



PEPR eNSEMBLE 2024 - Proposition de sujet de thèse

Acronyme/Titre	CIAGE : Collaborer avec une Intelligence Artificielle Générative en Éducation
Institution d'accueil	IMT Atlantique campus de Brest / Lab-STICC / équipe MOTEL
Direction / encadrement	<i>Directeurs</i> : Jean-Marie Gilliot et Grégory Smits <i>Encadrement</i> : Maria-Teresa Segarra Montesinos

Contexte du projet de recherche

De nombreux outils d'IA ont atteint un niveau de maturité suffisant pour être intégrés dans nos activités quotidiennes, qu'elles soient professionnelles ou personnelles. Ces outils modifient profondément la nature de nos méthodes de travail et les résultats que nous produisons. L'usage d'IA Génératives (IAG) s'est démocratisé très rapidement et conduit à d'importantes réflexions sur leur intégration en tant que collaborateur pour la production de contenus. L'éducation est évidemment un domaine qui a été profondément bouleversé par l'arrivée de ces technologies sous forme d'assistantes « omniscientes » pour accompagner les apprenants dans des tâches aussi diverses que la production d'écrits, la résolution d'exercices, l'apprentissage de langues étrangères, la programmation, etc. [1,2]. Il paraît désuet de freiner voire d'interdire l'usage de ces outils. Cependant, les méthodes pédagogiques doivent évoluer pour prendre en compte ces IAG comme des outils incontournables dont l'utilisation doit être observée, analysée puis guidée. Ce projet vise à considérer la maîtrise des IAG dans un contexte pédagogique comme une compétence importante à acquérir pour aboutir à des productions, éventuellement plus complètes et de meilleure qualité, sans négliger les objectifs pédagogiques des situations d'apprentissage dans lesquelles elles sont utilisées. Dans la plupart des situations d'apprentissage, les apprenants sont évalués sur la base de ce qui a été produit, ce livrable pouvant désormais inclure une proportion non négligeable de contenu généré automatiquement par une IA. **Un enjeu majeur en éducation est d'analyser les mécanismes d'interaction et de collaboration entre apprenants et IAG [1]. L'étude des différentes stratégies collaboratives permettrait de mieux appréhender l'appropriation du contenu généré par l'IAG et donc l'atteinte des objectifs pédagogiques visés au sein d'une équipe d'apprenants.**

Objectifs scientifiques et techniques

L'objectif du projet est d'avoir une **meilleure compréhension de la place prise par les IAG dans les activités pédagogiques en groupe** et d'**évaluer l'impact de leur utilisation dans l'acquisition des objectifs d'apprentissage associés** [1]. Plus précisément, nous considérons la production de code informatique, tâche pour laquelle les outils d'aide à la programmation (Github Copilot, DataCamp, Blackbox, Chat GPT, etc.) prennent une place de plus en plus importante. Ces deux axes d'analyse sont évidemment liés car, dans un contexte pédagogique, l'apport d'un outil dépend de son usage et de sa capacité à accompagner ou accélérer l'acquisition de compétences. Pour atteindre cet objectif, il faudra tout d'abord disposer :

- D'outils pour tracer les interactions entre IAG et apprenants, ainsi que pour tracer les contributions (ajout, modification, suppression) de l'IAG et de l'apprenant sur le code délivré,
- D'indicateurs pour transformer ces traces en données exploitables qui caractérisent la nature des interactions avec l'IAG (validation, correction, production) et des modifications apportées à la production,
- De méthodes d'analyse des données (analyse, classification et prédiction de séquences) pour les transformer en connaissances sur les différents schémas de collaboration entre apprenants et IAG.

