

JOBIM 2024 - MINI-SYMPOSIUM

L'IA EN SCIENCES DU VIVANT - IMPACT SUR NOS METIERS

Jeudi 27 juin 2024, 14h - 16h

Organisé par Le CATI BIOSBiol :

Gabryelle Agoutin (INRAE, Toulouse), Lucas Auer (INRAE, Nancy), Philippe Bardou (INRAE, Toulouse), Cédric Cabau (INRAE, Toulouse), Patrice Déhais (INRAE Toulouse), Cervin Guyomar (INRAE, Toulouse), Sarah Maman (INRAE, Toulouse) et Géraldine Pascal (INRAE, Toulouse).



L'IA et plus particulièrement l'IA générative représente sans conteste un des événements majeurs qui bouleverse et bouleversera nos métiers. La généralisation de l'accès à une multitude d'outils tels que ChatGPT, Copilot ou Gemini modifie fondamentalement notre quotidien. Dans de nombreux aspects de nos activités, allant du développement informatique à l'aide à la rédaction, en passant par la bibliographie, ces outils deviennent des assistants incontournables. Cependant, ces changements ne sont pas exempts de défis. Des considérations éthiques et environnementales, la transparence des algorithmes, et la nécessité d'une collaboration étroite entre chercheurs et développeurs sont autant de points essentiels à aborder pour garantir une utilisation responsable et éthique de l'IA en recherche.

Sources illustration :

<https://www.numerama.com/tech/722953-un-intelligence-artificielle-qui-aide-a-coder-a-ete-developpee-par-github-et-openai.html>

L'informaticien se transforme en « détective de la tromperie ». - PublicDomainPictures / Pixabay / pixabay

Dans ce symposium, nous proposons deux temps, un premier qui visera à explorer le fonctionnement des LLM (Large Language Models), d'aborder les limites générales des approches de machine learning, de voir les usages et applications de ces LLMs mais également les risques. Dans un second temps, nous proposons de faire un focus sur l'utilisation des IA génératives dans le contexte de publication scientifique en exposant des cas de mauvais usages et des pratiques frauduleuses qui polluent la littérature scientifique.

PROGRAMME : L'IA EN SCIENCES DU VIVANT - IMPACT SUR NOS METIERS

14H : présentation du mini-symposium par les organisateurs.

14H05 – 14H50 : [VINCENT GUIGUE](#) (AgroParisTech Paris-Saclay), Professeur en informatique au sein d'AgroParisTech Paris-Saclay. Ses travaux, centrés sur les architectures d'apprentissage automatique, se structurent en trois axes : l'analyse de séries temporelles, depuis sa thèse ; le traitement automatique de la langue naturelle et les tâches d'extraction d'information ; l'apprentissage de profils pour les systèmes de recommandation. Depuis une vingtaine d'années, le machine learning est un domaine de recherche très dynamique. Mais la rapidité d'évolution des approches ces dix dernières années est stupéfiante : une telle rupture n'a probablement jamais été observée dans l'histoire des sciences. Ces avancées déferlent sur la société dans son ensemble et posent de nombreux défis. La [recherche](#) de Vincent Guigue est très appliquée et repose essentiellement sur l'apprentissage de représentation qui permet de décomposer et analyser différents types d'entrées, tout en imposant des contraintes métiers.

Vincent Guigue présentera :

- Le fonctionnement des LLM et les spécificités de chatGPT
- Les limites générales (techniques) des approches de machine learning
- Les usages et applications des LLM et les développements techniques autour des autres IA génératives
- Les risques engendrés

14H50 – 15H05 : questions réponses

15H05 – 15H50 : [GUILLAUME CABANAC](#) (IRIT, Université de Toulouse) est professeur d'informatique à l'Université Toulouse III – Paul Sabatier et titulaire d'une chaire de recherche fondamentale de l'Institut universitaire de France (IUF) intitulée « [dépollution de la littérature](#) »

[scientifique](#) ». Ses travaux visent à identifier des publications non fiables par fouille de textes, notamment au sein du projet ERC Synergy '[Nanobubbles](#)' questionnant le processus d'auto-correction de la science. Il développe le '[Problematic Paper Screener](#)' qui traque des publications non fiables, pourtant publiées et souvent vendues par les maisons d'édition de premier plan. Au quotidien, ses signalements d'articles problématiques, contenant notamment des expressions torturées, ont donné lieu à des centaines de rétractations. Sa recherche a été distinguée dans le palmarès "[Nature's 10](#)" présentant « dix personnes qui ont contribué à façonner la science en 2021 » selon la revue Nature.

Guillaume Cabanac présentera :

- Des cas de mésusages de la génération de textes et de données avec ChatGPT dans des articles scientifiques
 - Publication de données fictives
 - Usurpation d'identité d'auteurs
 - Phénomène des paper mills.
- Problematic Paper Screener : un outil pour détecter les publications suspectes
- Des éléments sur les bonnes pratiques

15H50 – 16H05 : questions réponses et mots de clôture



Vincent Guigue



Guillaume Cabanac