

## IA générative : le MICS lance CroissantLLM un modèle de langage souverain, ouvert, éthique et frugal

*Cette innovation est portée par les professeurs Céline Hudelot et Pierre Colombo dans le cadre des travaux de thèse de Manuel Faysse, en collaboration avec Nuno Miguel Guerreiro et Patrick Fernandes.*

*Elle sera présentée le 7 mars à Paris La Défense dans le cadre de l'événement « Les Ateliers de l'IA », édition spéciale coorganisée par Le Digital Lab de CentraleSupélec et Illuin Technology.*

Les équipes de recherche du [laboratoire MICS](#) de CentraleSupélec ont développé avec [Illuin Technology](#) un nouveau modèle de langage (LLM) appelé CroissantLLM, dorénavant disponible sur la plateforme Hugging Face.

Cette brique technologique d'IA va permettre aux entreprises d'intégrer dans leurs processus métiers une IA souveraine, frugale et de culture française.

Cinq caractéristiques distinguent CroissantLLM des IA génératives actuellement proposés sur le marché, qui sont détaillés dans [l'article scientifique](#) publié :

- Il est **souverain** : c'est une équipe de recherche française qui l'a développée. Elle s'est appuyée sur le calculateur national Jean Zay au sein du GENCI ; les jeux de données sont également français et publics, donc connus et traçables.
- Il est **ouvert et responsable** : tout est en open source et donc accessible, des algorithmes aux modèles et jeux de données utilisés, pour une transparence totale.
- Il maîtrise la **langue française** : il a été entraîné sur autant de contenus en français que de contenus en anglais, ce qui lui permet d'intégrer et de maîtriser les spécificités de la langue et de la culture françaises.
- Il est **éthique** : l'équipe de recherche a veillé au respect des règles fixées par le récent AI Act.
- Il est **frugal** sans pour autant faire de concessions sur la vitesse : il compte seulement 1.3B de paramètres (contre +180 pour ChatGPT 3.5 par exemple). Il n'a donc pas besoin d'importantes puissances de calcul pour s'exécuter, ce qui lui permet de tourner sur des smartphones et ordinateurs personnels et pas uniquement en sollicitant plusieurs GPU. Par voie de conséquence, son fonctionnement consomme peu d'énergie. Pour autant, les benchmarks réalisés montrent qu'il est le modèle francophone le plus performant pour sa taille.

*« Ce travail est le fruit d'une collaboration étroite entre académie et industrie, illustrant l'importance de la synergie dans l'avancée de la recherche en IA. CroissantLLM résulte en effet des travaux menés par CentraleSupélec en collaboration avec plusieurs partenaires académiques de renom tels que Sorbonne Université, INESC-ID, Instituto Superior Técnico, Carnegie Mellon University et Institut DATAIA. Elle a aussi pu voir le jour grâce au soutien précieux de partenaires industriels : ILLUIN Technology ainsi que Unbabel, Diabolocom, et EqualAI », commente Céline Hudelot, professeure, directrice du laboratoire MICS.*

### Contacts presse :

Claire Flin : [claireflin@gmail.com](mailto:claireflin@gmail.com) – 06 95 41 95 90  
Marion Molina : [marionmolinapro@gmail.com](mailto:marionmolinapro@gmail.com) - 06 29 11 52 08

Pour permettre aux entreprises de tester CroissantLLM et de s'emparer pleinement du potentiel offert par cette IA générative, une présentation et une démo seront faites le 7 mars à 14h20 dans le cadre de l'événement « Les Ateliers de l'IA » organisé à Paris La Défense et en distanciel.

Programme détaillé sur <https://www.illuin.tech/ateliers-de-lia-2024/>

Adresse : Amphithéâtre du centre de conférences de Cœur Défense situé 110, esplanade de la Défense, 92400 Courbevoie – M° et RER Grande Arche La Défense

Accréditation presse auprès de Claire Flin ([claireflin@gmail.com](mailto:claireflin@gmail.com)) et Marion Molina ([marionmolinapro@gmail.com](mailto:marionmolinapro@gmail.com))



L'équipe de recherche du MICS

**A propos de CentraleSupélec - [www.centralesupelec.fr](http://www.centralesupelec.fr)**

CentraleSupélec est un établissement public à caractère scientifique, culturel et professionnel, né en janvier 2015 du rapprochement de l'École Centrale Paris et de Supélec. Aujourd'hui, CentraleSupélec se compose de 4 campus en France (Paris-Saclay, Metz, Rennes et Reims). Elle compte plus de 5 000 étudiants, dont 3 800 élèves ingénieurs, et regroupe 18 laboratoires ou équipes de recherche. Fortement internationalisée (25 % de ses étudiants et près d'un quart de son corps enseignant internationaux), l'école a noué plus de 170 partenariats avec les meilleures institutions mondiales. Ecole leader dans l'enseignement supérieur et la recherche, CentraleSupélec constitue un pôle de référence dans le domaine des sciences de l'ingénierie et des systèmes, classée parmi les meilleures institutions mondiales. Elle a cofondé l'Université Paris-Saclay en 2020 et préside le Groupe des Écoles Centrale (CentraleSupélec, Centrale Lyon, Centrale Lille, Centrale Nantes et Centrale Méditerranée) qui opère les implantations internationales (Pékin (Chine), Hyderabad (Inde), Casablanca (Maroc)).

**Contacts presse :**

Claire Flin : [claireflin@gmail.com](mailto:claireflin@gmail.com) – 06 95 41 95 90  
Marion Molina : [marionmolinapro@gmail.com](mailto:marionmolinapro@gmail.com) - 06 29 11 52 08