Thèse CIFRE entre l'équipe MOCAH-LIP6 (Sorbonne Université) et l'entreprise OuiActive

Adaptation dynamique des interactions d'un agent conversationnel (LLM) par le couplage d'un modèle de l'apprenant et d'un modèle praxéologique du domaine : une approche hybride pour un étayage pédagogique personnalisé

Mots-clés : Agents conversationnels, Modèles de langage (LLMs), Modèle de l'apprenant, Modèle praxéologique du domaine, Systèmes Tuteurs Intelligents (ITS), Étayage adaptatif, AIED.

1. Contexte

L'émergence des grands modèles de langage (LLMs) offre des perspectives intéressantes pour le développement d'agents conversationnels (chatbots) éducatifs capables d'interagir en langage naturel de manière fluide, réactive et engageante. Contrairement aux interfaces traditionnelles des environnements numériques d'apprentissage, un chatbot basé sur un LLM peut adapter son discours, répondre à des questions et fournir des explications détaillées à la demande de l'apprenant.

Cependant, les chatbots fondés sur des LLM généralistes, reposant principalement sur des techniques de *prompting*, présentent des capacités limitées à adapter leurs interventions aux caractéristiques de chaque apprenant et aux contraintes propres au domaine des connaissances. Ils peuvent ainsi conduire à une forme d'accompagnement relativement générique, susceptible de s'apparenter à une approche *one-size-fits-all* (Cohn et al., 2026 ; Shi et al., 2025). Dans cette perspective, l'intégration de composants issus des systèmes tuteurs intelligents (ITS), notamment un modèle du domaine et un modèle de l'apprenant, constitue une piste pour dépasser cette limitation. Le modèle du domaine permet de structurer les connaissances et les contraintes propres à la discipline, tandis que le modèle de l'apprenant permet de représenter son niveau de maîtrise, ses connaissances, ses difficultés et leur évolution au cours de l'apprentissage. La combinaison de ces deux modèles peut ainsi permettre au LLM de sélectionner et d'adapter ses interventions en fonction de l'état de l'apprenant et du contenu à enseigner, plutôt que de s'appuyer sur une stratégie pédagogique unique définie principalement au niveau du prompt (Kong et al., 2026; Park et al., 2024).

L'agent conversationnel actuellement implémenté au sein du système OuiActive s'inscrit dans une démarche socratique pour favoriser l'accompagnement par rapport à la délivrance immédiate de la réponse, l'enjeu est d'enrichir cette première démarche pour parvenir à un système articulant, de manière à guider et contraindre l'agent conversationnel, le modèle du domaine (modèle praxéologique) exploité dans l'environnement, un modèle de l'apprenant (à construire) et les interactions en temps réel de l'apprenant avec le système. Ce travail bénéficiera de l'existence d'un environnement déjà existant, largement déployé et d'un volume significatif de traces déjà existant.

2. Problématique

La question centrale de cette thèse est la suivante :

> **Comment coupler un modèle dynamique de l'apprenant et un modèle praxéologique du domaine avec un modèle de langage (LLM) afin d'adapter dynamiquement les interactions conversationnelles (aides, contenus) au profil et aux besoins de l'apprenant sur la plateforme OuiActive ?**

Cette problématique se situe à la croisée de trois disciplines :

- L'Intelligence Artificielle en Éducation (AIED) : architectures de Systèmes Tuteurs Intelligents (ITS) et modélisation de l'apprenant ;
- Le Traitement Automatique du Langage Naturel (TALN / IA générative) : contrôle, contrainte et conditionnement des LLMs (RAG, IA agentique, alignement) ;
- La didactique : représentations des savoirs et à la relation aux contenus d'un domaine de connaissances (mathématiques, sciences, etc.)

3. Objectifs de la thèse

Cette thèse vise quatre objectifs principaux :

1. Conception et structuration du modèle du domaine et du modèle de l'apprenant

Développer une représentation explicite du domaine (graphe de concepts, erreurs fréquentes, tâches) et implémenter un module de suivi dynamique de l'apprenant (*Knowledge Tracing*) alimenté par les interactions conversationnelles et l'ensemble des activités des apprenants sur la plateforme de OuiActive. Un travail d'ingénierie des traces d'interaction de la plateforme devra être mené.

2. Architecture de couplage entre les modèles du domaine et de l'apprenant et le LLM

Concevoir un mécanisme d'interfaçage (par exemple via du RAG, des frameworks d'agents ou du prompting structuré) permettant au LLM d'interroger le modèle de l'apprenant et le modèle du domaine avant de formuler une réponse ou une suggestion.

