# Analyse numérique de l’engagement d’apprenants dans des situations d’apprentissage en ligne

# Contexte du projet scientifique

Ce travail s’inscrit dans un projet plus large conduit au CReSTIC dans le cadre la campagne "Compétences psychosociales" (CPS) de Santé Publique France [7] menée par huit ministères signataires du projet [2], et à laquelle le CReSTIC participe, notamment avec l’Observatoire du Bien-Etre de l’Université de Lyon 2.

Ce stage de Master se situe en amont et en parallèle des recherches menés dans le cadre de cette campagne et consiste à choisir et mettre en place les outils adéquats en analyse statistique de données et d’extraction de connaissances à partir de traces d’interaction avec un site d’apprentissage en ligne afin de mesurer l’impact d’outils de formation en ligne et de mesurer l’engagement des apprenants dans leur démarche.

En effet, la campagne CPS prévoit le développement d’une plateforme numérique d’apprentissage qui s’adressera à deux types de personnes : les adultes en position d’éducation (parents, enseignants, éducateurs, infirmiers, …), et les jeunes de 16 à 25 ans issus de tous milieux sociaux.

Notre rôle est de recueillir, avec le consentement des usagers, les traces d’activité sur cette plateforme, et de construire des indicateurs centrés sur l’apprentissage.

Le but est d’établir des profils d’usagers, afin d’évaluer l’adéquation de l’outil numérique proposé, et de faire des préconisations pour son évolution. Nous disposons de données anonymisées recueillis dans un contexte similaire et qui est représentatif de l’usage de ce type de plateforme. Le travail sur ces données, mené dans le cadre de ce stage de Master, servira de base pour engager les travaux

# Sujet de stage

* Il s’agira de concevoir et appliquer des outils d’analyse statistique des données et d’extraction de connaissances à partir des données de cette expérimentation.
* Des algorithmes d’analyse de données et de classification seront utiles pour mettre en exergue des corrélations, interdépendances ou autres similarités dans les parcours d’apprentissage.

Dans la mesure où l’environnement relève de recherche auprès de populations vulnérables (des enfants) et qu’une partie des données recueillies pourrait être qualifié de personnelles, les travaux seront soumis à la règlementation en vigueur sur la protection des personnes et au RGPD et seront par conséquent validés sur leur conformité par les comités *ad hoc* requis.

## *Support théorique*

Le domaine des *Learning Analytics* se situe au croisement entre l’Informatique et les Sciences de l’Education et consiste à mettre en place des outils d’observation et de mesure permettant de quantifier et d’analyser les impacts de l’environnement pédagogique sur la qualité de l’apprentissage des apprenants [1, 3, 5]. La partie Informatique concerne donc tous les processus, méthodes et outils (physiques et logicielles) permettant de capter (vidéo, audio, texte) des données pertinentes à l’analyse du processus d’apprentissage, ainsi que tous les algorithmes et méthodes contribuant à fournir une compréhension fine, mesure et analyse du processus en question [6, 8].

Il s’agira donc de développer les méthodes d’Apprentissage Automatique sur les données brutes captées qui permettront de classifier les situations d’apprentissage observées en différents niveaux de succès, ou éventuellement de faire des prédictions sur les étapes pédagogiques correctives à mettre en place [4, 9].

En contribution à une nouvelle branche scientifique qui est l’*Engagement Analytics*, il s’agira de mesurer l'engagement des apprenants, et de réfléchir à un protocole d'évaluation de ces mesures.

# Objectifs

Réalisation, documentation et évaluation d’algorithmes d’analyse de données, de classification et de façon générale d’apprentissage non-supervisé ou semi-supervisé. Analyse de la qualité des résultats obtenus et préconisation d’améliorations ou de développements de nouveaux algorithmes d’apprentissage et de classification.

# Période de stage

De février 2025 à septembre 2025 en fonction des contraintes/agenda des candidates et candidats.

# Lieu de stage

Université de Reims Champagne-Ardenne, CReSTIC (Centre de Recherche en Sciences et Technologies de l’Information et de la Communication – <https://crestic.univ-reims.fr>)

# Gratification

La ou le stagiaire percevra une gratification de stage selon les règles en vigueur relatifs à l’encadrement et la gratification de stages.

# Prérequis

La candidate ou le candidat a une bonne connaissance en analyse de données numérique et de programmation. Elle ou il maîtrise Python ou R.  
La curiosité, l’autonomie et le sens de travail en équipe seront également des facteurs déterminants.

# Références Bibliographiques

1. Alwahaby, H., Cukurova, M., Papamitsiou, Z., Giannakos, M. (2022). The Evidence of Impact and Ethical Considerations of Multimodal Learning Analytics: A Systematic Literature Review. In: Giannakos, M., Spikol, D., Di Mitri, D., Sharma, K., Ochoa, X., Hammad, R. (eds) The Multimodal Learning Analytics Handbook. Springer, Cham. <https://doi.org/10.1007/978-3-031-08076-0_12>
2. Bulletin officiel Santé - Protection sociale - Solidarité n° 2022/18 du 31 août 2022 <https://sante.gouv.fr/fichiers/bo/2022/2022.18.sante.pdf>
3. Cerratto Pargman T, McGrath C. Mapping the Ethics of Learning Analytics in Higher Education : A Systematic Literature Review of Empirical Research. Journal of Learning Analytics [Internet]. 2021;8(2):123–39. Available from: <http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-192646>
4. Friend Wise, A., Sarmiento, J.P., Boothe, M. Jr. (2021). Subversive Learning Analytics. In LAK21: 11th International Learning Analytics and Knowledge Conference (LAK21). Association for Computing Machinery, New York, NY, USA, 639–645.

<https://doi.org/10.1145/3448139.3448210>

1. Guzmán-Valenzuela, C., Gómez-González, C., Rojas-Murphy Tagle, A. *et al.* Learning analytics in higher education: a preponderance of analytics but very little learning? *Int J Educ Technol High Educ* **18**, 23 (2021). <https://doi.org/10.1186/s41239-021-00258-x>
2. Kew, S.N., Tasir, Z. Learning Analytics in Online Learning Environment: A Systematic Review on the Focuses and the Types of Student-Related Analytics Data. *Tech Know Learn* **27**, 405–427 (2022). <https://doi.org/10.1007/s10758-021-09541-2>
3. Lamboy, B., Arwidson, P., du Roscoät, E., Fréry, N., Lecrique, J.-M., Shankland, R., Tessier, D., Williamson M.-O. (2021). Les compétences psychosociales : un référentiel pour un déploiement auprès des enfants et des jeunes. Synthèse de l'état des connaissances scientifiques et théoriques réalisé en 2021. Santé Publique France. <https://www.santepubliquefrance.fr/docs/les-competences-psychosociales-un-referentiel-pour-un-deploiement-aupres-des-enfants-et-des-jeunes.-synthese-de-l-etat-des-connaissances-scientif>
4. Lemay, D.J., Baek, C., Doleck, T. (2021). Comparison of learning analytics and educational Data mining: A topic modeling approach, Computers and Education: Artificial Intelligence, Volume 2, 100016, ISSN 2666-920X, <https://doi.org/10.1016/j.caeai.2021.100016>.
5. Shorfuzzaman, M., Hossain, M.S., Nazir, A., Muhammad, G., Alamri, A. (2019). Harnessing the power of big data analytics in the cloud to support learning analytics in mobile learning environment, Computers in Human Behavior, Volume 92, Pages 578-588, ISSN 0747-5632, <https://doi.org/10.1016/j.chb.2018.07.002>.