



Workshop – le 3 décembre 2026 – Maison de la
Simulation – Bâtiment DIGITEO – CEA Saclay



« De l'assistance à l'autonomie : l'impact de l'IA agentique sur les processus de recherche »

Programme détaillé

AI-Based Scientific Hypothesis Generation: Lessons Learned from Generating 100 Hypotheses per Scientific Problems

Franck Cappello (ANL - USA)

Résumé :

The progress of LLMs and multi-agent systems in knowledge breadth and reasoning capabilities position them as potential resource for automatic hypothesis generation. Several AI-base Hypothesis Generation Tools (HGTs) have been specifically designed for this objective. Because HGTs are order of magnitude faster than humans, it becomes possible to generate multiple hypotheses per scientific problem. Indeed, some of the existing HGTs produce thousands of hypotheses per problems. This talk addresses the following questions: does it make sense to generate tens or even a hundred hypotheses per problem?

We presents the lessons learned in generating 100 hypotheses per problem for 6 different research problems.

Bio :

Franck conducts research in methods, modeling, and algorithms in the context of parallel and distributed computing for science leading to production quality software. He led the Exascale Computing Project (ECP) VeloC and SZ development, both receiving R&D100 awards. In the context of AI for science, Franck is leading the Best Practices for Scientific Workflows in the Genesis Mission. He is an IEEE and ACM fellow and recipient of the IEEE CS Charles Babbage Award.

Des agents pour la découverte scientifique : vers une recherche autonome à l'échelle institutionnelle

B. Bolliet (Department of Physics, University of Cambridge - UK)

Résumé :

Nous présenterons comment des systèmes multi-agents peuvent être conçus pour mener des travaux de recherche scientifique. Nous montrerons comment nous les utilisons pour résoudre des problèmes de recherche complexes et ouverts, avec, dans certains cas, des performances supérieures à celles d'humains travaillant sans leur assistance. Les agents et les systèmes que nous développons agissent comme de véritables multiplicateurs de capacités et génèrent des idées qui sont nouvelles, y compris pour nous. Il est toutefois essentiel de souligner qu'ils ne

fonctionnent pas de manière isolée : ils augmentent les capacités des chercheurs humains, ils ne les remplacent pas. Nous présenterons ensuite notre vision de IA scientifiques auto-organisés, évoluant en parallèle et faisant progresser la science en coévolution avec les humains.

Bio :

Dr Boris Bolliet est enseignant-chercheur à l'Université de Cambridge, où il dirige des recherches sur l'intelligence artificielle agentique. Après un doctorat en cosmologie à l'Université de Grenoble, il a poursuivi une carrière en astrophysique, avec des recherches postdoctorales à l'Université de Manchester, puis à l'Université de Columbia, à New York. Il a notamment participé à de grandes expériences internationales consacrées à l'étude du fond diffus cosmologique, telles que l'Atacama Cosmology Telescope (ACT) et le Simons Observatory (SO).

Depuis 2024, il développe l'utilisation d'agents IA pour la recherche scientifique à travers ses travaux sur Cmbagent, un système agentique primé à NeurIPS 2025 dédié à la génération et à l'exécution autonome de code scientifique, et Denario, un système multi-agents conçu pour automatiser plusieurs étapes clés du processus de recherche scientifique, de la génération d'idées à l'écriture et à la revue d'articles.

IA et Travail : quelles conséquences sur le travail en recherche ?

F. Barcellini (Professeure des Universités en Ergonomie au Cnam. Directrice du Centre de Recherche sur le Travail et le Développement)

Résumé :

En 2024 Flore Barcellini a coordonné une étude intitulée Le travail et l'emploi à l'épreuve de l'IA : État des lieux et analyse critique de la littérature à la demande de l'Institut de Recherches Économiques et Sociales. Ce rapport offre une analyse critique et nuancée des enjeux liés à l'introduction de l'IA dans les entreprises et institutions, allant au-delà des discours simplistes sur les bénéfices de l'automatisation au travail et la crainte de perte d'emplois. Il met en lumière l'importance cruciale des choix organisationnels et de la participation des travailleurs dans le déploiement de l'IA. Les effets de l'IA sur l'emploi et les conditions de travail ne sont pas prédéterminés, mais dépendent largement des décisions prises par les entreprises/institutions et les acteurs sociaux.

