métier, tandis que LightGBM obtient la meilleure performance globale. En termes de RMSE, ce dernier améliore la prédiction d'environ 8 % par rapport aux experts, ce qui montre que les modèles de machine learning peuvent constituer un appui crédible à la décision pour les professionnels du secteur.

Tableau 1. Scores RMSE ($\times 10^3$) des modèles

	Random Forest	XGBoost	LightGBM	Multi-Layer Perceptron	Bayesian Ridge	Régression Linéaire
Train	124	149	175	364	346	257
Val	686	682	678	766	777	683

Enfin, l'analyse d'explicabilité met en avant des facteurs récurrents, en particulier le Notorious score, le nombre d'écrans et plusieurs variables issues des enquêtes spectateurs. Ces résultats soulignent l'intérêt de combiner variables intrinsèques, marketing et données d'opinion pour améliorer la prévision des entrées.

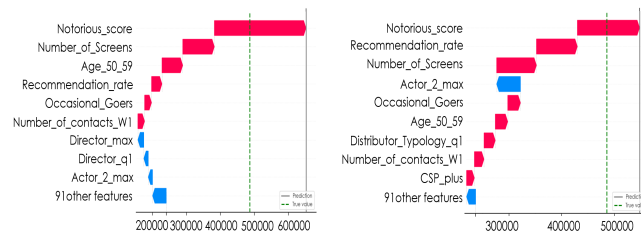


Figure 1. Impact des variables pour un film ayant réalisé plus de 500 000 entrées, selon les valeurs SHAP, pour Random Forest (gauche) et LGBM (droite).

5. Conclusion & travaux futurs

Nous avons étudié une approche de machine learning pour prédire les ventes de tickets au cinéma. Nos résultats montrent que les données d'études apportent une valeur significative à la prévision. Les meilleures performances sont obtenues en combinant des variables marketing, liées au film et d'opinion du public, ce qui permet d'atteindre un niveau comparable à celui des experts du secteur, voire légèrement supérieur, sur le jeu de test. L'analyse d'explicabilité met en évidence le rôle central d'une variable combinant la notoriété marketing et l'intention de voir le film. En perspective, nous cherchons à intégrer l'environnement de sortie afin de mieux capturer les dynamiques de concurrence entre films, notamment les phénomènes de cannibalisation, avec comme piste l'usage d'outils issus de la théorie des jeux.

Remerciements

Cette recherche a été soutenue par l'entreprise Vertigo. Nous remercions l'équipe Vertigo pour l'accès aux données d'enquête et ses précieux éclairages sur l'industrie du cinéma.

Bibliographie

- Al Fahoum A., Ghobon T. A. (2023). Performance Predictions of Sci-Fi Films via Machine Learning. *Applied Sciences*, vol. 13, n° 7, article 4312.
- Chao Q., Kim E., Li B. (2023). Movie Box Office Prediction with Self-Supervised and Visually Grounded Pretraining. *arXiv*, abs/2304.10311.
- Choi J. Y., Lee G. H. (2022). Comparative Analysis of Machine Learning Models for Prediction of Movie's Box Office Success. *ICIC Express Letters, Part B: Applications*, vol. 13, n° 11, p. 1207-1214.
- de Souza T. L. D. E., Nishijima M., Pires R. (2023). Revisiting predictions of movie economic success: Random Forest applied to profits. *Multimedia Tools and Applications*, vol. 82, n° 25, p. 38397-38420.
- Kumar S. B., Pande S. D. (2024). Explainable Neural Network analysis on Movie Success Prediction. *EAI Endorsed Transactions on Scalable Information Systems*, vol. 11, n° 4.
- Rahul V. S., Tejas M., Prasanth N. N., Raja S. P. (2023). Early Success Prediction of Indian Movies Using Subtitles: A Document Vector Approach. *International Journal of Image and Graphics*, vol. 23, n° 4, article 2350030.
- Sahu S., Kumar R., Long H. V., Shafi P. M. (2023). Early-production stage prediction of movies success using K-fold hybrid deep ensemble learning model. *Multimedia Tools and Applications*, vol. 82, n° 3, p. 4031-4061.
- Shahid M. H., Islam M. A. (2023). Investigation of time series-based genre popularity features for box office success prediction. *PeerJ Computer Science*, vol. 9, article e1603.
- Sharma A. S., Roy T., Rifat S. A., Mridul M. A. (2021). Presenting a Larger Up-to-Date Movie Dataset and Investigating the Effects of Pre-Released Attributes on Gross Revenue. *Journal of Computer Science*, vol. 17, n° 10, p. 870-888.

