



Enseignement

VIP-Bot :

un agent conversationnel pédagogique ancré dans les contenus des cours

Permettre aux enseignantes et aux enseignants de créer et de déployer des agents conversationnels entraînés exclusivement sur leurs propres ressources et activités pour réduire les hallucinations d'IA¹ et renforcer l'alignement pédagogique.

Type de cas

Enseignement, recherche.

Groupe de recherche

VIP Research Group, dirigé par le professeur Maiga Chang

 Université d'Athabasca (Alberta, Canada)

[Lien](#) 



Potentiel de transférabilité

Élevé : architecture modulaire à code source ouvert⁶, applicable à divers contextes institutionnels et disciplinaires, en formation à distance comme en présentiel. La transférabilité dépend de la qualité des contenus fournis et d'un accompagnement adéquat.



Coûts

Modérés ou élevés, principalement en ressources humaines (programmation, recherche, tests, validation). Des frais récurrents modérés sont à assumer pour l'infrastructure de serveurs et les coûts d'inférence⁷, variables selon le degré d'utilisation.



Niveau de dif culté

Modéré ou élevé : aucune expertise avancée en IAG nécessaire chez le personnel enseignant pour la configuration de la plateforme. Toutefois, le fonctionnement du système repose sur des mécanismes techniques moyennement complexes (indexation, vectorisation sémantique⁸ des contenus, maintenance de l'infrastructure).

Objectif

Permettre aux enseignantes et aux enseignants sans expertise avancée en IAG² de configurer et de déployer des agents conversationnels³ pédagogiques entraînés exclusivement sur les ressources et les activités de leurs propres cours, grâce à une architecture de GAR⁴, afin de réduire les hallucinations d'IA, les biais et le désalignement pédagogique associés au recours aux GML⁵ généralistes. La recherche vise évaluer la perception des utilisatrices et des utilisateurs (personnel enseignant et population étudiante) à l'égard des réponses produites, en comparaison des outils d'IAG généralistes, et à identifier les forces et les limites techniques du système afin d'orienter ses développements futurs, notamment en matière de qualité des réponses et de performance.

Contact



Maiga Chang

Professeur, Faculté des Sciences et Technologies, Université d'Athabasca

Référence

Nikou, S. A., Guliya, A., Van Verma, S. et Chang, M. (2024). A generative artificial intelligence empowered chatbot: System usability and student teachers' experience. *International Conference on Intelligent Tutoring Systems (ITS)*. Cham : Springer Nature Switzerland. 

Public cible

Personnes étudiantes en formation à distance dans des établissements d'enseignement postsecondaire, en particulier les futures personnes enseignantes souhaitant déployer un agent pédagogique balisé sur les contenus et activités de leurs cours.

Contexte

VIP-Bot (Verified Information Platform Bot) s'inscrit dans le cadre du projet de recherche ChatbotLLM⁹, une plateforme institutionnelle développée à l'Université d'Athabasca et permettant aux enseignantes et aux enseignants de créer des agents conversationnels éducatifs à partir de leurs propres cours. Le dispositif a été expérimenté à l'université par de futures enseignantes et futurs enseignants. À l'instar d'autres initiatives institutionnelles telles que YU AURA ou Piccolo, développées dans des contextes semblables, VIP-Bot fait partie d'une démarche visant à proposer des agents fiables tout en encadrant les usages de l'IAG et en conservant un contrôle sur les données, les contenus et les finalités pédagogiques.

État d'avancement

Projet pilote réalisé auprès d'étudiantes et d'étudiants, en 2023-2024, dans le cadre d'un cours de formation pédagogique. Expérimentation comparative menée présentement à l'Université de Strathclyde (Royaume-Uni).

Résumé



VIP-Bot (Verified Information Platform Bot) est un agent conversationnel pédagogique déployé à partir de la plateforme de recherche ChatbotLLM, conçu en réponse aux limites des GML généralistes et développé à l'Université d'Athabasca sous la direction du professeur Maiga Chang. Prenant la forme d'une application Web, il permet aux enseignantes et aux enseignants de créer des agents conversationnels ancrés dans les ressources et les activités de leurs propres cours sans avoir besoin d'une expertise avancée en IAG.

L'enseignante ou l'enseignant constitue le corpus de son cours et configure les règles d'interrogation. Le système interroge les contenus à l'aide de mécanismes de similarité sémantique¹⁰ pour produire des réponses ancrées dans les sources. Les étudiantes et étudiants utilisent l'agent conversationnel pour obtenir des explications et des clarifications liées au cours suivi.

Un projet pilote a été mené en 2023-2024 dans des cours universitaires, notamment de formation à l'enseignement. Une expérimentation est réalisée présentement à l'Université de Strathclyde (Royaume-Uni) pour comparer les perceptions des utilisatrices et des utilisateurs à l'égard des réponses de VIP-Bot ainsi que de celles produites par les GML grand public. Le projet a fait l'objet d'une publication à l'International Conference on Intelligent Tutoring Systems (ITS) en 2024 (Springer Nature Switzerland).

Description



VIP-Bot est une application accessible par une interface Web ainsi que par une intégration à la plateforme Discord¹¹.

Cet agent conversationnel permet aux enseignantes et aux enseignants de générer des contenus avec l'IAG de manière fiable, en se basant uniquement sur leurs propres cours, afin de produire des réponses ancrées dans un corpus défini suivant quatre étapes :

1. Création d'un sujet correspondant au périmètre documentaire (cours ou programme);
2. Dépôt des contenus dans divers formats (documents texte, présentations, tableaux, images et documents numérisés, fichiers audio, adresses URL, pages Web);

3. Traitement automatisé, extraction, segmentation et vectorisation sémantique des contenus;
4. Récupération des informations pertinentes par similarité sémantique et production de réponses selon une approche de génération augmentée par récupération.

Les interactions des étudiantes et des étudiants avec l'agent se font en langage naturel et à tout moment. Si aucun contenu pertinent n'est identifié, l'agent indique explicitement qu'il ne peut pas répondre.

Pourquoi avoir retenu ce cas ?

- Redonner aux enseignantes et aux enseignants le contrôle direct sur leurs données, les sources, les contenus et les modalités d'usage de l'IAG.
- Répondre à des enjeux de protection des données personnelles et de gouvernance.
- S'adapter facilement à différents contextes de cours par une architecture modulaire.
- Disposer d'un code source libre dans cette architecture modulaire.
- Camper l'enseignante ou l'enseignant dans son rôle de conceptrice ou de concepteur responsable du contenu de ses cours.
- S'ancrer dans une véritable démarche itérative de recherche et de développement.

Pourquoi adopter cette approche ?

- Accessibilité et inclusion : plusieurs interfaces sont offertes selon les profils (application Web ou Discord).
- Confidentialité et contrôle fin des corpus : les contenus sont privés par défaut.
- Usage de l'IAG dans un cadre contrôlé et sécurisant pour les étudiantes et étudiants.
- Aucune expertise avancée en IAG nécessaire chez les enseignantes et enseignants pour la configuration.

Avantages et limites



Avantages

- Réduction des hallucinations d'IA et du désalignement pédagogique : les réponses sont ancrées dans les ressources du cours.

- + Contrôle exercé par le personnel enseignant sur les sources et les contenus mobilisés.
- + Accès en tout temps pour les étudiantes et étudiants au soutien de VIP-Bot, peu importe la disponibilité du personnel enseignant.
- + Transparence : si l'agent ne trouve aucun contenu pertinent, il indique explicitement qu'il ne peut pas répondre.
- + Compatibilité avec les exigences d'intégrité scolaire et de protection des données personnelles au Canada. Architecture modulaire à code source ouvert : reproductible et adaptable à différents contextes institutionnels.
- + Architecture modulaire à code source ouvert et reproductible pouvant s'adapter à différents contextes institutionnels.
- + Double interface : application Web (navigateur) et intégration à la plateforme Discord, ce qui permet l'adaptation aux profils et aux préférences des utilisatrices et des utilisateurs.
- + Retombées scientifiques consignées (ITS, 2024).

Défis et limites

- Performance technique : le temps de réponse et la gestion des dépassements de délai¹² (time-out) nécessitent des optimisations continues.
- Qualité et pertinence des réponses fortement dépendantes de la qualité des contenus fournis, ce qui suppose un investissement en amont dans la structuration des ressources pédagogiques.
- Accompagnement nécessaire lors de l'appropriation par les enseignantes et enseignants.
- Déploiement institutionnel plus long et complexe que le prototypage.
- Nécessité d'encadrer les attentes des utilisatrices et des utilisateurs quant aux capacités du système.

Fonctionnement



Le fonctionnement de VIP-Bot repose sur cinq étapes structurées allant de la configuration par l'enseignante ou l'enseignant à l'interaction des étudiantes et des étudiants avec l'agent conversationnel.

1. **Création d'un sujet** : L'enseignante ou l'enseignant crée, dans le système, un sujet correspondant, par exemple, à un cours universitaire ou à un programme de formation.

Ce sujet constitue le périmètre documentaire à partir duquel l'agent conversationnel pourra répondre.

