

Vers des Parcours d'Apprentissage Adaptatifs et Personnalisés en Entreprise : Évaluation et Recommandation par Agents IA

Houssam Kanso¹ Hugues Ali Mehenni¹ Mohamad Al Assaad¹

(1) Talan Research and Innovation Center, Paris, France

houssam.kanso@talan.com, hugues.ali-mehenni@talan.com,

mohamad.al-assaad@talan.com

RÉSUMÉ

L'apprentissage personnalisé est devenu un enjeu majeur de l'éducation moderne, notamment avec les avancements de l'intelligence artificielle dans l'enseignement et l'évaluation. Cet article présente une architecture multi-agents qui combine des modèles de langage, la génération automatisée d'évaluations, l'évaluation des réponses des apprenants et la recommandation de formations à suivre afin de favoriser des parcours d'apprentissage adaptatifs dans le cadre de l'apprentissage et la montée en compétences continue des salariés dans les entreprises. Le système s'appuie sur des évaluations conçues par des experts et des données de formations enrichies pour identifier les lacunes des apprenants et recommander des ressources pédagogiques pertinentes. En associant la recherche sémantique et des agents d'IA modulaires, l'approche proposée vise à fournir un soutien à l'apprentissage personnalisé et évolutif pour les environnements éducatifs.

ABSTRACT

Toward Adaptive and Personalized Learning Paths in the Workplace : Evaluation and Recommendation via AI Agents

Personalized learning has become a major challenge in modern education, particularly with the advances of artificial intelligence in teaching and assessment. This article presents a multi-agent architecture that combines language models, automated assessment generation, learner response evaluation, and training recommendation in order to promote adaptive learning pathways within the context of learning and continuous employee upskilling in companies. The system relies on assessments designed by experts and enriched training data to identify learners' gaps and recommend relevant educational resources. By combining semantic search and modular AI agents, the proposed approach aims to provide scalable and personalized learning support for educational environments.

MOTS-CLÉS : Intelligence Artificielle, Apprentissage Personnalisé, Grands Modèles De Langage, Recommandation De Cours, Évaluation Automatisée.

KEYWORDS: Artificial Intelligence, Personalized Learning, Large Language Models, Course Recommendation, Automated Assessment.

1 Introduction

L'évolution des modèles éducatifs, témoigne d'une transition historique majeure allant de l'instruction sur mesure à la scolarisation de masse, pour aboutir aux solutions technologiques contemporaines. Avant le XVIII^e siècle, le système privilégiait un apprentissage personnalisé reposant sur le tutorat

individuel et l'enseignement en petits groupes. Au cours des XVIII^e et XIX^e siècles, cette approche a évolué avec l'introduction de programmes d'études standardisés afin d'encadrer des groupes d'élèves de plus en plus larges. Cette tendance s'est accentuée au XX^e siècle, période marquée par une augmentation spectaculaire du nombre d'inscriptions, ce qui a eu pour conséquence directe une réduction de l'attention individuelle accordée à chaque étudiant (Dockterman, 2018). Enfin, l'ère numérique actuelle se caractérise par l'émergence de l'apprentissage en ligne et de l'Intelligence Artificielle (IA), marquant de nouvelles tentatives de personnalisation pédagogique, bien que des défis subsistent encore pour transformer durablement les méthodes d'enseignement.

Ces dernières années, l'IA a eu un impact considérable dans le domaine de l'éducation, en particulier avec l'utilisation des grands modèles de langage (LLMs) et de l'IA générative (GenAI), conduisant à une nouvelle ère de l'éducation. Des systèmes de tutorat intelligents à la génération automatisée de contenu, ces technologies sont intégrées dans les pratiques d'enseignement et d'apprentissage afin de fournir des expériences évolutives, adaptatives et personnalisées (Mousaviniasab *et al.*, 2018). Cependant, l'IA présente des limites (Qiu & Liu, 2025), notamment lorsqu'il s'agit de la génération d'examens pertinents et de l'évaluation de ces examens. Dans ce papier, nous traitons ce problème et explorons l'utilisation de l'IA en proposant une solution multi-agents capable d'évaluer les apprenants et de leur proposer les meilleures formations à suivre en se basant sur leur profil et leur évaluation.

Notre article est organisé de la façon suivante. La section 2 présente quelques travaux connexes. Dans la section 3, nous présentons notre architecture multi-agent d'apprentissage personnalisé avec évaluation et recommandation basées sur les LLMs. Dans la section 4, nous présentons l'implémentation de preuve de concept de notre architecture ainsi que nos premiers résultats. Enfin, nous concluons et présentons les perspectives futures dans la section 5.

2 État de l'art

L'apprentissage personnalisé assisté par l'intelligence artificielle a connu une évolution remarquable ces dernières années, transformant les approches traditionnelles de l'éducation. L'émergence des LLMs a ouvert de nouvelles perspectives dans trois domaines clés. Premièrement, la génération automatisée d'évaluations : (Maity & Deroy, 2024) démontrent que les LLMs peuvent générer automatiquement des contenus pédagogiques et des questions d'examen, bien que la validation de la qualité pédagogique des évaluations générées demeure un défi critique (Yan *et al.*, 2024). Deuxièmement, l'évaluation des réponses : (Seneviratne & Manathunga, 2025) montrent que les LLMs peuvent évaluer efficacement les réponses d'apprenants avec une bonne fiabilité. Troisièmement, la recommandation personnalisée : (Li *et al.*, 2023) démontrent l'efficacité des approches basées sur les embeddings sémantiques pour recommander des ressources pédagogiques alignées avec le profil et les lacunes des apprenants. L'architecture multi-agents constitue une approche prometteuse pour combiner ces composantes : (Zhang *et al.*, 2025) montrent comment décomposer les tâches éducatives en agents spécialisés améliore la modularité et la scalabilité. Par ailleurs, l'IA explicable devient de plus en plus importante dans les systèmes d'apprentissage. Elle permet de mieux comprendre les décisions des modèles, d'améliorer la confiance des enseignants et de rendre les recommandations plus transparentes. Par exemple, certaines méthodes peuvent expliquer les résultats de manière individuelle pour chaque apprenant (Alibayeva & Allayarova, 2025). Dans le même contexte, les systèmes d'apprentissage adaptatifs en formation continue évoluent vers des approches centrées sur l'utilisateur. Ils prennent en compte des informations comme le niveau de l'apprenant, le temps disponible ou

ses contraintes professionnelles afin d'adapter les contenus. Cela permet d'ajuster la difficulté, de recommander des ressources, et de proposer des outils de suivi comme des tableaux de bord ou des rappels (Fromm & Ifenthaler, 2024). Néanmoins, plusieurs défis subsistent : (Miranda, 2025) soulèvent des préoccupations quant aux biais dans les évaluations générées et la fidélité pédagogique des contenus automatisés, tandis que la validation empirique des systèmes personnalisés complets dans des contextes professionnels reste limitée. Notre travail s'inscrit dans ce contexte en proposant une architecture intégrant la génération adaptative d'évaluations, la données venant d'experts et la recommandation sémantique de formations dans le cadre de l'apprentissage professionnel en entreprise.

3 Architecture

L'architecture proposée est composée de plusieurs pipelines coordonnés qui forment ensemble un système intelligent de recommandation d'apprentissage basé sur des agents. L'objectif principal de cette architecture est d'extraire les connaissances à partir des évaluations d'experts et des contenus pédagogiques disponibles, d'évaluer le niveau de compétence d'un utilisateur à travers une évaluation adaptative, puis de recommander des parcours d'apprentissage personnalisés en conséquence.

3.1 Pipeline d'ingestion des évaluations d'experts

Le pipeline présenté dans la Figure 1 est responsable de la transformation des données brutes issues des évaluations qui sont écrites par des experts en représentations vectorielles structurées pour quelles puissent être récupérées et utilisées comme exemples par les agents pour une meilleure performance. Ce pipeline d'ingestion garantit que les connaissances issues des experts sont efficacement indexées et rendues accessibles pour la génération et l'évaluation adaptatives des évaluations.

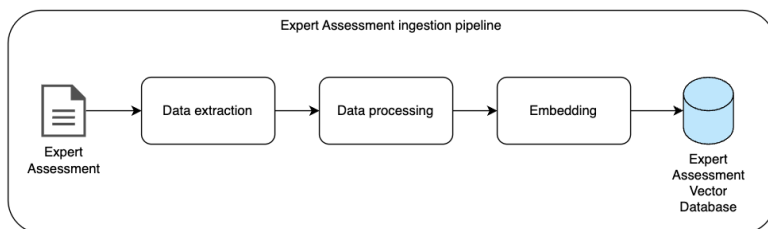


FIGURE 1 – Pipeline d'ingestion des évaluations d'experts

Extraction des données : Les documents bruts d'évaluation des experts (contenant les questions, les réponses attendues et les grilles d'évaluation) sont analysés et transformés en un format exploitable par la machine.

Traitement des données : Les données extraites sont nettoyées et standardisées, garantissant la cohérence du format des questions, du balisage des métadonnées et de la représentation des réponses. Chaque évaluation contient un ensemble de questions à choix multiples (à réponses simples ou multiples), associées à des niveaux de difficulté, des compétences ciblées, ainsi qu'un barème de

points. Cette structure permet de représenter de manière normalisée des questionnaires pédagogiques variés, indépendamment du domaine abordé, afin de faciliter leur exploitation dans des contextes d'évaluation automatique et d'analyse des compétences.

Embedding : Les évaluations traitées sont converties en représentations vectorielles à l'aide d'un modèle sémantique. Ces embeddings sont stockés dans la *base de données vectorielle des évaluations d'experts*, permettant une récupération basée sur la similarité pour l'apprentissage few-shot et le prompting contextuel dans les étapes suivantes.

3.2 Pipeline d'ingestion des formations

Le pipeline d'ingestion des formations est illustré dans la Figure 2. Il est conçu pour pré-traiter et enrichir les métadonnées des formations avant leur intégration dans le système de recommandation. Il transforme des données de formations brutes (généralement sous forme tabulaire) en une base de données enrichie sémantiquement et interrogeable. Ce pipeline permet une ingestion automatisée et scalable de diverses sources de formations, tout en garantissant la qualité des métadonnées et leur richesse sémantique.

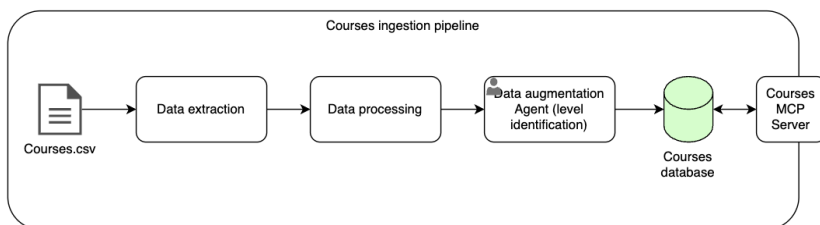


FIGURE 2 – Pipeline d'ingestion des formations

Extraction des données : Le système lit et extrait les attributs des formations tels que le titre, la description, les thématiques, les prérequis et la durée. Le corpus de formations utilisé dans notre cas est une liste de plus de 24 000 formations couvrant de nombreux domaines, tels que le développement informatique, la science des données, l'intelligence artificielle, la cybersécurité, le cloud computing, le marketing digital, la gestion de projet, la bureautique, le design graphique, la photographie, la musique, le développement personnel, la finance, l'entrepreneuriat, ainsi que les langues et la préparation aux certifications professionnelles.

Traitement des données : Les données extraites sont ensuite nettoyées, normalisées et structurées afin d'assurer une représentation uniforme des caractéristiques des formations.

Agent d'augmentation des données : Un agent identifie et attribue automatiquement à chaque formation un niveau de compétence estimé (par exemple débutant, intermédiaire, avancé) en analysant les descriptions des cours et les objectifs d'apprentissage. Ce niveau est ensuite validé par un expert humain.

Base de données des formation et serveur MCP : Les formations traités et enrichis sont stockés dans la base de données, qui est connectée au serveur MCP. Cela permet l'interrogation en temps réel

et la récupération des ressources d'apprentissage les plus pertinentes pour chaque utilisateur.