Ce rapport met également en garde contre des risques de l'IA pour le travail (subordination accrue, risques sur la fiabilité des actions, perte de sens du travail, diminution de la créativité et uniformisation de la pensée et des produits).

Les conclusions et les préconisations de ce rapport pourront être discutés au regard des enjeux et des spécificités des usages de l'IA et de l'IA agentique dans travail de recherche.

Bio :

Flore Barcellini est directrice du Centre de Recherche sur le Travail et le Développement et co-responsable de l'axe Mutations et Transitions du Travail et des Usages. Elle mène des recherches-actions pluridisciplinaires - avec les sciences de l'ingénieur et l'agronomie notamment - sur la conduite des transitions du travail dans les cadres des transformations sociétales (agroécologie, énergétique), techniques (technologies de l'industrie du futur, IA, robotique...) et organisationnelles (travail communautaire – Fab Lab/ Open Source software,

transformation du dialogue social). Dans ce cadre, elle a conduit une dizaine de projets de recherche avec des grands groupes industriels, des institutions publiques, des PME ou des bailleurs sociaux. Au Cnam, elle participe également à la formation d'ergonomes, des professionnels spécialistes de la conduite de projets innovante - participative et donnant une valeur à l'expérience des professionnels, usagers et citoyens- dans le cadre des transformations techniques, organisationnelles et sociétales du travail. Ingénieure physicienne de formation initiale, elle s'est formée à l'ergonomie au Cnam.

Vers des agents bio-informatiques contextualisés : graphe de connaissances traçable pour l'interprétation de signatures géniques

Guido Ugozzoni (CEA/DRF/IRIG - FR)

Résumé :

Dans le cadre d'essais cliniques en oncologie de précision, les technologies omiques permettent de mesurer le profil moléculaire de dizaines de milliers de gènes au sein de cohortes de patients. Ces données sont ensuite analysées pour extraire notamment des gènes marqueurs pronostiques, permettant d'anticiper la progression de la maladie, ou prédictifs, utilisés comme tests compagnons pour la réponse à un traitement médicamenteux. Cependant, pour une même pathologie, ces listes de gènes marqueurs varient d'une étude à l'autre, ce qui complique leur transfert en clinique. En effet, une liste de gènes marqueurs n'a de sens qu'au regard des conditions dans lesquelles elle a été établie : type de cancer, stade tumoral, composition de la cohorte, thérapie évaluée, sens de l'effet et issue clinique. Or ces conditions ne sont documentées que dans la littérature scientifique, dispersées entre des milliers de publications, ce qui rend très difficile la recherche de cohérence et de reproductibilité entre les gènes marqueurs identifiés dans différentes études.

Pour permettre cette recherche de consensus, nous travaillons à transformer un ensemble d'articles scientifiques en une base de connaissances structurée, interrogeable de façon automatisée. Cinq exigences guident cette approche : contrôler ce qu'un modèle de langage a le droit d'affirmer, en le contraignant à un cadre précis définissant les types d'informations et de liens autorisés, pour qu'il retranscrive fidèlement la littérature sans rien y ajouter ; préserver le sens des résultats scientifiques, qui associent souvent plusieurs éléments à la fois (un gène, un type de cancer, un stade, une cohorte, une thérapie, un effet observé) ; enrichir cette base à l'aide d'annotations génomiques externes, afin de faciliter l'intégration des connaissances entre publications ; représenter ces connaissances sous forme d'un graphe, offrant une structuration explicite et interrogeable ; et garantir la traçabilité, chaque affirmation devant rester reliée à l'article dont elle provient.

Cette approche est illustrée par la consolidation et l'interrogation des connaissances relatives à la résistance aux immunothérapies et aux thérapies ciblées pour les patients atteints de cancer du rein à cellules claires.

Bio :

G. Ugozzoni est chercheur au laboratoire Biosciences et Bioingénierie pour la santé (BGE) du CEA, qu'il a rejoint en décembre 2024 pour y travailler sur l'IA appliquée à la santé. Formé à la physique statistique et à la biologie computationnelle, il développe des méthodes

d'apprentissage automatique pour la biologie moléculaire et le design de protéines à visée thérapeutique.