3. Génération d'interventions et d'aides pédagogiques adaptatives

Développer une politique d'intervention capable de moduler l'étayage fourni par le chatbot (de l'indice à l'explication guidée) en fonction de l'écart mesuré entre le profil de l'apprenant et le modèle du domaine.

4. Évaluation expérimentale de l'efficacité pédagogique et de la qualité des interactions

Mesurer l'impact du système adaptatif sur la maîtrise des concepts, la réduction des erreurs récurrentes et l'expérience utilisateur par rapport à un chatbot LLM générique non conditionné, appliqué à deux disciplines différentes.

4. Cadre théorique

Ce travail s'appuie sur trois cadres fondamentaux :

- **Architecture classique des Systèmes Tuteurs Intelligents (ITS)**

L’architecture d’un ITS (Wenger, 1987; VanLehn, 2011) articulant modèle apprenant, modèle du domaine et modèle tuteur sert de socle pour structurer les entrées et contraintes fournies au LLM.

- **Théorie Anthropologique du Didactique (TAD) et Modélisation Praxéologique**

La notion de praxéologie (Chevallard, 1999) fournit un cadre de formalisation rigoureux pour le Modèle du Domaine et la structuration du savoir. Elle permet de modéliser l'activité de l'apprenant sous la forme de concepts tels que les types de tâches (ce qui est à réaliser), les techniques (comment le réaliser) et le logos (qui justifie et explique les techniques). Ce cadre permet de contraindre le LLM pour qu'il identifie précisément si le besoin d'aide porte sur la technique à appliquer ou sur la théorique sous-jacente.

- **Knowledge Tracing**

Les modèles de suivi des connaissances (ex. *Bayesian Knowledge Tracing*, *Deep Knowledge Tracing*) permettent d'actualiser la probabilité de maîtrise d'un concept en temps réel à partir de l'analyse des réponses textuelles et comportementales des apprenants.

5. Candidature

- **Encadrement scientifique** : Sébastien Jolivet, Vanda Luengo et Amel Yessad
- **Lieu de la thèse** : Entreprise OuiActive (Givors - Lyon Sud) et LIP6 (campus Jussieu - Paris)
- **Financement** : Contrat CIFRE
- **Début de thèse** : au printemps 2027, en fonction de l’avancement du processus de traitement par l’ANRT
- **Profil recherché** : Étudiants en Master 2 d'Informatique ou élèves-ingénieurs en dernière année (BAC+5), notamment en spécialité Intelligence Artificielle, TALN et Machine Learning et intéressés par l’IA en éducation
- **Dossier de candidature :** envoyer un CV, une lettre de motivation ainsi que les relevés de notes du M1 et M2 à Vanda Luengo (vanda.luengo@lip6.fr), Sébastien Jolivet (Sebastien.Jolivet@unige.ch), Amel Yessad (amel.yessad@lip6.fr) et Yadem Benmessaoud (info@ouiactive.com)
- **Date limite de candidature :** 15 décembre 2026

Références bibliographiques

Bosch, M., Chevallard, Y., García, F. J., & Monaghan, J. (Éds.). (2019). *Working with the Anthropological Theory of the Didactic in Mathematics Education : A Comprehensive Casebook*. Routledge.

Brusilovsky, P. (2001). User modeling and user-adapted interaction.

Cohn, C., Guo, S., Rayala, S., Wang, H. D., Mohammed, N., Timalsina, U., ... & Biswas, G. (2026, June). Evidence-decision-feedback: Theory-driven adaptive scaffolding for llm agents. In *International Conference on Artificial Intelligence in Education* (pp. 1-16). Cham: Springer Nature Switzerland

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, *4*(4), 253-278.

Chevallard, Y. (1999). L’analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en Didactique des Mathématiques*, *19*(2), 221-265.

Mazzullo, E., Bulut, O., Wongvorachan, T., & Tan, B. (2023). Learning analytics in the era of large language models. *Analytics*, *2*(4), 877-898.

Kong, L., Sun, J., Zhou, J., & Ng, V. LLM-based intelligent tutoring systems: A survey.

Park, M., Kim, S., Lee, S., Kwon, S., & Kim, K. (2024, May). Empowering personalized learning through a conversation-based tutoring system with student modeling. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1-10).

Shi, Y., Yu, K., Dong, Y., & Chen, F. (2025). Large language models in education: a systematic review of empirical applications, benefits, and challenges. *Computers and Education: Artificial Intelligence*, 100529.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational psychologist*, *46*(4), 197-221.

Wenger, E. (2004). Artificial intelligence and tutoring systems. *international Journal of Artificial intelligence in Education*, *14*, 39-65.

Woolf, B. P. (2010). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann.