

# Offre de stage M2 IHM & Text-Mining

**Type de stage :** M2

**Titre :** Déploiement et évaluation d'un annotateur immersif et embarqué, application à la reconnaissance d'entité

**Mots-clés :** Annotation, XR, extraction d'information

## **Contexte :**

Les techniques d'extraction d'information (*text-mining*) récentes les plus performantes reposent sur de l'apprentissage automatique supervisé. Or, l'évaluation et l'entraînement de ces méthodes nécessitent une phase d'annotation humaine laborieuse. Cela est particulièrement vrai pour les domaines spécialisés, tels que la microbiologie, où la production et la validation des annotations nécessitent l'intervention d'experts, et où les données produites sont d'un grand intérêt [Deléger2016]. La Réalité Virtuelle (RV) et par extension la Réalité Étendue (XR) permettent d'améliorer les performances d'utilisateurs ainsi que leur engagement lors de la réalisation de tâches spécifiques comme de la rééducation ou de l'apprentissage humain [David2017, Mayer2023].

## **Projet :**

Nous proposons dans ce stage de développer et d'évaluer un outil d'annotation textuelle immersif en Réalité Virtuelle (RV) similaire à VAnnotator [Spiekermann2018] pouvant tirer parti des possibilités d'annotation en Réalité Étendue (XR) [Borhani2023]. La tâche d'annotation spécifique visée sera la reconnaissance d'entités. Cette dernière consiste à identifier et classifier des entités nommées dans un texte (ex : identifier dans la phrase suivante que la mention "*Escherichia coli*" est une bactérie - "*Escherichia coli est trouvée dans les intestins des animaux*"). L'accent sera mis sur la mise en place de l'annotateur immersif ainsi que de son évaluation comparée avec les outils existants.

## **Objectifs du stage :**

- Revue de la littérature des solutions d'annotation immersives.
- Transfert de données textuelles issues d'un [corpus BB4](#) pour la reconnaissance d'entités et développé à MaIAGE dans un casque de RV.
- Installation et configuration d'un annotateur standard [Colucci2024] dans un environnement immersif, pour la visualisation des données, le développement de fonctionnalités et pour améliorer l'annotation en RV.
- Mise en place d'expérimentations utilisateurs pour comparer l'utilisation immersive et non-immersive de Doccano (*cybersickness*, utilisabilité, etc.).
- Rédaction du manuscrit pour diffusion et documentation.

Le stage pourra être l'occasion de préparer un poster pour une conférence des domaines concernés (ex : IHM, TALN).

**Compétences requises :**

- Compétences en développement informatique (préférentiellement web, Python et Android)
- Compétences en IHM (évaluation de l'expérience utilisateur, analyse d'applications interactives)
- Une expérience en développement XR serait un plus
- Capacité de travail en autonomie
- Un attrait pour la recherche scientifique

**Formations visées :** informatique (IHM, IA, etc.) ou équivalent.

**Durée :** 6 mois

**Date de début prévisionnelle :** février/mars 2024

**Lieu :** laboratoire MaIAGE, centre INRAE de Jouy-en-Josas (78)

**Gratification :** selon la grille réglementaire

Ce travail a bénéficié d'un financement ANR PIA : ANR-20-IDEES-0002

Pour postuler, veuillez transmettre CV et courte lettre de motivation à :

- Arnaud Ferré ([arnaud.ferre@inrae.fr](mailto:arnaud.ferre@inrae.fr)), chercheur en TAL à MaIAGE, équipe Bibliome
- Alexandre Kabil ([alexandre.kabil@cnrs.fr](mailto:alexandre.kabil@cnrs.fr)), ingénieur de recherche au LISN, équipe Ex)Situ

**Bibliographie**

- **[Deléger2016]** Deléger, L., Bossy, R., Chaix, E., Ba, M., Ferré, A., Bessieres, P., & Nédellec, C. (2016, August). Overview of the bacteria biotope task at bionlp shared task 2016. In Proceedings of the 4th BioNLP shared task workshop (pp. 12-22).
- **[David2017]** David, L., Bouyer, G., & Otmane, S. (2017, March). Towards an upper limb self-rehabilitation assistance system after stroke. In Proceedings of the Virtual Reality International Conference-Laval Virtual 2017 (pp. 1-4).
- **[Mayer2023]** Mayer, R. E., Makransky, G., & Parong, J. (2023). The promise and pitfalls of learning in immersive virtual reality. *International Journal of Human-Computer Interaction*, 39(11), 2229-2238.
- **[Spiekermann2018]** Spiekermann, C., Abrami, G., & Mehler, A. (2018, May). VAnnotatoR: a gesture-driven annotation framework for linguistic and multimodal annotation. In Proceedings of the Annotation, Recognition and Evaluation of Actions (AREA 2018) Workshop, AREA.
- **[Borhani2023]** Borhani, Z., Sharma, P., & Ortega, F. R. (2023). Survey of annotations in extended reality systems. *IEEE Transactions on Visualization and Computer Graphics*.
- **[Colucci2024]** Colucci Cante, L., D'Angelo, S., Di Martino, B., & Graziano, M. (2024, July). Text Annotation Tools: A Comprehensive Review and Comparative Analysis. In International Conference on Complex, Intelligent, and Software Intensive Systems (pp. 353-362). Cham: Springer Nature Switzerland.