

Communiqué de presse – Issy-les-Moulineaux – 25 janvier 2025

Message important concernant la plate-forme publique lucie.chat

*Ces dernières heures sur les réseaux sociaux, plusieurs commentaires et posts ont attiré notre attention. Il est évident que nous n'avons pas suffisamment bien communiqué et clarifié **sur ce que LUCIE peut ou ne peut pas faire dans son état actuel**, ainsi que la nature des travaux réalisés jusqu'à présent. Il est donc important de revenir sur certains points clés pour éviter toute confusion et polémique inutile.*

1 Un projet de recherche académique encore en phase initiale

LUCIE est avant tout un projet de **recherche académique** visant à démontrer les capacités à développer des communs numériques d'IA générative. À ce jour, **aucun travail spécifique n'a été réalisé avec l'Éducation Nationale** pour personnaliser ou adapter le modèle à un usage éducatif.

Toute utilisation dans un contexte de production est donc prématurée.

LUCIE est le fruit de plusieurs mois de travail de toute une communauté, la communauté OpenLLM-France, qui regroupe de nombreux chercheurs et plusieurs entreprises dont LINAGORA.

Une partie de cette communauté s'est réunie en consortium autour de LINAGORA pour répondre à un Appel à Projet France 2030 visant justement à financer des travaux de recherche en vue de développer "des communs numériques dans le domaine des IA génératives".

Lors de la phase d'instruction de notre réponse à cet AAP, il nous a été demandé de nous concentrer sur des cas d'usages liés au domaine de l'éducation, puisqu'un des membres du consortium est une association de ce domaine.

2 Ce que LUCIE est et ce qu'elle n'est pas

LUCIE, dans sa version actuelle, est un modèle "brut", développé à des fins de recherche. À ce stade, elle ne dispose pas de :

- Instruction approfondie : le modèle fonctionne avec des réglages minimaux.
- RHLF (Renforcement par Apprentissage Humain) : aucune optimisation n'a encore été réalisée pour calibrer les réponses.
- Garde-fous (guardrails) : aucune prévention systématique contre des usages inappropriés.

LINAGORA

Villa Good Tech, 37 Rue Pierre Poli 92130 Issy-les-Moulineaux

www.linagora.com

Les réponses générées par LUCIE ne sont donc pas garanties et certaines contiennent des biais et des erreurs.

Encore une fois, LUCIE ne doit pas être utilisée dans des contextes éducatifs ou de production dans sa version actuelle.

LUCIE est un modèle de langage et pas un modèle de connaissance. Comme il est d'usage pour tout LLM en production, le modèle doit être déployé avec un système prompt robuste associé à du RAG pour obtenir des résultats satisfaisants.

3 Une mise en ligne prématurée et une décision nécessaire

Conscients que la phase d'instruction n'était que partielle, nous avons pensé, à tort, qu'une mise en ligne publique de la plateforme lucie.chat était néanmoins possible dans la logique d'ouverture et de co-construction des projets Open Source.

Cette ouverture était souhaitée pour exposer une première itération du modèle (qui est très attendu) et aussi un appel à la collecte des **données d'instruction** qui nous manquent cruellement. **Le véritable défi se situe en effet dans les données d'instruction**, c'est-à-dire sur la façon dont un modèle doit interagir avec un utilisateur humain. Les datasets publics existants sont principalement en langue anglaise, pas en Open Source, et posent des problèmes de valeurs et de culture.

Là où les acteurs des Big Techs semblent utiliser des millions de paires de questions/réponses (principalement en anglais et sans que ces données soient open source) pour aligner leurs modèles, nous, nous nous appuyons sur quelques dizaines de milliers de paires en français couvrant peu de domaines et donc peu de cas d'usage.

Nous savons bien entendu que les capacités de "raisonnement" (y compris sur des problèmes mathématiques simples) ou encore la capacité à générer du code de la version actuelle de LUCIE sont insatisfaisantes. Nous aurions dû informer les utilisateurs de la plate-forme de ces limites de telle manière à ne pas créer d'attente inutile.

