

Communiqué de presse – 21 septembre 2023



## Le Projet LLM4All financé par l'ANR : un consortium gagnant apportant les dernières avancées en matière d'IA générative au service de la santé

LINAGORA, le CNRS, au travers du Laboratoire lorrain de recherche en informatique et ses applications (Loria – CNRS/Inria/Université de Lorraine) et du Laboratoire d'informatique de l'École polytechnique (LIX - CNRS/Institut Polytechnique de Paris) l'AP-HP et HuggingFace sont très heureux d'annoncer que l'ANR financera l'initiative LLM4All portée par un consortium d'acteurs stratégiques et complémentaires issus du monde de la recherche et de la technologie.

Deux cas d'usage **en matière de santé** ont été identifiés : le résumé automatique de réunions et l'analyse des appels d'urgence du SAMU. Pour le premier, un grand modèle de langue (*Large Language Model* ou *LLM*) actualisé sera formé sur un ensemble de données de réunions qui sera mis à disposition **sous licence Open Source**. Pour le second, un modèle combinant les aspects RAP et NLP sera spécifiquement adapté pour analyser un ensemble d'appels d'urgence simulés, enrichi par des données externes, y compris des éléments vocaux et des ontologies médicales.

« Avec LLM4All, les partenaires du consortium associent leurs capacités en R&D et ingénierie combinant innovation, efficacité et accessibilité pour répondre aux besoins croissants en termes d'IA Générative multimodale souveraine et Open Source. » **Michel-Marie Maudet, Directeur Général de LINAGORA**

« Le CNRS soutient fortement les initiatives de recherche sur des thématiques liées au traitement automatique de la langue, en particulier les systèmes d'IA génératifs, à travers ses laboratoires et en mettant son supercalculateur Jean Zay à disposition pour l'entraînement des modèles. En outre, le CNRS est engagé dans le développement et la valorisation du logiciel libre. La collaboration avec des acteurs importants du monde du logiciel libre en France fait donc partie intégrante de notre effort pour la production des modèles de type LLM en langue française. » **Adeline Nazarenko - Directrice de l'Institut des sciences de l'information et de leurs interactions (INS2I) du CNRS**

### LINAGORA

Villa Good Tech  
37 Rue Pierre Poli 92130 Issy-les-Moulineaux  
[www.linagora.com](http://www.linagora.com)

« LLM4All est à l'origine issu d'un rapprochement stratégique entre LINAGORA et le CNRS. Je suis ravi que cette collaboration se soit, par la suite, ouverte à de nouveaux acteurs clés. LLM4All est porté par une double ambition : en matière de recherche pure, mieux comprendre comment et pourquoi les LLM apprennent et oublient et dans son application, contribuer à réduire les coûts d'adaptation de ces modèles. Par ailleurs, d'un point de vue stratégique, LLM4All a vocation à ouvrir les LLM selon une vision Open Source chère à LINAGORA et bien sûr, aussi partagée par l'ensemble des partenaires du projet. » **Christophe Cerisara – Chercheur et Coordinateur du projet LORIA-Synalp**

## **LLM4All répond à un double challenge technique vers des grands modèles de langage plus efficaces et plus accessibles**

Au cours de ces derniers mois, les LLM ont démontré des capacités émergentes impressionnantes, notamment dans la génération de multiples formes de contenus, l'apprentissage en contexte et la décomposition de problèmes complexes en parties pouvant être traitées par des IA, soulignant l'importance de la taille comme critère déterminant pour la qualité des modèles. L'initiative LLM4All se lance par conséquent avec comme ambition majeure de développer des LLM de grande envergure tout en solutionnant deux défis cruciaux :

Le premier défi est la **mise à jour automatique des modèles** :

Contrairement à une idée reçue, il ne suffit pas de continuer à former un LLM avec de nouvelles données. Sans des processus d'apprentissage adéquats, ces modèles risquent d'« oublier » des informations précédentes. LLM4All s'attaque à ce défi en explorant des approches innovantes, telles que l'intégration de réseaux neuronaux expansifs et de techniques de parcimonie. L'objectif est de fournir des **LLM multilingues constamment mis à jour**, en s'appuyant sur de nouvelles techniques développées dans le cadre du projet.

