

Liagora, le labo qui veut faire "le pont entre la recherche fondamentale en IA et les outils industriels"

A quelques jours du sommet pour l'action sur l'IA, les initiatives se multiplient dans le domaine. En se regroupant sous le même étendard, Illuin Technology et le laboratoire MICS de CentraleSupélec espèrent développer des systèmes sécurisés, moins énergivores et open source capables de se mesurer aux plus grands.

Célia Séramour - 28 janvier 2025 \ 15h50



© Liagora - De gauche à droite : Céline Hudelot, Directrice du Laboratoire MICS, Robert Vesoul, co-fondateur et dirigeant d'Illuin Technology et Romain Soubeyran, directeur de CentraleSupélec, lors de l'inauguration du laboratoire commun, Liagora.

LIAGORA : Laboratoire d'IA générative pour l'orchestration de recherches à base d'agents. L'acronyme est, certes un peu long, mais est plutôt précis quant au projet. Résultant d'un rapprochement entre Illuin Technology, société née dans le berceau de CentraleSupélec et MICS, l'un des laboratoires de la fameuse école d'ingénieurs, il s'agit d'un dispositif de recherche public/privé qui se veut ambitieux.

L'objectif est d'"être un pont entre les sujets de recherche fondamentale en IA générative, et la recherche appliquée, afin d'en faire des outils industriels pour les entreprises", annonce fièrement Robert Vesoul. Ce dernier porte une double casquette : il est à la fois co-fondateur et dirigeant d'Illuin Technology mais aussi co-directeur de la chaire [Innovation](#) à CentraleSupélec.

Les agents, le RAG et la sécurité en haut de la feuille de route

Concrètement, ce laboratoire, qui réunira notamment des post-docs, des doctorants, des research data scientists et des chercheurs, compte développer des assistants IA qui s'affranchissent des limites actuelles observées : hallucinations, besoin de recourir à l'approche RAG, manque de contexte long terme, peu d'utilisation des règles métier, etc. Une pierre apportée à l'édifice colossal qu'est l'IA "*responsable, éthique et digne de confiance*". Des principes qui s'accompagnent d'une démarche frugale, visant à développer des systèmes moins coûteux en énergie et en infrastructure. Partant de ces valeurs, Liagora se concentra sur trois thématiques déjà identifiées.

D'abord, la conception et l'orchestration d'agents basés sur des LLM avec la traduction de FAQ et corpus en règles métier, l'utilisation de règles métier et d'outils dans des assistants. La deuxième thématique concerne l'évaluation et la conception de systèmes de RAG : pipelines de RAG textuelles complexes et différenciées et pipelines de RAG multimodales. Enfin, la sécurisation des systèmes développés et des assistants qui en découlent par une hybridation de règles, de moteurs d'IA générative, sécurisation des entrées et sorties du système.

L'open source a la primeur

Liagora est en partie soutenu par l'Agence Nationale de la Recherche (ANR), établissement public placé sous la tutelle du ministère de l'Enseignement supérieur et de la Recherche. L'ANR finance les activités de recherche - plus précisément les salaires des chercheurs issus du laboratoire MICS - et demande en contrepartie un format de coopération d'une durée de cinq ans aux deux parties. "*Cela nous convenait bien. Dans cinq ans, quels seront les projets d'IA générative ? Nous aurons probablement un programme qui aura varié*", commente Robert Vesoul.

Le laboratoire entend par ailleurs soutenir l'écosystème open source et mettre en avant les communs numériques. De fait, les jeux de données, benchmarks et modèles développés seront publiés en open source, prévient le dirigeant d'Iluin Technology. Des exemples passés montrent d'ailleurs l'engagement de cette nouvelle entité sur le sujet.

Des liens académiques et de recherches forts

Depuis 2018, Illuin Technology et le laboratoire MICS travaillent ensemble au travers de nombreux programmes et projets. "*Nous avons des liens académiques et de recherche avec Paris-Saclay et CentraleSupélec très fort. Cela nous permet de faire des projets communs, d'avoir des initiatives académiques assez originales pour garder et aller capter tous ces jeunes talents dans le domaine de l'IA*", résume Robert Vesoul.

Ils comptabilisent de fait 7 publications communes depuis 2023 et embarquent de nombreux chercheurs des deux côtés, à l'instar de Manuel Faysse et sa thèse CIFRE portant sur l'amélioration des méthodes de recherche d'information dans des corpus multimodaux. Parmi les travaux les plus repris à l'international, notons [CroissantLLM](#) - un modèle de langage bilingue (français-anglais) optimisé pour des applications industrielles -, et [ColPali](#), une technologie de RAG multimodal améliorant la compréhension de documents enrichis de tableaux, images et diagrammes.

Robert Vesoul estime que cette initiative permet de se mesurer à des compétiteurs plus féroces et souvent mieux soutenus financièrement. *"Quand nous comparons les équipes de recherche que nous pouvons mettre en collaboration au niveau européen, nous arrivons à des choses qui sont proches des capacités de recherche d'un Meta ou autre. Il n'y a pas à rougir. Selon lui, le cadre dans lequel s'inscrit Liagora est un moyen de faire de la recherche de très bon niveau. Nous n'allons pas faire des modèles de fondation et des LLM dans le cadre de ce laboratoire, mais nous entraînerons des modèles pointus. Le fait que nous soyons dans ce cadre public-privé aide aussi : nous aurons à disposition les capacités de calcul du supercalculateur Jean Zay".*