From Digital Traces to Shared Understanding: An AI-based approach to group awareness in distributed teams

Kaoutar Akhsass^{1,2}

Supervisors: Manuele Kirsch-Pinheiro¹, Luiz Angelo Steffene²

¹ *Centre de Recherche en Informatique (CRI), Université Paris 1 Panthéon-Sorbonne, Paris, France*

kaoutar.akhsass@univ-paris1.fr — manuele.kirsch-pinheiro@univ-paris1.fr

² *LICIIS, Université de Reims Champagne Ardenne, Reims, France*

luiz-angelo.steffene@univ-reims.fr

RESUME. La généralisation du travail à distance et du travail hybride rend de plus en plus difficile, pour les équipes distribuées, le maintien d'une compréhension partagée de l'activité collective dans des environnements fragmentés entre multiples plateformes. Cet article présente un projet de thèse situé à l'intersection du Travail Coopératif Assisté par Ordinateur (TCAO/CSCW), de la conscience de groupe (group awareness) et de l'Intelligence Artificielle, qui vise à concevoir un système basé sur l'IA capable d'identifier, de structurer et de personnaliser les informations pertinentes pour la coordination à partir de traces numériques hétérogènes issues des environnements de travail contemporains.

ABSTRACT. The widespread adoption of remote and hybrid work has made it increasingly difficult for distributed teams to maintain a shared understanding of collective activity across fragmented multi-platform environments. This paper presents a thesis project situated at the intersection of Computer-Supported Cooperative Work (CSCW), group awareness, and Artificial Intelligence, which aims to design an AI-enabled system capable of identifying, structuring, and personalizing coordination-relevant information from heterogeneous workplace traces.

MOTS-CLÉS : conscience de groupe, CSCW, télétravail, IA générative, traces numériques, environnements multiplateformes.

KEYWORDS: Group awareness, CSCW, remote work, generative AI, digital traces, multi-platform environments

1. Introduction

The widespread adoption of hybrid and remote work has transformed collaborative practices and increased information fragmentation across platforms. This transformation mostly occurred as an emergency response rather than a planned shift (Wang et al., 2020). In France, the share of employees working remotely has grown from 9% in 2019 to 26% in 2023 (Beatriz and Erb, 2024), while 52.9% of European companies reported using remote meeting tools in 2024 (Eurostat, 2025). This dispersion generates cognitive overload and technostress (Jackson et al., 2022). It directly undermines group awareness, i.e., the shared understanding team members have of one another's activities, availability, and interdependencies (Dourish and Bellotti, 1992). While CSCW has produced strong conceptual and technical foundations for supporting awareness, existing approaches remain only partially adapted to the multi-platform, asynchronous, and highly fragmented environments that characterize contemporary distributed work.

Situated at the intersection of CSCW and AI, this thesis project proposes to design a system capable of identifying, structuring, and personalizing coordination-relevant information from heterogeneous digital traces found in contemporary work environments. It is guided by the following research question: *How can Artificial Intelligence support group awareness in multi-platform remote work environments by identifying, structuring, and personalizing coordination-relevant information, while respecting privacy and regulatory constraints?*

2. Scientific Background and Related Work

CSCW has established group awareness as a central coordination challenge in collaborative environments. Dourish and Bellotti (1992) define awareness as the understanding that individuals have of their group, their colleagues, and their activities providing context for aligning individual contributions. In distributed settings, this awareness no longer arises naturally and must be actively supported through technological mechanisms embedded in groupware systems. Prior research has demonstrated both the value and the limits of awareness mechanisms in collaborative environments (Kirsch-Pinheiro et al., 2003; 2004).

At the same time, contemporary collaborative work rarely unfolds on a single platform. Recent studies show that distributed teams work across multiple tools and collaboration hubs, generating information silos and overload that impaired continuity in shared understanding (Jackson et al., 2022). Group awareness is therefore no longer simply a problem about displaying cues within a single shared interface, but about identifying and articulating coordination-relevant information across fragmented multi-platform environments (Lee et al., 2022).

Recent AI advances open new perspectives for this challenge. Process mining (van der Aalst, 2012) and context mining (Rabah et al., 2020) enable the extraction of meaningful coordination patterns from activity traces, while Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) (Lewis et al., 2020) enable

contextual summarization and grounded access to heterogeneous distributed traces. Together, these technologies point toward interactive, adaptive awareness support systems that go beyond static dashboards.

Existing awareness support, whether based on static dashboards in groupware platforms (Gutwin and Greenberg, 2002) or on AI-based features such as meeting recaps (Asthana et al., 2025), remains tied to a single artefact and overlooks the multi-platform integration and role-based personalization required by distributed work.