Approche et positionnement

L'étude des interactions entre apprenants et IAG reposera sur l'analyse de traces générées par un module intégré à un éditeur de programmation (module en cours de développement au Lab-STICC). Le recours à des méthodes de fouille de données et d'apprentissage automatique pour analyser et comprendre le comportement d'apprenants est une démarche classique dans le domaine des *Learning Analytics* [3]. L'analyse automatique des traces aura pour objectif l'identification des différents types de comportement collaboratif avec une IAG [4], mais également l'évaluation de l'appropriation du contenu issu de l'IAG par l'apprenant sur la base des modifications qu'il effectuera pour intégrer ce contenu au reste du code [5]. Considérant disponible une évaluation finale des compétences des apprenants, un modèle prédictif d'apprentissage automatique sera exploité pour identifier l'impact des usages (variables observées) des IAG sur l'acquisition des compétences (variable prédite) [1]. **L'originalité du projet est de s'intéresser à la fois aux interactions avec l'outil d'IAG ainsi qu'aux modifications faites sur le contenu qu'il génère pour caractériser l'usage qu'il en est fait. Cette direction de recherche s'inscrit pleinement dans le projet 3 sur l'étude des collaborations avec des systèmes intelligents et, plus précisément, dans les thématiques 2 et 3 à travers l'analyse des interactions entre une IAG et un apprenti développeur, ainsi que l'évaluation de l'impact des IAG sur l'acquisition de compétences de programmation.**

Organisation du projet, durée, jalons

Le projet est organisé autour de quatre grandes étapes : 1) Mise en place des contextes d'expérimentation pour générer les traces d'interaction et de travail collaboratif entre un apprenti développeur et une IAG (T0 à T6, livrable : open data). 2) Analyse des traces pour identifier et caractériser les profils de collaboration IAG-apprenant (T6 à T18, livrable : publication dans les conférences EDM ou ECTEL). 3) Étude de l'impact des profils de collaboration sur l'acquisition de compétences (T18 à T30, livrable : publication dans un journal tel qu'AIED). 4) Valorisation et finalisation du projet (T30 à T36, livrable : thèse).

Partenariat : présentation et rôle des co-encadrants

L'équipe MOTEL du Lab-STICC est spécialisée dans l'analyse de traces d'éducation en vue d'améliorer l'expérience d'apprentissage. Le contexte d'expérimentation dans lequel les traces seront captées sera issu d'enseignements gérés par les membres de l'équipe au sein de l'IMT Atlantique. G. Smits apporte dans cet encadrement son expertise en apprentissage automatique et en fouille de données, compétences nécessaires pour traduire les traces en connaissances sur les interactions entre apprenants et outils d'IAG..

Références

- [1] A. Tlili *et al.*, "What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education," *Smart Learn. Environ.*, vol. 10, no. 1, p. 15, 2023.
- [1] L. Yan, R. Martinez-Maldonado, and D. Gasevic, "Generative artificial intelligence in learning analytics: Contextualising opportunities and challenges through the learning analytics cycle," in *Proceedings of the 14th Learning Analytics and Knowledge Conference*, 2024, pp. 101–111.
- [2] W. C. H. Hong, "The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research," *J. Educ. Technol. Innov.*, vol. 5, no. 1, 2023.
- [3] A. Barthakur, S. Dawson, and V. Kovanovic, "Advancing learner profiles with learning analytics: A scoping review of current trends and challenges," in *LAK23: 13th International Learning Analytics and Knowledge Conference*, 2023, pp. 606–612.
- [4] A. Shibani, R. Rajalakshmi, F. Mattins, S. Selvaraj, and S. Knight, "Visual Representation of Co-Authorship with GPT-3: Studying Human-Machine Interaction for Effective Writing.," *Int. Educ. Data Min. Soc.*, 2023.
- [5] A. Haddouche, F. Djelil, C. Hoffmann, N. Mandran, and C. D'ham, "Proposal of Indicators for Measuring Collaborative Writing in a Digital Learning Environment," in *15th International Conference on Computer Supported Education*, SCITEPRESS-Science and Technology Publications, 2023, pp. 495–502.