2. **Dépôt des contenus** : L'enseignante ou l'enseignant dépose les contenus associés au sujet au moyen de l'interface de gestion. Les formats acceptés sont les suivants :

- Documents texte (fichiers .pdf, .doc, .docx ou .txt);
- Présentations (fichiers .ppt ou .pptx);
- Tableurs (fichiers .xls ou .xlsx);
- Images et documents numérisés (traités par reconnaissance optique de caractères¹³);
- Fichiers audio (MP3, transcrits automatiquement);
- Adresses URL et pages Web.

Les contenus sont privés par défaut, avec la possibilité d'une ouverture ou d'un partage contrôlé par la personne qui les a déposés.

3. **Traitement, indexation et récupération de l'information** : Une fois les contenus déposés, le système déclenche un traitement automatisé comprenant la conversion de tous les fichiers vers un format textuel, l'extraction et le nettoyage du texte (suppression des caractères non pertinents), la détection et l'exploration contrôlée des liens Web intégrés, la segmentation du corpus de même que la vectorisation sémantique. Lorsqu'une requête est formulée, le système mobilise des mécanismes de similarité sémantique et la GAR pour repérer les éléments pertinents dans le corpus. Si aucun contenu pertinent n'est identifié, l'agent indique explicitement qu'il ne peut pas répondre.
4. **Génération de la réponse** : Les passages pertinents sont sélectionnés et une réponse est assemblée sur une base extractive. Un GML généraliste (GPT-3.5 ou Gemini) peut être utilisé uniquement pour améliorer la lisibilité et la fluidité de la réponse, sans en modifier le sens ni ajouter d'information externe. Cette étape de reformulation est paramétrable et peut être désactivée par l'enseignante ou l'enseignant.
5. **Interaction** : Les étudiantes et étudiants interagissent avec l'agent conversationnel à toute heure, par le biais de l'interface Web ou sur Discord, en posant des questions en langage naturel. Ils reçoivent des réponses basées sur les ressources déposées par l'enseignante ou l'enseignant.



Démarche de développement

Le projet consiste en une initiative de recherche appliquée portée par le professeur Maiga Chang et une petite équipe incluant des étudiantes et étudiants. La démarche distingue la plateforme générique (ChatbotLLM) des agents déployés (comme VIP-Bot) pour offrir un cadre reproductible et transférable tout en donnant aux enseignantes et aux enseignants un rôle actif dans la conception pédagogique.

Recherche scientifique

La recherche a fait l'objet d'une publication scientifique à la conférence ITS en 2024 (Springer Nature Switzerland). Elle a permis d'étudier la perception des étudiantes et des étudiants en formation à l'enseignement quant à la facilité d'utilisation et à l'utilité de l'outil.

Expérimentation comparative en cours

Des tests et une évaluation sont menés présentement à l'Université de Strathclyde (Royaume-Uni), dans le cadre d'un cours, afin de comparer les perceptions des utilisatrices et des utilisateurs à l'égard des réponses produites par VIP-Bot ainsi que de celles générées par les GML grand public.

Canaux de déploiement

- + Accessible par un navigateur Web.
- + Intégré à Discord, un canal privilégié par la communauté étudiante.

Infrastructure technique hybride

- **Hébergement** : Application Web hébergée en mode infonuagique¹⁴, accessible par un navigateur ou Discord, et serveurs locaux utilisés pour le traitement et le stockage des données pédagogiques ainsi que la gestion du système.
- **Technologies d'interface** : Interface développée à l'aide de technologies Web standards (HTML, CSS, JavaScript).

- **Traitement des données** : Langage PHP, langage Python et outils spécialisés pour la transformation, l'indexation de même que la vectorisation des contenus pédagogiques (documents texte, documents PDF, fichiers audio).
- **Architecture** modulaire à code source ouvert et reproductible (accessible sur demande).
- **GML** pour la reformulation (optionnel) : GPT-3.5 (OpenAI) ou Gemini (Google), utilisés uniquement pour améliorer la lisibilité et la fluidité de la réponse sans modification du sens.

Formats de contenus acceptés

- Documents texte (fichiers .pdf, .doc, .docx ou .txt).
- Présentations (fichiers .ppt ou .pptx).
- Tableurs (fichiers .xls ou .xlsx).
- Images et documents numérisés (traités par la reconnaissance optique de caractères).
- Fichiers audio (MP3, transcrits automatiquement).
- Adresses URL et pages Web.

Impacts



Impacts pédagogiques

- + Les enseignantes et enseignants considèrent VIP-Bot comme un outil de médiation pédagogique permettant de prolonger l'accompagnement des étudiantes et des étudiants sans se substituer à l'évaluation enseignante ni au jugement professionnel.
- + Ils estiment que l'outil renforce leur contrôle sur les contenus et les modalités d'usage de l'IA dans leurs cours.
- + Les étudiantes et étudiants jugent l'outil motivant, facile à utiliser et utile, notamment pour générer des idées, clarifier des concepts et amorcer l'exploration de la littérature.