3.3 Pipeline principal de recommandation et d'évaluation

Le pipeline principal de recommandation et d'évaluation est représenté dans la Figure 3. Il intègre l'interaction utilisateur, l'évaluation personnalisée, l'évaluation des réponses et la recommandation de cours. Il exploite les connaissances stockées dans la base de données vectorielle des évaluations d'experts ainsi que le serveur MCP des cours afin de produire des recommandations d'apprentissage adaptées. Le système repose sur 4 agents, où chaque agent est spécialisé dans une tâche précise et contribue à une étape bien définie du processus. Le flux de travail s'articule autour des agents et processus suivants :

Agent d'extraction du profil : Analyse les CV des utilisateurs et leurs réponses afin de construire un profil détaillé capturant leur parcours, leurs centres d'intérêt et leurs objectifs d'apprentissage.

Agent générateur d'évaluations : Utilise les exemples récupérés depuis la base vectorielle des évaluations d'experts pour générer des questions d'évaluation personnalisées, alignées avec le domaine cible et le niveau de compétence de l'utilisateur.

Agent évaluateur : Compare les réponses de l'utilisateur aux réponses de référence des experts afin de produire un rapport d'évaluation mettant en évidence les points forts, les points faibles et les axes d'amélioration.

Agent de recommandation de formations : Combine le rapport d'évaluation et le profil utilisateur pour interroger le serveur MCP des formations et sélectionner les formations les plus pertinentes et adaptées au niveau de l'apprenant. Il fournit un ensemble personnalisé de recommandations de formations alignées avec le niveau actuel de l'utilisateur et sa trajectoire d'apprentissage.

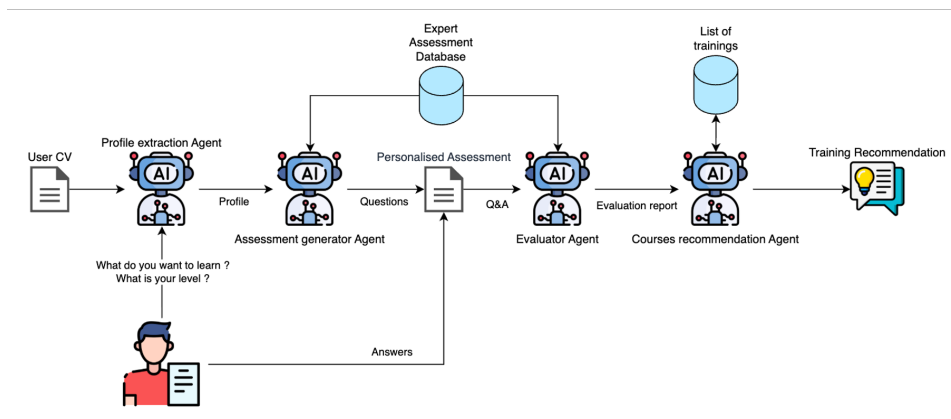


FIGURE 3 – Pipeline principal de recommandation et d'évaluation

L'architecture présente plusieurs avantages, notamment la combinaison entre l'évaluation basée sur l'expertise et la recommandation de parcours fondée sur les données. Elle est également modulaire, ce qui permet de faire évoluer les composants d'ingestion et d'évaluation. De plus, elle offre une bonne scalabilité pour traiter de grands volumes de données liées aux évaluations et aux formations.

4 Implémentation

Afin de valider et de mettre en avant le fonctionnement de l'architecture proposée, une implémentation complète de celle-ci a été réalisée.

```
== Démarrage de l'évaluation adaptative ==
-- Vague 1 / 3 (difficulté visée: easy) --

Question 1 / 2 | ID=pyb-4
Sujet : strings | Difficulté : easy
Quelle expression renvoie la longueur de la chaîne 'python' ?
A. length('python')
B. len('python')
C. size('python')
D. count('python')
Votre réponse (texte libre ou lettre du QCM) : A

Question 2 / 2 | ID=pyb-1
Sujet : variables | Difficulté : easy
Quelle est la sortie du code suivant : x = 3; y = 2; print(x + y) ?
A. 2
B. 3
C. 5
D. 6
Votre réponse (texte libre ou lettre du QCM) : B
[adapt] Distraction appliquée pour la question pyb-2
```

FIGURE 4 – Exemple de la phase d'évaluation

Dans la phase d'évaluation, l'utilisateur répond à des questions qui sont proposées par l'agent générateur de questions. La figure 4 représente un exemple dans lequel l'utilisateur répond à deux questions de niveau facile. Une fois que l'utilisateur répond au premier niveau de questions, il reçoit une autre vague de questions d'un niveau différent. Ce processus permet d'évaluer le niveau de l'utilisateur en commençant avec des questions faciles et en s'adaptant en fonction des réponses.

Dans la phase d'analyse, l'agent responsable de l'évaluation des réponses de l'utilisateur évalue chacune des réponses de l'utilisateur et identifie le niveau de connaissance de l'utilisateur dans un domaine ou une compétence bien définie. La figure 5 montre un exemple de l'analyse des réponses de l'utilisateur, où l'utilisateur a répondu à toutes les questions de manière incorrecte et où le niveau de l'utilisateur a été défini comme débutant.

```
Profile:
UserProfile(name=None, years_experience=0.5, current_level='beginner', known_tools=['python', 'sql'], interests=['data', 'ml', 'backend'], goal='I want to learn python')
Mastery:
MasteryState(subject='python', level='beginner', score=0.0)

Score global: 0.0%

Détail des questions:
- pyb-4: 0% x - Mauvaise réponse. Réponse attendue : len('python').
- pyb-1: 0% x - Mauvaise réponse. Réponse attendue : 5.
- pyb-2: 0% x - Mauvaise réponse. Réponse attendue : for i in range(3): print(i).
- pyb-5: 0% x - Mauvaise réponse. Réponse attendue : nums.append(4).
- pye-1: 0% x - Mauvaise réponse. Réponse attendue : Contrôler la création des classes elles-mêmes en personnalisant le type de la classe.
- pye-5: 0% x - Mauvaise réponse. Réponse attendue : Utiliser un profileur (comme cProfile) pour identifier les points chauds avant d'appliquer des optimisations ciblées..
```

FIGURE 5 – Exemple de la phase d'analyse des réponses

Dans la phase de recommandation, l'agent responsable de proposer des formations est invoqué. Il se base sur les réponses de l'utilisateur, le profil qui a été extrait ainsi que le niveau de connaissance qui a été identifié pour proposer une formation à ce dernier. Dans la figure 6, deux formations ont été proposées à l'utilisateur. La première s'appelle "Python pour débutants" et la deuxième, "Python 101". Une explication est également fournie à l'utilisateur pour lui expliquer le parcours à suivre.

Malgré l'ensemble des mécanismes proposés dans cette architecture, plusieurs limites potentielles doivent être considérées. En particulier, les systèmes reposant sur des agents génératifs peuvent

```
Recommendations:
- Python pour débutants: This course covers the fundamental aspects of Python programming, including basic syntax, variables, loops, and functions. It's ideal for beginners looking to establish a solid foundation in Python.
- Python 101: First Steps: This video series introduces Python's basic types, control flow, and data structures. Perfect for beginners who want to start coding in Python and understand its core concepts.

Adaptive Content:
To begin your journey in Python programming, start with 'Python pour débutants' offered by OpenClassrooms. This MOOC is designed to help you grasp the fundamental syntax and concepts such as variables, loops, and functions. Spend time understanding each topic thoroughly, as they form the backbone of Python programming. Once you complete this course, move on to 'Python 101: First Steps' from Coursera. This video series will expand your knowledge by introducing types, control flow, and basic data structures. These are crucial for writing efficient and effective Python code. As you progress, practice coding regularly to reinforce your learning. Use online coding platforms to solve beginner-level Python challenges, which will help solidify your understanding. Additionally, explore Python documentation and community forums to gain insights and tips from experienced programmers. By following these formations sequentially, you'll build a strong foundation in Python, setting the stage for more advanced topics in the future.
```

FIGURE 6 – Exemple de la phase de recommandation

introduire des biais liés aux données d'entraînement ou aux exemples utilisés pour la génération des questions et des recommandations. De plus, le risque d'hallucinations dans la génération automatique de contenus d'évaluation ou de suggestions de formations peut affecter la fiabilité du système. Dans ce contexte, il apparaît nécessaire de compléter ces travaux par des études qualitatives approfondies afin d'évaluer la qualité des questions générées, la pertinence des formations recommandées et l'impact global du système sur les parcours d'apprentissage des utilisateurs.

5 Conclusions et perspectives

Dans cet article, nous avons présenté une architecture basée sur des agents pour la recommandation personnalisée d'apprentissage et la génération d'évaluations adaptatives à l'aide de LLMs et de pipelines augmentés par les données. Notre approche combine l'ingestion d'évaluations d'experts, la récupération sémantique vectorielle, l'évaluation automatisée et la recommandation de formation au sein d'un cadre capable d'adapter les parcours d'apprentissage aux profils et aux niveaux de compétence des utilisateurs. Notre travail démontre le potentiel de la combinaison de mécanismes d'évaluation basés sur les LLMs avec des données éducatives structurées pour soutenir des environnements d'apprentissage adaptatifs, tout en conservant une architecture modulaire et extensible.

Comme perspectives futures, nous prévoyons d'évaluer le système proposé auprès de véritables apprenants dans des contextes éducatifs réalistes afin d'analyser son efficacité, son utilisabilité, son impact pédagogique et le taux de satisfaction des utilisateurs. Nous envisageons également d'intégrer des objectifs d'apprentissage explicites pour chaque formation et d'établir des liens entre les questions d'évaluation et ces objectifs. Cet alignement permettrait un suivi plus précis des compétences et améliorerait la pertinence des parcours d'apprentissage recommandés. Enfin, nous souhaitons étendre le système avec des capacités de génération de contenu personnalisé, telles que la création automatique de du contenu de formations et de supports de remédiation ciblés sur les faiblesses identifiées de l'apprenant.

Références

- ALIBAYEVA R. N. & ALLAYAROVA S. U. (2025). From data to pedagogy : Leveraging explainable artificial intelligence to enhance trust, transparency, and effectiveness in intelligent learning systems. *Qubahan Techno Journal*.
- DOCKTERMAN D. (2018). Insights from 200+ years of personalized learning. *NPJ Science of Learning*, 3.

- FROMM Y. M. & IFENTHALER D. (2024). Designing adaptive learning environments for continuing education : Stakeholders' perspectives on indicators and interventions. *Computers in Human Behavior Reports*.
- LI X., HENRIKSSON A., DUNELD M., NOURI J. & WU Y. (2023). Evaluating embeddings from pre-trained language models and knowledge graphs for educational content recommendation. *Future Internet*, **16**, 12.
- MAITY S. & DERROY A. (2024). The Future of Learning in the Age of Generative AI : Automated Question Generation and Assessment with Large Language Models.
- MIRANDA S. (2025). Artificial intelligence from google environment for effective learning assessment. *Inf.*, **16**, 462.
- MOUSAVINASAB E., ZARIFSANAIEY N., KALHORI S. R. N., RAKHSHAN M., KEIKHA L. & SAEEDI M. (2018). Intelligent tutoring systems : a systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, **29**, 142 – 163.
- QIU Y. & LIU C. (2025). Capable exam-taker and question-generator : the dual role of generative ai in medical education assessment. *Global Medical Education*, **2**, 135 – 148.
- SENEVIRATNE H. & MANATHUNGA S. (2025). Artificial intelligence assisted automated short answer question scoring tool shows high correlation with human examiner markings. *BMC Medical Education*, **25**.
- YAN L., SHA L., ZHAO L., LI Y., MARTINEZ-MALDONADO R., CHEN G., LI X., JIN Y. & GAŠEVIĆ D. (2024). Practical and ethical challenges of large language models in education : A systematic scoping review. *British Journal of Educational Technology*, **55**(1), 90–112.
- ZHANG X., ZHANG C., SUN J., XIAO J., YANG Y. & LUO Y. (2025). Eduplanner : Llm-based multiagent systems for customized and intelligent instructional design. *IEEE Transactions on Learning Technologies*, **18**, 416–427.