Le second défi est de rendre ces modèles **plus accessibles en termes de coûts de calcul** :

LLM4All propose une gamme de solutions adaptées à différents besoins, allant d'optimisations spécifiques pour équilibrer vitesse et mémoire, à des méthodes collaboratives pour accélérer les calculs, en passant par des techniques parcimonieuses et des méthodes de distillation pour des tâches spécifiques. Mais LLM4All ne s'arrête pas là. Le projet vise également à valider ces avancées technologiques sur des applications concrètes en **langue française**.

## **LLM4All : Deux cas d'usage appliqués au domaine de la santé et en français**

Au-delà des modèles génériques qui seront diffusés, ces approches seront validées sur deux cas d'usages en français, respectivement le résumé automatique de réunions et l'analyse des appels d'urgence du **SAMU**.

Pour le premier, un LLM à jour sera entraîné sur un grand corpus disponible de réunions et diffusé en **Open Source**. Pour le second, un LLM sera adapté à un corpus d'appels d'urgence simulé et combiné à des informations externes, dont la voix et des ontologies médicales.

« Avec LLM4all, on va intégrer un modèle de la pathologie d'urgence dans le LLM en désignant des tâches

### **LINAGORA**

Villa Good Tech  
37 Rue Pierre Poli 92130 Issy-les-Moulineaux  
[www.linagora.com](http://www.linagora.com)

*auto-supervisées et semi-supervisées adaptées à nos objectifs médicaux, en collaboration étroite avec des experts du domaine. » Ivan Lerner - Assistance Publique – Hôpitaux de Paris | AP-HP · Département de Santé Publique, Informatique Médicale et Biostatistique*

## **LLM4All : vers la construction de – CLAIRE –, un modèle Open Source en langue française**

Ainsi, LLM4All prépare actuellement la construction d'une première version d'un modèle dénommée – CLAIRE – visant à tester ces différentes approches sur un ensemble de datasets « propres » **qui respectent les exigences visées par l'IA Act**. L'objectif du consortium LLM4All est d'entraîner CLAIRE sur les prochaines semaines afin d'obtenir un modèle opérationnel avant la fin de mois d'octobre.

### Note aux éditeurs :

*Les travaux visés par le projet LLM4All s'inscrivent pleinement dans l'actualité de cette rentrée 2023 :*

*- Début septembre, on a vu en l'espace d'une même semaine apparaître des modèles à 180B de paramètres pour Falcon et des modèles à 1B pour Llama. La période actuelle est celle de la recherche empirique par des tests en vue d'établir dans les prochaines semaines de nouvelles lois ou à minima de nouvelles approches de construction des LLM ;*

*- L'apprentissage automatique en continu est aussi au cœur de la recherche des LLM : on parle désormais de ReST (Reinforced Self-Training for Language Modeling) et de RLAIIF (Scaling Reinforcement Learning from Human Feedback with AI Feedback → <https://arxiv.org/pdf/2309.00267.pdf>) visant à avoir des boucles de feedback plus rapides et plus efficaces ;*

*- Enfin avec les techniques de quantization telles que GPTQ (Accurate Post-Training Quantization for Generative Pre-trained Transformers → <https://arxiv.org/abs/2210.17323>), on cherche à diminuer drastiquement la taille des modèles afin de les exécuter sur des infrastructures plus économiques et plus efficaces d'un point de vue énergétique ;*

### **LINAGORA**

Villa Good Tech  
37 Rue Pierre Poli 92130 Issy-les-Moulineaux  
[www.linagora.com](http://www.linagora.com)



Pour plus d'informations, veuillez contacter :  
Michel-Marie MAUDET [contact@openllm-france.fr](mailto:contact@openllm-france.fr) +33660469852

### **À propos d'OpenLLM France :**

LINAGORA est à l'initiative de la communauté OpenLLM France qui a pour but de créer une IA générative multimodale souveraine et open source. Nous croyons en l'importance de la transparence et du partage dans le développement des technologies d'IA.