Within CSCW, supporting group awareness across heterogeneous, multi-platform data remains an open challenge. The proposed system should suggest the coordination cues that best match each user’s role and context, raising issues of trace integration, activity log processing, decentralization, role-based personalization, and compliance with the GDPR and the AI Act.

3. Thesis Project

The thesis assumes that an AI-enabled system can identify, structure, and personalize coordination-relevant information from heterogeneous digital traces in distributed work environments. It seeks to determine what information best supports group awareness in multi-platform settings, how it can be adapted to collaborators’ roles and context, and under which conditions such mechanisms remain compatible with privacy and regulatory requirements.

To achieve this objective, the project follows an iterative methodology combining conceptual framing, data analysis, model design, and evaluation. It will first identify the forms of information that best support awareness in multi-platform environments, then move to collecting and contextualizing digital traces, before exploring how generative AI can select, summarize, and personalize this information for each collaborator. The final stages will involve user-centred studies in realistic settings to assess the relevance, usefulness, and acceptability of the proposed approach.

The technical building blocks listed above form a complete recommendation task pipeline. Process and context mining structure platform logs, LLMs generate role and context aware summaries from unstructured data, and RAG grounds the generated recommendations in retrieved traces to ensure traceability and reduce hallucinations.

The research will first rely on publicly available datasets, such as MeetingBank (Hu et al., 2023), to prototype and evaluate the early stages of the approach under controlled conditions, then progressively collect real collaborative traces, including meeting recordings, chat exchanges, and platform event logs.

4. Conclusion

This paper presents the foundations and main directions of a PhD thesis. It aims to design an AI-based awareness support system that identifies, structures, and personalizes coordination-relevant information from heterogeneous workplace traces. The literature review shows that while CSCW offers strong bases for awareness

support, existing approaches remain only partially adapted to contemporary distributed work. Recent advances in generative AI, particularly LLMs and RAG, open new perspectives for synthesizing distributed traces and supporting contextually relevant awareness, while also raising challenges regarding privacy, selectivity, and personalization. Next steps will focus on building an experimental corpus, from public datasets to more contextualized collaborative traces, and on designing and evaluating AI-based mechanisms for awareness support in realistic multi-platform environments.

Bibliography

- Asthana S., Hilleli S., He P. and Halfaker A. (2025). Summaries, Highlights, and Action Items: Design, Implementation and Evaluation of an LLM-powered Meeting Recap System. *Proceedings of the ACM on Human-Computer Interaction*, vol. 9, pp. 1-29.
- Beatriz M. and Erb L.-A. (2024). Comment évolue la pratique du télétravail depuis la crise sanitaire ? *DARES*. [Online]. Available: <https://dares.travail-emploi.gouv.fr>
- Dourish P. and Bellotti V. (1992). Awareness and Coordination in Shared Workspaces. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW '92)*, Toronto, Canada, pp. 107-114.
- Eurostat (2025). 53% of EU enterprises held remote meetings in 2024. [Online]. Available: <https://ec.europa.eu/eurostat>
- Gutwin C. and Greenberg S. (2002). A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW)*, vol. 11, n° 3-4, pp. 411-446.
- Hu Y. et al. (2023). MeetingBank: A Benchmark Dataset for Meeting Summarization. *Proceedings of ACL 2023*, pp. 16409-16423.
- Jackson V., Van der Hoek A. and Prikładnicki R. (2022). Collaboration Tool Choices and Use in Remote Software Teams. *CHASE 2022 / ACM*, pp. 76-80.
- Kirsch-Pinheiro M., De Lima J. V. and Borges M. (2003). A framework for awareness support in groupware systems. *Computers in Industry*, vol. 52, pp. 47-57.
- Kirsch-Pinheiro M., Gensel J. and Martin H. (2004). Representing context for an adaptive awareness mechanism. *International Conference on Collaboration and Technology*, Springer, Berlin, pp. 339-348.
- Lee C. Y. P., Zhang Z. and Herskovitz J. (2022). CollabAlly: Accessible Collaboration Awareness in Document Editing. *Proceedings of ACM CHI 2022*, pp. 1-17.
- Lewis P., Perez E. and Piktus A. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Adv. in Neural Information Proc. Systems*, vol. 33, pp. 9459-9474.
- Rabah N. B., Kirsch-Pinheiro M. and Le Grand B. (2020). Machine Learning for a Context Mining Facility. *IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 1-7.
- van der Aalst W. (2012). Process Mining: Overview and Opportunities. *ACM Transactions on Management Information Systems*, vol. 3, pp. 1-17.
- Wang B., Liu Y., Qian J. and Parker S. K. (2020). Achieving effective remote working during the COVID-19 pandemic: A work design perspective. *Applied Psychology*

