

VENDREDI

19 | **JUIN**
2026

8h30 à 15h30

Mener ses recherches avec l'IA générative

**Enjeux
méthodologiques
et éthiques**

**Repères pour
un usage éclairé**



2e Journée d'étude
organisée par le

Laboratoire Parole et Langage
5 avenue Pasteur – Aix-en-Provence

Inscrivez-vous ici !



UMR 7309 • CNRS • AMU
**PAROLE ET
LANGAGE**



amU
Aix Marseille Université

Programme

Matin

8h30 ACCUEIL CAFE (Jardin)

9h00 - 9h20 OUVERTURE (B011)

9h20 - 10H00 : Anissa Hamza-Jamann (ATILF – Université de Lorraine) : « Mon opinion est claire : ça dépend ! » : l'IA au fil des pratiques de recherche en sciences du langage.

Cette communication s'intéresse à ce que les chercheurs en sciences du langage font - ou ne font pas - avec l'intelligence artificielle générative. À partir d'une enquête exploratoire menée auprès d'enseignants-chercheurs, de doctorants et de chercheurs, elle donne à voir un tableau loin des polarisations qui dominent souvent le débat public. On y observe des pratiques discrètes et hésitantes, des positionnements ambivalents qui résistent aux catégories tranchées, des rapports à l'IA qui se construisent au fil des usages. Ces ambivalences éclairent quelque chose de plus profond : la manière dont les chercheurs négocient, souvent silencieusement, l'intégration ou non de nouveaux outils à leurs gestes et pratiques de recherche.

10H00 – 10H30 : Anna Kalinina (LiLPa – Université de Strasbourg) : Les LLMs comme assistants d'annotation linguistique : apports méthodologiques et limites pour l'analyse des expressions polylexicales

Cette communication présente une expérimentation portant sur l'usage des grands modèles de langue (LLMs) pour l'annotation d'expressions polylexicales nominales en français. Plusieurs modèles génératifs ont été évalués selon différentes configurations de prompting afin d'analyser leur capacité à reproduire une catégorisation linguistique fine.

Au-delà des performances quantitatives, cette étude interroge les usages méthodologiques des IA génératives dans la recherche en linguistique : transformation d'un guide d'annotation en protocole de prompting, influence des consignes sur le raisonnement des modèles, reproductibilité des annotations et place de la validation humaine. Les résultats montrent que les LLMs constituent des outils prometteurs pour la pré-annotation et l'assistance à l'analyse, mais que leurs décisions restent sensibles à la formulation des instructions et aux frontières sémantiques fines entre catégories. Cette communication discutera ainsi des apports et des limites des IA génératives dans les pratiques d'annotation linguistique, ainsi que des enjeux méthodologiques et éthiques liés à leur intégration dans les protocoles de recherche.

10H30 – 11H00 PAUSE CAFE (Jardin)

11H00 - 11H30 : Frédéric Béchet (LIS – Aix-Marseille Université) : Simuler l'humain avec des grands modèles de langue : Applications à l'annotation de corpus et aux simulations d'interactions humain-humain

L'utilisation des LLM comme substituts humains (human proxies) devient une alternative crédible pour l'annotation de corpus pour des tâches relativement simples où la consistance des annotations et la facilité à obtenir des données en très grand nombre permettent d'analyser de grands corpus pour y rechercher des phénomènes particuliers ainsi qu'entraîner des petits modèles par apprentissage supervisé à grande échelle.

Au-delà de la création de données, ces modèles permettent de simuler des interactions humaines en jouant le rôle de «personnage» défini par des profils complexes. Cette approche par simulation interactive est une piste prometteuse pour le développement et l'évaluation de modèles d'interaction humain-machine et humain-humain. Cette présentation donnera des exemples concrets de ces deux cas d'usage sur des corpus provenant de séries télé pour l'annotation linguistique (projet ANR MINERAL) et sur le domaine médical pour la simulation d'interactions (projet ANR REGARD).

11H30 - 12H00 : Amélie Guardiola & Caroline Péron (SCD – Aix-Marseille Université) : De la requête à la bibliographie : ce que l'IA fait (et défait) dans la recherche d'information scientifique

Cette intervention propose un état des lieux de l'usage des outils d'IA générative dans la recherche d'informations académiques. Elle s'appuie sur une cartographie des outils disponibles et de leurs usages pour la recherche, l'identification de ressources et la structuration de l'information scientifique. Au-delà des apports en termes de gain de temps et d'exploration, il convient de mettre en lumière certains points de vigilance : biais de visibilité, fiabilité des références générées, gestion opaque des données.

L'objectif de cette intervention est de fournir des repères concrets pour un usage raisonné de ces outils dans un contexte académique.

12H00 - 12H30 : Nicolas Santiago (Institut Laënnec – Aix-Marseille Université) : L'intégrité scientifique à l'heure de l'intelligence artificielle : vers une refondation de la responsabilité sociale des chercheurs par le biais de nouveaux dispositifs déclaratifs et réflexifs.

L'essor des modèles d'IA générative, pourrait faciliter la production d'articles frauduleux et rendre leur détection plus difficile, ce qui menace la fiabilité de la recherche. Préserver l'intégrité scientifique exige non seulement d'améliorer les outils de détection et de renforcer la transparence des données, mais aussi d'accroître la responsabilité des chercheurs.

000

12H30 - 13H30 Déjeuner (Jardin)



Après-Midi

De 13H30 à 15H30

Ateliers

Eric Olivier (CIPE – Aix-Marseille Université)

1 Quelques notions et modalités d'usage des IA génératives.

L'IA générative s'est imposée dans tous les registres du travail universitaire sans que l'institution y soit pleinement préparée. L'adoption de ces systèmes est loin de faire l'unanimité, mais il semble désormais nécessaire d'en connaître les principes d'usage ainsi que certains concepts clés.

Cet atelier a pour objectif de présenter les notions et techniques de base permettant une utilisation raisonnée des IA génératives dans le travail universitaire. Nous verrons notamment comment « parler aux IA », à travers quelques principes de base du prompting, quels sont l'utilité et les enjeux de la recherche générative, comment interroger expérimentalement l'« intelligence » des IA génératives, ainsi que les possibilités et les limites offertes par leurs capacités agentiques.

Nicolas Santiago

2 A propos de l'adoption de nouveaux dispositifs déclaratifs et réflexifs concernant les usages de l'IA dans la recherche et leurs effets potentiels sur l'intégrité scientifique : mise en discussion des enjeux et de la coordination des instances et des acteurs de la recherche.

Face aux nouvelles menaces pour l'intégrité scientifique causées par l'usage de l'IA dans la recherche, il est nécessaire d'engager les acteurs de la recherche vers le développement d'instruments à la fois normatifs et réflexifs, capables de restaurer un cadre de confiance que la seule détection ne suffit plus à garantir. L'atelier aura pour objectif de mettre en discussion les questions posées par la documentation proposée et d'engager une réflexion sur le rôle des instances et des acteurs de la recherche concernés par cet instrument.

Amélie Guardiola & Caroline Péron

3 Utilisation d'outils IA pour la recherche d'informations scientifiques.

L'atelier vous propose de tester des outils d'IA directement avec vos sujets de recherche pour explorer les potentialités (avantages et limites) dans la recherche d'informations. Nous testerons aussi les fonctionnalités de génération de bibliographie scientifique.



Brochure réalisée par Florian Bechard



Paul Pouzergues paul.pouzergues@univ-amu.fr



Claudia Starke claudia.starke@univ-amu.fr