Chimie numérique et matériaux énergétiques

C. Wespiser (CEA/DAM/DIF - FR)

Résumé :

Les matériaux énergétiques sont constitués de molécules pouvant, sous diverses sollicitations, libérer leur énergie chimique sur l'échelle temporelle de la microseconde. Ils sont utilisés à la fois pour des applications civiles telles que l'aérospatial ou les générateurs de gaz, et dans des domaines reliés à la défense. Les dangers inhérents à l'étude expérimentale de tels matériaux ralentissent cependant la compréhension des phénomènes physiques et chimiques sous-jacents à leur réactivité. Des modèles génératifs à la synthèse de précurseurs en passant par la simulation moléculaire, nous verrons dans cette présentation différents travaux entrepris au CEA-DAM visant à intégrer les technologies numériques dans ce domaine où seules quelques centaines de molécules ont été caractérisées expérimentalement. Les perspectives offertes par différentes technologies d'automatisation de la chimie seront discutées.

Bio :

J'ai obtenu mon doctorat en chimie-physique en 2020 à l'Université de Sherbrooke, au Québec. J'ai ensuite rejoint le CEA-Le Ripault, où j'ai travaillé sur la conception assistée par IA de nouvelles molécules organiques. Ce faisant, je me suis intéressé à diverses problématiques chimiques pouvant être abordées par des algorithmes d'apprentissage automatique, allant de la prédiction des propriétés moléculaires et des modèles génératifs incluant la faisabilité synthétique, à l'analyse de données expérimentales pour l'optimisation des conditions réactionnelles. En 2025, j'ai intégré le Laboratoire Matière en Conditions Extrêmes au CEA DAM Île-de-France. J'y ai élargi mes recherches à l'utilisation de l'apprentissage automatique pour enrichir les simulations de dynamique moléculaire. En combinant approches physiques et méthodes issues de la science des données, mes travaux visent à approfondir la compréhension de la réactivité chimique en phase condensée sous différents angles, afin d'accompagner les chimistes dans la conception efficace de nouvelles molécules fonctionnelles.

Diverses expérimentations de l'ia pour du développement logiciel et la conception de circuits

M. Duranton (CEA/DRT/LIST - FR)

Résumé :

La présentation présente l'état des expérimentations et réflexions sur l'utilisation de l'IA au Département des Systèmes et Circuits Intégrés Numériques (DSCIN) selon 3 directions :

- L'utilisation de l'IA pour la conception logicielle. Par exemple, un projet de plus d'un million de lignes de code a intensivement utilisé l'IA lors de son développement, et l'objectif de ce projet est de faire une plateforme pour distribuer les calculs d'IA, une approche RSI (Recursive Self-Improvement) en quelque sorte.
- L'utilisation de l'IA pour la conception de circuit : c'est un domaine assez différent du développement logiciel car les bases d'apprentissage des modèles d'IA classiques sont pauvres pour cet usage, et il n'y a (encore) que peu d'acteurs dans le domaine.

- La mise en œuvre de différentes manières de fournir les solutions d'ia, allant de solutions fonctionnant totalement en local sur les machines des utilisateurs, en passant par l'utilisation de ressources départementales ou du CEA (Maïa) allant pour certains projets ouverts à l'utilisation de ressources externes. Le choix du harnais (Claude Code, Codex, OpenCode, Pi, Hermès, DeepSeek, etc.) est aussi essentiel pour obtenir de bons résultats et certains harnais sont plus adaptés à certains modèles et à certains usages que d'autres.

Bio :

Marc Duranton est « senior fellow » au CEA et membre de la division « Département des Systèmes et Circuits Intégrés Numériques » du CEA, où il est responsable de la stratégie et participe à la mise au point de solutions (accélérateurs matériels et outils logiciels) destinées à l'intelligence artificielle et aux systèmes distribués, de l'Internet des objets (IoT) au calcul haute performance (HPC). Il a auparavant travaillé chez Philips et NXP, où il a dirigé le développement de la gamme de puces L-Neuro, des processeurs numériques utilisant des réseaux neuronaux artificiels, ainsi que de plusieurs coprocesseurs vidéo pour le processeur VLIW TriMedia. Il est responsable de la feuille de route de la communauté HiPEAC (<https://www.hipeac.net/vision/>).