Intelligence Artificielle Explicable pour la détection de cyberattaques et la reconnaissance de plans via des règles probabilistes entraînées de manière adverse

Antoine Leblanc¹

Encadrants : Bénédicte Le Grand¹, Jacques Robin², Nourhène Ben Rabah¹

1. Centre de Recherche en Informatique (CRI), Université Paris 1 Panthéon-Sorbonne

90 rue de Tolbiac, 75013 Paris, France

antoine.leblanc@univ-paris1.fr, benedicte.le-grand@univ-paris1.fr, nourhene.ben-rabah@univ-paris.fr

2. esieaLab, École Supérieure d'Informatique Electronique Automatique (ESIEA)

4 All. Katherine Johnson, 94200 Ivry-sur-Seine, France

jacques.robin@esiea.fr

RESUME. Face à l'augmentation et à la complexité croissante des cyberattaques, les systèmes de détection actuels peinent à fournir des explications exploitables aux analystes. Ce travail de thèse propose une approche d'IA explicable combinant une ontologie de référence en cybersécurité, nommée ROCADe, et un modèle de raisonnement basé sur la Programmation Logique Probabiliste sous Contraintes (CPLP). L'objectif est de corréliser des alertes hétérogènes pour reconnaître non seulement des actions isolées, mais des plans d'attaque complets, tout en offrant une transparence sur le processus de décision du modèle.

ABSTRACT. Faced with the increase and growing complexity of cyberattacks, current detection systems struggle to provide analysts with actionable explanations. This thesis proposes an explainable AI approach combining a cybersecurity reference ontology, called ROCADe, and a reasoning model based on Constraint-Based Probabilistic Logic Programming (CPLP). The goal is to correlate heterogeneous alerts to recognize not only isolated actions, but entire attack plans, while providing transparency into the model's decision-making process.

MOTS-CLES : IA Explicable, Programmation Logique Probabiliste, Cybersécurité, Ontologie de Référence, Détection d'Attaques, Reconnaissance de Plans

KEYWORDS: Explainable AI, Probabilistic Logic Programming, Cybersecurity, Reference Ontology, Attack Detection, Plan Recognition

1. Introduction

Le paysage de la cybersécurité est marqué par une explosion du nombre de rançongiciels (+175 % entre 2023 et 2024 (Microsoft, 2024)) et l'émergence d'attaques structurées de type APT (Advanced Persistent Threat). Les systèmes de détection et de réponse génèrent une quantité massive d'alertes souvent déconnectées, rendant la tâche des analystes des centres opérationnels de sécurité (SOC) ardue.

Ma thèse, débutée en décembre 2024 dans le cadre du projet ANR ANCILE, s'inscrit dans ce contexte. Elle vise à développer des modèles capables de détecter des techniques offensives et de reconnaître des plans d'attaque de manière explicable, en utilisant des règles probabilistes entraînées de façon adverse.

2. Problématique et Questions de Recherche

Les approches d'apprentissage profond appliquées à la détection d'intrusions (Apruzzese et al., 2023) atteignent de bonnes performances mais souffrent d'un déficit d'explicabilité reconnu (Rudin, 2019), particulièrement problématique en contexte SOC où l'analyste doit justifier chaque décision. À l'inverse, les systèmes à base de règles sont transparents mais peinent à gérer l'incertitude et la variabilité des attaques réelles. Mes travaux s'articulent autour de deux questions de recherche (QR) :

– **QR1 (terminée)** : Quelle ontologie utiliser pour représenter uniformément des modèles probabilistes explicables d'actions et de plans de cyberattaques ?

– **QR2 (en cours)** : Quelle est l'efficacité de la Programmation Logique Probabiliste sous Contraintes (C) pour la détection d'attaques et la reconnaissance de plans, et comment ses explications se comparent-elles à celles des modèles d'apprentissage automatique réputés interprétables ?

3. Etat de l'art et Contribution Ontologique : ROCADE

3.1. Evaluation et Crédibilité des Ontologies de Cybersécurité

Une étape préalable a consisté à catégoriser les ontologies de cybersécurité existantes. L'application du *Framework for Ontologies Classification* (F4OC) (Martins et al., 2022) sur un corpus de 87 ontologies a révélé qu'il néglige une dimension essentielle : la crédibilité. F4OC valorise les ontologies rigoureuses sémantiquement (ex. COVER, ROSE (Oliveira et al., 2022)) au détriment de standards industriels (MITRE ATT&CK, STIX) qui bénéficient d'un fort soutien institutionnel et d'une adoption opérationnelle. Nous avons donc proposé un critère de crédibilité décomposé en quatre dimensions — soutien académique, adoption industrielle, soutien institutionnel et validation par les praticiens (Leblanc et al., 2026).

