

Titre. Technologies intelligentes pour le débriefing médical

Note : il est obligatoire que le candidat dispose d'un niveau satisfaisant en français car les interactions avec l'équipe médicale se feront essentiellement dans cette langue. Les entretiens de sélection se feront essentiellement dans cette langue.

Laboratoire hôte : Laboratoire d'Informatique de l'Université du Mans (LIUM), 20 Avenue René Laennec, 72000 Le Mans.

- **Encadrants :** Emmanuel G. BLANCHARD, Sébastien IKSAL

Entreprise hôte : Centre Hospitalier du Mans (CHM), 94 Av. Rubillard, 72037 Le Mans Cedex 9, France.

- **Encadrants :** Jean-Christophe CALLAHAN, Christophe GUITTON

Financement : Cette thèse est conditionnée à l'obtention d'un financement CIFRE (thèse industrielle)

Mots-clés. Environnement Informatique pour l'Apprentissage Humain (EIAH), Artificial Intelligence in Education (AIED), Simulation Virtuelle Rapide (SVR), débriefing médical, data visualisation, agent pédagogique autonome, ingénierie de données, indicateurs pédagogiques, ingénierie d'ontologies, Large Language Model (LLM), Retrieval Augmented Generation (RAG), Analytiques éducatives

Nb. Dans ce qui suit, le masculin est employé comme genre neutre.

Résumé. Un bon système de santé est l'un des éléments les plus fondamentaux de toute société développée. Cela passe par des professionnels (médecins et autres praticiens) bien formés et en nombre suffisant. Or, dans de nombreux pays, un nombre trop faible de professionnels ont été formés dans les dernières décennies pour assurer le renouvellement normal des praticiens tout en satisfaisant les besoins grandissants des populations. Différents verrous ont été levés récemment (en France, fin du numérus clausus pour les médecins (Roberto, 2022) mais cela n'a fait que déplacer le problème. L'afflux de nouveaux candidats se heurte au nombre limité d'experts pour les former sachant que ceux-ci doivent en parallèle assurer leur mission de soin. Les médecins experts sont donc parmi les professionnels les plus occupés qui soient.

Les Dr Blanchard (Université du Mans, France) et Wiseman (université McGill, Canada) travaillent depuis plus de 10 ans sur la conception de technologies adaptées aux caractéristiques et contraintes propres aux professionnels médicaux (Blanchard, Wiseman, Lajoie, 2010 ; Wiseman, Blanchard, Lajoie, 2016). Cela a récemment abouti à l'émergence du concept de **Simulations Virtuelles Rapides (SVR)**, (Blanchard & Wiseman, 2024) qui s'articule autour d'une philosophie dite « suffisamment réaliste » visant à rechercher la complexité minimale (technologie, représentation des connaissances...) permettant d'obtenir un impact d'apprentissage satisfaisant. Cette démarche et les technologies éducatives qui en découlent ont déjà séduit plusieurs centres de formation médicaux parmi lesquels le Centre de Simulation Médical d'Angers et le Centre de Simulation Médicale Steinberg de l'université McGill au Canada.

Le **Centre Hospitalier du Mans (CHM)** s'y intéresse également tout en amenant des challenges additionnels. Étant donné que cette institution n'est pas universitaire, ses experts médicaux

n'ont aucun temps réservé pour l'enseignement tout en ayant à encadrer de nombreux étudiants à divers niveaux de leur formation. Dans ce contexte, le CHM a identifié l'adoption de SVR comme une opportunité d'améliorer qualitativement et quantitativement son offre de formation et l'accompagnement des étudiants qui lui sont confiés.

Le Laboratoire d'Informatique de l'Université du Mans (LIUM) et le CHM se sont donc naturellement rapprochés avec comme objectif commun de développer des modules permettant d'améliorer la phase de débriefing médical suivant un enseignement à base de SVR. Cette thèse portera donc sur ce sujet et sera découpée en deux phases :

- **Phase 1. Technologies pour le débriefing de groupe.** L'objectif est d'aider un enseignant médical à rapidement interpréter la performance d'un groupe d'étudiants à qui aurait été assigné une formation au moyen de SVR. En collaboration avec les experts médicaux, cette phase doit permettre d'identifier des métriques et indicateurs permettant de caractériser différentes dimensions de la performance d'un groupe et, in fine, à quel point l'enseignement par SVR a été un succès. Les travaux se situeront essentiellement dans les domaines de l'ingénierie, l'analyse et la visualisation de données (Amburst et al., 2021 ; Paiva et al., 2018) sans nécessairement s'y restreindre.
- **Phase 2. Technologies pour le débriefing autonome.** Sur la base des métriques et indicateurs identifiés dans la phase 1, l'objectif est de développer une technologie de type agent pédagogique autonome permettant de remplacer ponctuellement un médecin dans des tâches de débriefing. Les travaux se situeront dans le domaine de l'intelligence artificielle générative et pourraient faire appel à concepts tels que les modèles de langage (Naveed et al, 2023 ; Magister et al., 2023), le Retrieval Augmented Generation (Lewis et al., 2020), les flux de travail à base d'agents (Yang et al., 2023). Cependant, l'usage d'une technique en particulier n'est pas arrêté à l'heure où nous écrivons ces lignes, étant donnée la rapidité extrême des évolutions dans ce domaine.