Nous n'aurions pas du sortir le service lucie.chat sans ces explications et précautions d'usage. Nous avons été emportés par notre propre enthousiasme. Nous allons donc nous y reprendre pour mieux expliquer notre démarche. Pour cela, nous avons besoin de calme et de temps. **Aussi, nous préférons fermer temporairement l'accès à la plateforme lucie.chat.**

Nous demandons à tout le monde de respecter le travail des chercheurs et des ingénieurs qui œuvrent pour développer des systèmes d'IA de confiance, plus éthiques et plus transparents.

Nous restons à l'écoute de vos retours et vous remercions pour votre **compréhension** et votre **soutien** dans cette aventure collective. Ensemble, faisons de LUCIE un véritable bien commun numérique.

Nous restons à votre disposition pour répondre à vos questions et échanger davantage

LINAGORA

Villa Good Tech, 37 Rue Pierre Poli 92130 Issy-les-Moulineaux

www.linagora.com

sur ce projet. Deux posts permettent néanmoins de connaître l'état actuel de nos travaux ainsi que les prochaines étapes du projet :

- <https://tinyurl.com/3sn6st7a>
- <https://tinyurl.com/3hcxud3c>

Vous pouvez nous contacter à l'adresse suivante : ✉ contact@openllm-france.fr

À propos d'OpenLLM France

OpenLLM France (<https://www.openllm-france.fr>) est une initiative 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/tZf7BR4dY7>

Pour plus d'informations, merci de contacter :

Michel-Marie MAUDET (contact@openllm-france.fr)

À propos de LINAGORA

LINAGORA est une Entreprise à Mission, pionnier Français du Logiciel Libre. Sa Mission depuis 2000, est d'inventer, développer, mettre en œuvre et maintenir des technologies Open Source, pour préserver l'indépendance technologique de ses clients et la souveraineté numérique des États en particulier de la France et de l'Europe. Cette mission se conjugue avec la volonté d'avoir un maximum d'impacts positifs pour ses équipes, la société et la planète.

Depuis sa création, LINAGORA défend un numérique éthique, une « 3ème Voie Numérique », respectueuse des droits de chacun, responsable, inclusive et durable, capable d'apporter une réelle alternative aux géants américains et chinois.

Parmi ses logiciels phares LINAGORA, développe **Twake Workplace** qui se positionne comme une alternative 100 % libre aux solutions des GAFAM. Twake Workplace est disponible sous la forme d'une plate-forme complète ou bien module par module. Elle comprend notamment :

- **Twake Mail**, une puissante messagerie moderne basée sur le protocole JMAP et le serveur d'e-mail JAMES de la fondation Apache dont LINAGORA assure le leadership technique ;
- **Twake Chat**, une solution de communications instantanées pour entreprise

LINAGORA

Villa Good Tech, 37 Rue Pierre Poli 92130 Issy-les-Moulineaux

www.linagora.com

développée sur le protocole MATRIX et compatible avec la solution de chat de l'État Français, Tchap ;

- **Twake Drive** : une plateforme collaborative très facile d'usage permettant le travail en groupe grâce à OnlyOffice.

En plus de Twake Workplace, LINAGORA développe **LinShare** qui est une solution de partage sécurisé de fichiers volumineux.

Depuis près de 10 ans, LINAGORA travaille également dans le domaine de la voix et développe ses propres algorithmes de transcription et modèles de langage. Ces technologies se retrouvent dans **LinTO**, une plateforme Open Source d'intelligence conversationnelle. LinTO permet l'enregistrement de réunions, leur transcription et leur édition en mode collaboratif afin de les organiser et de pouvoir en tirer partie grâce à l'intelligence artificielle.

LINAGORA est aussi un acteur clef du domaine de l'intelligence artificielle Open Source. LINAGORA propose une gamme complète de services et de produits cognitifs basés sur des modèles fondations sobres, compacts, Open Source et souverains. Ces derniers sont développés en mode communautaire avec un ensemble d'acteurs académiques, organisations publiques et privées réunis au sein de la communauté **OpenLLM France**.

En plus de son offre logicielle, LINAGORA propose son offre unique 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 grands 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.

Pour plus d'informations ou demande d'interview, merci de contacter :

Frédéric Hermelin – fhermelin@linagora.com

LINAGORA

Villa Good Tech, 37 Rue Pierre Poli 92130 Issy-les-Moulineaux
www.linagora.com