3.2. Positionnement par rapport à l'état de l'art

Les plateformes SOAR commerciales automatisent la réponse aux incidents via des *playbooks* codés en dur, sans modèle sémantique du raisonnement adverse (Islam et al., 2019). Côté ontologique, UCO (Syed et al., 2016) et WAVED (Akbar et al., 2023) intègrent les standards de facto (ATT&CK, CVE, CWE, STIX) mais n'offrent pas d'ancrage fondationnel, ce qui limite leur cohérence et leur extensibilité. À notre connaissance, aucune ontologie ne satisfait simultanément ancrage fondationnel, intégration des standards industriels et validation par les praticiens (Leblanc et al., 2026).

3.3. L'Ontologie ROCADe

Pour combler ce manque, nous co-construisons ROCADe (*Reference Ontology for Cybersecurity Attack and Defense*) avec des experts en modélisation conceptuelle (Université de Twente), en standardisation (LIST) et des analystes SOC (Gatewatcher). ROCADe est ancrée sur UFO et spécialise COVER et ROSE, tout en intégrant les standards MITRE ATT&CK, D3FEND, CAPEC et CVE/CWE. Notre contribution propre porte sur la formalisation des plans d'attaque et de défenses comme *Closed Intentions* UFO et sur la modélisation des dépendances historiques entre tactiques ATT&CK et méta-tactiques de la *Unified Kill Chain* (Pols & Dominguez, 2017). ROCADe est validée, via une implémentation opérationnelle en Prolog, par conformité aux standards ISO/IEC 27005:2022 et NIST CSF 2.0, et par revue d'un panel d'analystes SOC.

4. Raisonnement et Détection par Programmation Logique Probabiliste sous Contraintes (CPLP)

La Programmation Logique Probabiliste sous Contraintes (CPLP) (De Raedt & Kimmig, 2015) offre trois avantages pour notre cas d'usage. (1) **Explicabilité native** : chaque détection correspond à un prédicat logique prouvé ; la chaîne de règles activées lors de cette preuve constitue une explication directement lisible par l'analyste. (2) **Gestion de l'incertitude** via les probabilités, qui modélisent à la fois le succès incertain d'une étape d'attaque et le bruit des capteurs. (3) **Gestion de contraintes** temporelles et structurelles, intégrées au sein même du raisonnement.

Notre méthodologie projette les événements hétérogènes (Kubernetes, SDN, logs système) sur l'ontologie ROCADe, les transformant en faits logiques exploitables par le moteur CPLP. La reconnaissance de plans d'attaque identifie alors des signatures comportementales caractéristiques des APT et permet d'inférer le stade d'une attaque en cours, offrant à l'analyste un temps d'avance pour une remédiation proactive.

5. Conclusion et Perspectives

Mes prochains travaux porteront sur l'implémentation du moteur CPLP et son évaluation sur des jeux de données d'attaques complexes. Trois axes structureront cette phase : (i) la *complexité* computationnelle de l'inférence CPLP; (ii) la *comparaison* empirique avec des modèles d'apprentissage automatique interprétables et des méthodes post-hoc; (iii) l'*évaluation* qualitative des explications auprès des analystes SOC du projet ANCILE.

Remerciements: Ce travail a été financé par l'Agence nationale de la recherche (ANR) dans le cadre du projet ANCILE (subvention n° ANR-23-CE39-0010).