Le manifesto de cette initiative est disponible ici : <https://github.com/OpenLLM-France/Manifesto/tree/main>

Pour poursuivre nos échanges, nous vous invitons à rejoindre le serveur : <https://discord.gg/k2BQvGfjzA>

**Le prochain Meetup OpenLLM France #2 se tiendra à la Villa Good Tech le mercredi 11 octobre 2023 à partir de 18h30.**

### **À propos de LINAGORA**

LINAGORA est une Entreprise à Mission, pionnier Français du Logiciel Libre. Sa Mission depuis 23 ans, est d'inventer, développer, promouvoir et maintenir des technologies et services Open Source, qui contribuent à préserver l'indépendance technologique de ses clients, renforcent la souveraineté numérique des États, contribuent à diminuer l'impact énergétique des solutions mises en œuvre. Depuis sa création, LINAGORA défend un numérique éthique, une « 3ème Voie Numérique », respectueuse des droits des usagers, inclusive et durable, capable d'apporter une alternative aux géants américains et chinois.

Parmi les logiciels phares de LINAGORA, Twake Digital Workplace offre une alternative 100 % Open Source aux solutions des GAFAM ; LinShare, permet le partage sécurisé et le stockage de fichiers, et enfin Team Mail est la 1ère application d'e-mail basée sur le protocole JMAP et le serveur de messagerie JAMES développé par la Fondation Apache. Le protocole JMAP est un nouveau standard de mail permettant de réduire de 50 % son impact carbone.

Enfin, très active dans l'innovation technologique, LINAGORA met au point LinTO, une plateforme Open Source d'intelligence conversationnelle et de technologies vocales. Les APIs cognitives de LinTO (transcription & analyse du langage), sont particulièrement performantes dans les contextes d'échanges en entreprise impliquant de nombreux intervenants.

En plus de son offre Logiciel, LINAGORA propose son offre d'OSSA (Open Source Software Assurance) qui assure le Maintien en Conditions Opérationnelles (MCO) et le Maintien en Conditions de Sécurité (MCS) des logiciels Open Source utilisés dans les systèmes d'informations les plus critiques de ses clients. Enfin, LINAGORA accompagne ses clients avec une gamme complète de services professionnels : conseil, expertise technique, développement, formation.

Présent sur tout le territoire français, LINAGORA dispose également de bureaux au Canada, au Vietnam, en Tunisie et vend ses logiciels et services partout dans le monde.

### **À propos du CNRS**

Le Centre national de la recherche scientifique est une institution publique de recherche parmi les plus reconnues et renommées au monde. Depuis plus de 80 ans, il répond à une exigence d'excellence au niveau de ses recrutements et développe des recherches pluri et inter disciplinaires sur tout le territoire, en Europe et à l'international. Orienté vers le bien commun, il contribue au progrès scientifique, économique, social et culturel de la France. Le CNRS, c'est avant tout 32 000 femmes et hommes et 200 métiers. Ses 1000 laboratoires, pour la plupart communs avec des universités, des écoles et d'autres organismes de recherche, représentent plus de 120 000 personnes ; ils font progresser les connaissances en explorant le vivant, la matière, l'Univers et le fonctionnement des sociétés humaines. Le lien étroit qu'il tisse entre ses

#### **LINAGORA**

Villa Good Tech  
37 Rue Pierre Poli 92130 Issy-les-Moulineaux  
[www.linagora.com](http://www.linagora.com)



activités de recherche et leur transfert vers la société fait de lui aujourd'hui un acteur clé de l'innovation. Le partenariat avec les entreprises est le socle de sa politique de valorisation. Il se décline notamment via près de 200 structures communes avec des acteurs industriels et par la création d'une centaine de start-up chaque année, témoignant du potentiel économique de ses travaux de recherche. Le CNRS rend accessible les travaux et les données de la recherche ; ce partage du savoir vise différents publics : communautés scientifiques, médias, décideurs, acteurs économiques et grand public. Pour plus d'information : [www.cnrs.fr](http://www.cnrs.fr)