Impacts institutionnels

- + Approche structurée et gouvernance compatible avec les exigences d'intégrité scolaire, de cohérence pédagogique et de protection des données personnelles.
- + Solution de rechange crédible aux outils d'IA généralistes.

Impacts scientifiques

- + Retombées scientifiques par l'étude d'un modèle hybride d'agent conversationnel éducatif offrant un terrain d'analyse des usages et des perceptions.
- + Publication validée par des pairs : Nikou et autres, 2024.

Déploiement et transférabilité



Déploiement

- 2023-2024 : projet pilote dans des cours, notamment de formation à l'enseignement, à l'Université d'Athabasca.
- Expérimentation comparative à l'Université de Strathclyde (Royaume-Uni).
- Déploiement par une application Web (navigateur) et Discord.

Transférabilité

- + Architecture modulaire à code source ouvert (accessible sur demande), adaptable à différents contextes institutionnels et disciplinaires, en formation à distance comme en présentiel.
- + Architecture permettant une exploitation dans différents cours.
- + Approche applicable à d'autres usages organisationnels allant au-delà de l'enseignement.

Conditions de transférabilité

- Gouvernance institutionnelle claire : politiques de données, responsabilités définies, supervision des contenus.
- Qualité et structuration des ressources fournies par le personnel enseignant.
- Accompagnement adéquat du personnel enseignant.
- Infrastructure de serveurs ou accès infonuagique disponibles.



- **Ancrer l'agent conversationnel dans des contenus de cours validés et à jour.** La qualité des réponses dépend fortement de la qualité des contenus fournis par l'enseignante ou l'enseignant.
- **Rendre explicites les limites, les sources et les règles de fonctionnement de l'agent conversationnel** dès le déploiement. Les utilisatrices et utilisateurs doivent comprendre ce que l'agent peut ou ne peut pas faire.
- **Maintenir une gouvernance humaine claire** incluant la sélection, la mise à jour et la supervision des contenus par l'enseignante ou l'enseignant.
- **Anticiper les processus de gouvernance** et ceux d'adaptation aux politiques institutionnelles.
- **Déployer l'outil à partir de cas d'usage pédagogiques précis.**
- **Accompagner les utilisatrices et utilisateurs** à l'aide de formations, de consignes, d'exemples de requêtes et de balises d'usage.



Conseils pour les établissements souhaitant adopter une approche similaire :

- Prioriser la qualité et la structuration des contenus déposés plutôt que la sophistication du GML : il s'agit du principal déterminant de la qualité des réponses.
- Encadrer les utilisatrices et utilisateurs en leur expliquant clairement ce que l'agent conversationnel peut ou ne peut pas faire.
- Former les enseignantes et enseignants au regard de la formulation de requêtes pertinentes et de la structuration de contenus de qualité.
- Compléter les données quantitatives d'usage par des retours qualitatifs (sondages, entretiens) pour évaluer en profondeur l'utilité pédagogique de l'outil.
- Prévoir un déploiement institutionnel plus long et complexe que le prototype.
- Adapter le dispositif à son propre contexte pédagogique et institutionnel.



Poste	Type	Détail
Développement et validation (ressources humaines)	Unique (principal poste) Important	Temps cumulé de l'équipe de recherche : programmation, conception du système, tests avec les utilisatrices et utilisateurs, validation scientifique (dont la préparation de la publication pour la conférence ITS en 2024). Principal poste de coûts lié au projet.
Infrastructure de serveurs (hébergement infonuagique ou serveurs)	Récurrent Modéré	Hébergement de l'application Web et du serveur de traitement des contenus. Frais récurrents liés au stockage des corpus et à la disponibilité du service.
Inférence des GML de reformulation (optionnel)	Récurrent Variable, mais modéré	Frais d'appel aux GML généralistes (GPT-3.5 ou Gemini) pour la reformulation des réponses, facturés à l'usage. Étape paramétrable et coûts maîtrisables selon le degré d'activation.
Maintenance et fonctionnement courants	Récurrent Moyennement élevé	Maintenance de la plateforme (coûts en ressources humaines).



Publication scientifique

- Nikou, S. A., Guliya, A., Van Verma, S. et Chang, M. (2024). A generative artificial intelligence empowered chatbot: System usability and student teachers' experience. International Conference on Intelligent Tutoring Systems (ITS) (p. 330-340). Cham : Springer Nature Switzerland. [🔗](#)

Plateforme du VIP Research group

- Lien 

Répertoire GitHub (code source)

[sur demande]

Contact

Maiga Chang, professeur, Faculté des sciences et des technologies, Université d'Athabasca

Direction

Maiga Chang (professeur, Faculté des sciences et technologies, Université d'Athabasca ; directeur du VIP Research Group).