Bibliographie non numérotée

- Akbar, K. A. *et al.* (2023). The Design and Application of a Unified Ontology for Cyber Security. Int. Conf. on Information Systems Security (ICISS), p. 23-41. https://doi.org/10.1007/978-3-031-49099-6_2
- Apruzzese, G. *et al.* (2023). The Role of Machine Learning in Cybersecurity. ACM Digital Threats: Research and Practice, vol. 4, n° 1, p. 1-38. <https://doi.org/10.1145/3545574>
- De Raedt, L., and Kimmig, A. (2015). Probabilistic (logic) programming concepts. Machine Learning, vol. 100, n° 1, p. 5-47. <https://doi.org/10.1007/s10994-015-5494-z>
- Islam, C. *et al.* (2019). A Multi-Vocal Review of Security Orchestration. ACM Computing Surveys, vol. 52, n° 2, p. 1-45. <https://doi.org/10.1145/3305268>
- Leblanc, A. *et al.* (2026). Rethinking Cybersecurity Ontology Classification and Evaluation: Towards a Credibility-Centered Framework. Enterprise Design, Operations, and Computing (EDOC 2025), Springer LNCS.
- Martins, B. F. *et al.* (2022). A framework for conceptual characterization of ontologies and its application in the cybersecurity domain. Software and Systems Modeling, vol. 22, n° 1, p. 1-27. <https://doi.org/10.1007/s10270-022-01055-4>
- Microsoft (2024). Microsoft Digital Defense Report 2024. <https://aka.ms/MDDR2024>
- Oliveira, Í. *et al.* (2022). An Ontology of Security from a Risk-Treatment Perspective. Conceptual Modeling (ER 2022), Springer LNCS, p. 365-379. https://doi.org/10.1007/978-3-031-17995-2_26
- Pols, P. and Domínguez, F. (2017). The Unified Kill Chain: Designing a Unified Kill Chain for analyzing, comparing and defending against cyber attacks. Rapport technique, Fox-IT, HAN Univ., Pays-Bas.
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, vol. 1, n° 5, p. 206-215. <https://doi.org/10.1038/s42256-019-0048-x>
- Syed, Z. *et al.* (2016). UCO: A Unified Cybersecurity Ontology. AAAI Workshop on Artificial Intelligence for Cyber Security (AICS).

Détection et qualification des dégradations émotionnelles et relationnelles au sein d'un groupe familial dans un contexte d'aidance

Paul CARIOU¹

Encadrante : Manuele KIRSCH PINHEIRO¹

*1 Centre de Recherche en Informatique, Université Paris 1 Panthéon Sorbonne
90 rue de Tolbiac, 75013, Paris, France
paul@tamalou.io, manuele.kirsch-pinheiro@univ-paris1.fr*

RESUME. Dans un contexte de vieillissement de la population, les proches aidants jouent un rôle crucial mais sont confrontés à l'épuisement et à l'isolement. Les outils de santé actuels (DxCare, Epic, DMP) se concentrent sur les données médicales formelles, laissant de côté l'état émotionnel des proches. Or les échanges informels regorgent de données cliniques non captées. On peut se demander comment les outils d'IA peuvent améliorer l'analyse des échanges au sein des groupes d'aidants pour réduire leur charge et améliorer le suivi des personnes fragiles.

MOTS-CLES : Grands modèles de langage (LLM), intelligence artificielle médicale, structuration des données, coordination familiale, données cliniques, traitement automatique du langage naturel (NLP), signaux faibles, analyse des sentiments.

1. Introduction

Dans un contexte d'aidance, les environnements collaboratifs digitaux, bien qu'ils facilitent la communication, introduisent de nouvelles vulnérabilités : manque de visibilité sur les intentions des membres, rigidité des rôles, surcharge informationnelle, et difficulté à capter les signaux faibles qui émergent des interactions quotidiennes. Les objets techniques définissent des rôles, distribuent des compétences et naturalisent des relations sociales, ce qui peut accentuer les déséquilibres au sein des groupes (Akrich, 1992). Dans la santé, la charge administrative repose de plus en plus largement sur l'aidant principal, généralement la femme, la sœur ou la fille restée à proximité géographique du patient (Gösenbauer et al., 2025), accentuant les risques d'épuisement et les effets négatifs sur la santé du patient comme de l'aidant.

Les outils collaboratifs enrichis par l'Intelligence Artificielle (IA) offrent des perspectives inédites pour répondre à ces défis. Les *Large Language Models* (LLM) et les techniques d'analyse des sentiments permettent d'extraire et de structurer les informations implicites issues des conversations, comme les émotions, les tensions ou

les besoins non exprimés. L'objectif de cette thèse est de définir des marqueurs de dégradation émotionnelle et relationnelle dans les échanges numériques, de développer des algorithmes capables de les identifier et ainsi de permettre une détection précoce des situations de surcharge.

2. État de l'art

Des études récentes révèlent qu'une majorité d'aidants ressentent une détresse psychologique significative, aggravée par la dispersion géographique, la charge émotionnelle et la complexité des dynamiques relationnelles intrafamiliales (Smriti et al., 2024). Dans ce contexte de dématérialisation, les groupes familiaux sont confrontés à la nécessité de maintenir une conscience collective, ou *groupe awareness*, définie comme la capacité des membres à percevoir et comprendre les activités, intentions et états émotionnels des autres participants par (Dourish et Bellotti, 1992). Ce concept, central dans le domaine du *Computer-Supported Cooperative Work* (CSCW), permet d'assurer une coordination efficace, d'éviter les duplications de travail et de favoriser une adaptation spontanée aux imprévus.

