LISTIC

Offre de stage 2025/2026

Titre	Conception de méthodes de préservation des informations sensibles lors du traitement automatisé des documents numérisés par les LLMs.
Niveau	Master 2ème année / Ingénieur 5ème année
Date de début/ fin	~4-5 mois, entre mi-février et mi-juillet 2026
Ville, Pays	Annecy-le-Vieux, *France*
Laboratoire	Laboratoire d'Informatique, Systèmes, Traitement de l'Information et de la Connaissance - LISTIC
Description du sujet	**Contexte :** Les grands modèles de langage (*LLMs*, pour *Large Language Models*) représentent une approche émergente permettant l'extraction d'informations à partir de documents administratifs numérisés dans divers domaines, notamment la santé, la finance ou encore l'assurance. Toutefois, l'utilisation de ces modèles dans des contextes sensibles présentent sûrement un risque de divulgation d'informations confidentielles ou de vulnérabilité à des attaques [Yao 2024]. En effet, les études existantes traitent directement les images de documents comme Donut [Kim 2021] ou bien proposent de combiner les LLMs avec un OCR [Loukil 2024]. Elles accordent une attention particulière à l'exactitude des données extraites tout en négligeant la préservation des informations sensibles, notamment des données personnelles, stratégiques ou à risque. Ainsi, l'objectif de ce stage est de mettre en place un processus pour masquer les informations sensibles dans les sorties des LLMs afin de satisfaire aux exigences légales (RGPD) ou aux recommandations fondées sur des enjeux sociétaux et éthiques. Ce stage consiste à, dans un premier temps, étudier les méthodes existantes de préservation des informations sensibles lors du traitement automatisé des documents numérisés par les LLMs. Puis, dans un second temps, il vise à concevoir une méthode de pseudonymisation des informations sensibles dans les sorties des LLMs. Enfin, la solution proposée sera appliquée sur deux jeux de données ayant différentes caractéristiques, notamment des données synthétiques dans des documents financiers sous forme de textes et d'images. **Objectifs du stage :** 1. L'étudiant.e étudiera l'état de l'art sur les méthodes existantes de préservation de la privacy dans le cadre de l'utilisation des LLMs. 2. Il.elle développera une solution basée sur les techniques de privacy pour les LLMs et les graphes de connaissances pour sécuriser la gestion des données sensibles des documents financiers. **3.** Il.elle réalisera une analyse des performances de la solution proposée et une analyse comparative de différentes configurations. La valorisation des résultats obtenus fera l'objet d'une publication dans une conférence.

	Références bibliographiques : **[Yao 2024]** Yao, Y., Duan, J., Xu, K., Cai, Y., Sun, Z., & Zhang, Y. (2024). A survey on large language model (llm) security and privacy: The good, the bad, and the ugly. *High-Confidence Computing*, 100211. **[Kim 2021]** Kim, G., Hong, T., Yim, M., Park, J., Yim, J., Hwang, W., ... & Park, S. (2021). Donut: Document understanding transformer without ocr. *arXiv preprint arXiv:2111.15664*, *7*(15), 2. **[Loukil 2024]** Loukil, F., Cadereau, S., Verjus, H., Galfre, M., Salamatian, K., Telisson, D., ... & Le Van, O. (2024, November). LLM-centric pipeline for information extraction from invoices. In *2024 2nd International Conference on Foundation and Large Language Models (FLLM)* (pp. 569-575). IEEE.
Compétences requises	**Profil recherché :** • Étudiant en Master 2 en Informatique, Intelligence Artificielle, Traitement d'Image ou domaines connexes. • Connaissances en deep learning et traitement de graphes de connaissances. • Maîtrise de Python et de bibliothèques comme PyTorch/TensorFlow. • Intérêt pour les grands modèles de langage (LLM).
Gratification	Selon législation en vigueur
Tuteurs / Contacts	**Tuteurs de stage** : Faiza Loukil, Hervé Verjus, Kavé Salamatian **Contact** : {prenom.nom}@univ-smb.fr