Références

- Armbrust, M., Ghodsi, A., Xin, R., & Zaharia, M. (2021). Lakehouse: a new generation of open platforms that unify data warehousing and advanced analytics. In *Proc. of the Conference on Innovative Data Systems Research (CIDR2021)*, vol. 8, pp. 28-36.
- Blanchard, E.G., Wiseman, J. (2024). Rapid Virtual Simulations: Achieving 'Satisficing Learning Impact' with 'Realistic-Enough' activities in Health Science Education. <https://doi.org/10.48550/arXiv.2407.05179>
- Naveed, H., Khan, A., Qiu, S., Saqib, M., Anwar, S., Usman, M., Barnes, N., Mian, A. (2023). A comprehensive overview of large language models. <https://doi.org/10.48550/arXiv.2307.06435>
- Roberto, A. (2022). La fin du numérus clausus n'aura pas d'effet avant 2035. Fondation pour la recherche sur les administrations et les politiques publiques. <https://www.ifrap.org/emploi-et-politiques-sociales/la-fin-du-numerus-clausus-naura-pas-deffets-avant-2035>. Dernier accès le 21 juillet 2024.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-T., Rocktäschel, T., Riedel, S., Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP tasks. In *Procs of Neural Information Processing Systems (NeurIPS2020)*, vol. 33, pp. 9459–9474.
- Magister, L.C., Mallison, J., Adamek, J., Malmi, E., Severy, A. (2023). Teaching Small Language Models to Reason. In *Proc. of the 61st Annual Meeting of the Association for Computational Linguistics, vol. 2*, pp. 1773–1781.
- Paiva, R., Bittencourt, I. I., Lemos, W., Vinicius, A., & Dermeval, D. (2018). Visualizing learning analytics and educational data mining outputs. In *Procs of the 19th Int. Conf on Artificial Intelligence in Education (AIED2018), Part II*, pp. 251-256. Springer International Publishing.
- Yang, Z., Li, L., Wang, J., Lin, K., Azarnasab, E., Ahmed, F., Liu, Z., Liu, C., Zeng, M., Wang, L. (2023). MM-REACT: Prompting ChatGPT for Multimodal Reasoning and Action. <https://doi.org/10.48550/arXiv.2303.11381>

Programme de travail envisagé. La première partie de cette thèse sera dévolue à la prise en main du sujet par le doctorant. Un gros effort sera porté sur la réalisation d'un état de l'art investiguant les interconnexions entre les différentes dimensions du domaine de recherche interdisciplinaire abordé dès la phase 1, à savoir (a) l'éducation et le débriefing médical, (b) l'intelligence artificielle pour l'éducation ainsi que (c) l'ingénierie, l'analyse et la visualisation des données... Ces travaux permettront de raffiner l'objectif de la phase 1 pour en faire ressortir une problématique de recherche bien définie et proposer une solution qui sera alors implémentée. Suite à la réalisation de ce premier travail, un nouvel effort d'état de l'art portera cette fois sur les domaines des agents pédagogiques et les évolutions récentes (à ce moment-là) de l'intelligence artificielle générative. Ces travaux permettront cette fois de raffiner l'objectif de la phase 2 pour en faire ressortir une problématique de recherche bien définie et proposer une solution qui sera alors implémentée. La dernière partie de cette thèse sera centrée sur la rédaction de la thèse et sa soutenance par le candidat.

Profil recherché. Le candidat idéal doit être une personne motivée, dynamique et titulaire d'un diplôme d'ingénieur ou de master 2 en informatique (ou diplômé bac+5 pouvant justifier de compétences solides dans ce domaine). De plus, cette thèse se positionnant dans des domaines interdisciplinaires (EIAH, AIED, informatique cognitive, médecine), l'intérêt du candidat pour une ou plusieurs disciplines des Humanités sera apprécié, sans pour autant que ce critère impose une formation *a priori*. En dernier lieu, si la maîtrise du français écrit et orale est obligatoire pour candidater à ce poste, un très bon niveau en anglais écrit et une expérience passée dans l'écriture d'articles scientifiques sont des atouts.

Candidature. Le dossier de candidature doit comporter les éléments suivants : (i) un CV, (ii) une lettre de motivation, (iii) les notes de Master 1 et de Master 2 (ou équivalent), (iv) une ou plusieurs lettres de recommandation.

Pour soumettre votre dossier, envoyez un email à emmanuel.blanchard@univ-lemans.fr avec pour sujet : **Candidature SVR+AIED**. La date limite de soumission est le **10 septembre 2024**.