L'analyse des sentiments constitue un domaine clé du traitement automatique du langage naturel (NLP), largement abordé dans la littérature, et particulièrement utile pour extraire des informations subjectives à partir de textes informels (Al-Qabla et al., 2023; Nijhawan et al., 2022). Dans le contexte de soins, les ressentis du patient et de ses proches peuvent fournir des indications précieuses sur leurs conditions de vie. Néanmoins, les émotions sont des processus dynamiques, façonnés par le temps, les interactions et la régulation (Kuppens et al., 2017). Une étude fine et longitudinale est nécessaire pour évaluer l'état psychologique.

Dans le secteur du soin, le NLP, les systèmes multimodaux et les algorithmes d'IA ont grandement contribué à diagnostiquer et surveiller les troubles mentaux (Muetunda et al., 2024). Cependant, la formalisation des signaux faibles reste un enjeu majeur : les données numérisées restent limitées (Murtha et al., 2023 ; Shickel et al., 2021) et les dossiers médicaux structurés se concentrent sur des événements aigus, sans intégrer les informations du quotidien comme la fatigue, le sommeil ou l'humeur (Ryu et al., 2024 ; Schiltz et al., 2020). Des patients rapportent parfois une sévérité symptomatique supérieure à ce que leurs cliniciens documentent (Trojan et al., 2021), et des recommandations institutionnelles appellent désormais à intégrer ces données rapportées dans le suivi courant (Lordick et al., 2022).

3. Extraction et structuration des implicites conversationnels

Prenons l'exemple d'une famille coordonnant le maintien à domicile d'une personne âgée atteinte de démence légère. Ses trois enfants échangent quotidiennement sur une application conversationnelle. Sur plusieurs semaines, on observe des marqueurs progressifs : formulations de plus en plus courtes et factuelles de l'aidant principal ("elle a encore refusé de manger", "je gère"), diminution de la fréquence des messages des autres membres, expressions d'irritation contenue ("comme d'habitude c'est moi qui m'en occupe") et disparition progressive des

marques d'humour ou d'encouragement mutuel. Ces marqueurs lexicaux, syntaxiques, contextuels et relationnels constituent les unités d'observation de cette recherche.

Pris isolément, chacun de ces signaux est ambigu. C'est leur accumulation dans le temps qui forme une trajectoire émotionnelle : une modélisation longitudinale de l'état émotionnel et relationnel d'un individu ou d'un groupe, reconstituée à partir de la sédimentation de ces marqueurs sur plusieurs semaines. Dans notre exemple, la surcharge croissante de l'aidant principal combinée à la diminution de l'implication des membres périphériques signalent une situation à risque.

C'est cette convergence d'éléments gradués et contextualisés qui déclenche une alerte sous la forme par exemple d'une notification à un professionnel de santé signalant des marqueurs d'épuisement émotionnel sur le dernier mois. La conception d'un tel système suppose de résoudre trois verrous : la définition et la taxonomie des marqueurs (V1), leur détection automatique dans un corpus en français informel (V2), et leur agrégation en trajectoires temporelles cohérentes (V3). L'approche est interdisciplinaire, associant aidants, professionnels de santé et experts en IA, CSCW et sociologie, afin de développer des outils respectueux des contraintes éthiques des groupes.

Ce travail s'inscrit dans une collaboration avec Tamalou, application gratuite de coordination des soins (messagerie, calendrier partagé, médiathèque), conçue pour réduire la charge mentale de l'aidant principal. Nous explorons un LLM basé sur GPT, affiné par l'intégration de l'analyse des sentiments et de règles grammaticales à l'instar de MoLeSy (Zhang et al., 2022), afin d'identifier les signaux faibles dans les échanges et d'ouvrir la perspective d'une génération de notes structurées compatibles avec les DMP.

4. Discussions et perspectives

L'objectif majeur est de créer un pont entre le savoir familial et l'expertise médicale : meilleure qualité de vie pour l'aidant et le patient, optimisation du suivi et réduction des coûts pour le système de santé. La contribution centrale est de nature sémiologique et méthodologique : comment passe-t-on d'une interaction conversationnelle informelle à une donnée cliniquement interprétable ? Les LLM et l'analyse des sentiments servent d'outils opérationnels au service de cette formalisation.

Plusieurs limites et considérations éthiques sont à prendre en compte. Premièrement, la question des biais des données, due au nombre limité de familles volontaires, pourrait compromettre la généralisation de la démarche. Un nombre élevé de faux positifs peut également conduire à une mauvaise hiérarchisation des alertes et à une perte de crédibilité. Deuxièmement, la méfiance des familles envers le risque de surveillance est un risque majeur que nous comptons contrer par une bonne communication et le respect des normes (RGPD et IA Act).

Bibliographic

- Al-Qablan, T.A., Mohd Noor, M.H., Al-Betar, M.A., Khader, A.T. (2023). A survey on sentiment analysis and its applications”. *Neural Computing and Applications*, vol. 35, p. 21567–21601. <https://doi.org/10.1007/s00521-023-08941-y>.
- Akrich, M. (1992). The Description of Technical Objects. Bijker, W. & Law, J. Shaping Technology/Building Society. *Studies in Sociotechnical Change*, MIT Press, p.205-224.
- Dourish, P., Bellotti. V. (1992). « Awareness and Coordination in SharedWorkspaces ». In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work - CSCW'92*, 10714. Toronto, Ontario, Canada: ACM Press.
- Gösenbauer, B., Braun, A., Bilger, M. (2025). « Informal long-term care and its gender distribution: a systematic review ». *Journal of Integrated Care*, vol. 33, n° 5, p. 17–35.
- Kuppens P, Verduyn P. (2017). « Emotion dynamics ». *Curr Opin Psychol*. 2017 Oct;17:22-26. 10.1016/j.copsyc.2017.06.004. Epub 2017 Jun 21. PMID: 28950968.
- Lordick, F., Belka, C., Cordes, N., De Ruyscher, D., Guckenberger, M., Hurkmans, C., Niyazi, M., Petersen, C., Ricardi, U., Valentini, V. (2022). « The role of patient-reported outcome measures in the continuum of cancer clinical care: ESMO Clinical Practice Guideline ». *Annals of Oncology*, vol. 33, n° 9, p. 878–892.
- Muetunda, F., Sabry, S., Jamil, M.L., Pais, S., Dias G., Cordeiro, J. (2024). « AI-Assisted Diagnosing, Monitoring and Treatment of Mental Disorders: A Survey ». *ACM Transactions on Computing for Healthcare*, Vol. 5, n° 4, Article No. 23, p. 1 - 24.
- Murtha, J.A., Jawara, D., Funk, L.M. (2023). « Optimization of the Electronic Health Record for Research ». *Annals of Surgery Open*, vol. 4, n° 2. <https://doi.org/10.1097/AS9.0000000000000297>.
- Nijhawan, T., Attigeri, G., Ananthakrishna, T. (2022). « Stress detection using natural language processing and machine learning over social interactions ». *Journal of Big Data*, vol. 9, article n° 33. <https://doi.org/10.1186/s40537-022-00575-6>.
- Ryu, S., Park, S., Kim, M. (2024). « Challenges in and Opportunities for Electronic Health Record-Based Data Analysis and Interpretation ». *Gut and Liver*, vol. 18, n° 2, p. 204–210. <https://doi.org/10.5009/gnl230272>.
- Schiltz, N., Foradori, M., Reimer, A., Dolansky, M. (2020). « Functional Measures Are Severely Under-Captured in Electronic Health Records ». *Innovation in Aging*, vol. 4, Supplement 1, p. 925. <https://doi.org/10.1093/geroni/igaa057.925>.
- Shickel, B., Loftus, T., Ozrazgat-Baslanti, T., Rashidi, P., Bihorac, A. (2021). « Why Is the Electronic Health Record So Challenging for Research and Clinical Care? ». *Methods of Information in Medicine*, vol. 60, n° 1-02, p. 32–40.
- Smriti, Diva, Lu Wang, et Jina Huh-Yoo. « Emotion Work in Caregiving: The Role of Technology to Support Informal Caregivers of Persons Living With Dementia ». *Proceedings of the ACM on Human-Computer Interaction*, vol o CSCW1 (17 avril 2024), Article 48, pp. 48:1-48:34. <https://doi.org/10.1145/3637325>.
- Trojan, A., Leuthold, N., Thomssen, C., Rody, A., Winder, T., Jakob, A., Egger, C., Held, U., Jackisch, C. (2021). « The Effect of Collaborative Reviews of Electronic Patient-Reported Outcomes on the Congruence of Patient- and Clinician-Reported Toxicity in Cancer Patients Receiving Systemic Therapy ». *JMIR Cancer*, vol. 7, n° 3, e29271. <https://doi.org/10.2196/29271>.