

Évaluer avec les IAG

1.	Introduction	1
2.	Utiliser les IAG pour aider à la conception des évaluations	2
2.1.	Créer des quiz	2
2.1.1.	Créer des questions	3
2.1.2.	Importer des questions sur le Portail des études	4
2.2.	Relecture des évaluations	5
2.3.	Situations fictives	6
3.	Évaluer les étudiants à l'ère des IAG	7
3.1.	Poser un cadre	7
3.2.	Évaluations résilientes aux IAG	8
3.2.1.	Évaluations orales	8
3.2.2.	Évaluations processuelles	9
3.3.	Adapter ses évaluations pour intégrer les IAG	10
3.3.1.	Déplacer l'évaluation du produit vers le processus	10
3.3.2.	Favoriser les évaluations interactives et en contexte	10
3.3.3.	Encadrer l'usage des IAG	10
4.	Conclusion	11

1. Introduction

C'est l'une des premières inquiétudes qui s'est manifestée lors de la mise en ligne des IAG : si les étudiants, élèves, stagiaires, etc. peuvent utiliser des générateurs de texte pour faire leurs devoirs à leur place, comment peut-on encore évaluer le niveau d'un apprenant ?

Cette partie va porter sur **deux aspects de l'évaluation des étudiants à l'ère des IAG** : premièrement, nous verrons comment vous pouvez **utiliser les outils d'IAG pour vous aider à évaluer les étudiants**, et ensuite nous

verrons comment **adapter ses évaluations pour intégrer l'existence des IAG** (évaluations résilientes aux IAG, modification du type ou du contenu des évaluations).

2. Utiliser les IAG pour aider à la conception des évaluations

Comme dans la préparation d'un cours et l'accompagnement des étudiants, il est possible de se servir des outils d'IAG pour s'aider dans la conception d'évaluations. Il n'est évidemment **pas question ici de déléguer l'évaluation des étudiants** à un outil d'intelligence artificielle, mais de vous donner des pistes sur les différentes façons de se servir des IAG pour **gagner du temps ou proposer différents types d'évaluations** en cours d'année pour aider vos étudiants dans leurs apprentissages.

Nous allons voir ici comment utiliser les IAG pour s'aider dans la conception et l'intégration de quizz en ligne, mais aussi en relecture et en création de situations fictives cohérentes et adaptées au contexte.

2.1. Créer des quiz

Il existe de nombreuses études (dont [la plus exhaustive](#) publiée par Black et Williams fait toujours référence sur le sujet bien qu'elle soit datée de 1998) qui démontrent que les évaluations formatives sont des **dispositifs d'apprentissage efficaces**, peu importe la discipline.

Les évaluations formatives sont des évaluations sans « sanction », c'est-à-dire qu'elles servent juste à mesurer en cours d'année le niveau atteint par les étudiants en regard des objectifs du cours. Associées à un feedback pertinent, elles permettent aux étudiants de se situer par rapport aux objectifs finaux et à identifier leurs forces et leurs faiblesses dans leur apprentissage. Pour l'enseignant, elles permettent en plus d'obtenir régulièrement une « image » du niveau de ses étudiants, qui peut amener à une modification du plan de cours (passer plus vite sur une notion déjà bien comprise, ou au contraire passer plus de temps sur un point encore mal intégré).

Ces évaluations formatives, si elles sont donc efficaces, peuvent être **chronophages à mettre en place**. En règle générale, on choisit des modes d'évaluation automatisés pour laisser à l'étudiant le choix du moment où il veut les réaliser, et qui permettent un résultat et un feedback immédiat.

Créer des quiz peut prendre du temps : choix du type de question (QCM classique, texte à trous, appariement, réponse courte, vrai/faux, etc.), nombre assez important pour balayer tout le programme, discriminants plausibles : il y a tout une série de paramètres à prendre en compte, sans oublier que chaque question doit **comporter un feedback précis** en fonction des réponses choisies. De plus, saisir ensuite toutes vos questions sur le Portail des études peut être long également, puisqu'il faut créer chaque question une par une.

Vous pouvez vous aider des IAG pour :

- Créer des questions pour vos évaluations
- Générer des fichiers d'importation de questions pour vos espaces de cours ce qui crée automatiquement les questions en une seule fois.

2.1.1. Créer des questions

En utilisant [le cours créé précédemment sur la Révolution industrielle](#), et toujours en se basant sur les contenus et le plan généré par les IAG, nous allons maintenant demander à l'outil de nous générer des QCM :

Prompt : Tu es un enseignant en Histoire et expert en pédagogie universitaire. Pour des étudiants de troisième année de licence 3, tu vas faire un cours basé sur le plan que nous venons de créer ensemble. Génère 5 questions de quiz basées uniquement sur le contenu de la première séance de ce cours, dont les objectifs d'apprentissage sont :

Définir et dater le phénomène de la Révolution industrielle (RI) selon différentes écoles de pensée,

Distinguer les concepts clés : invention, innovation majeure et incrémentale

Commencer à construire un argument historiographique sur la pertinence du terme « révolution »

Le quiz doit pouvoir être facilement intégré sur une plateforme Moodle, avec l'activité Test. Je veux 5 questions de type QCM, 5 items de réponse à chaque fois, donnez la ou les bonnes réponses, et rédigez des feedbacks à donner aux étudiants en fonction de leur réussite ou de leur échec à chaque item de réponse. Les distracteurs doivent être plausibles mais incorrects, il ne faut pas de formulations ambiguës, les questions doivent couvrir l'ensemble des objectifs.

- **Rôle et contexte**

Comme à chaque fois, il faut donner à l'IAG un rôle en relation avec la demande, et un contexte d'application

« Tu es un enseignant en Histoire et expert en pédagogie universitaire. Pour des étudiants de troisième année de licence 3 »

- **Demande**

Il faut être précis dans la demande que vous formulez afin d'obtenir rapidement un résultat conforme à vos attentes : nombre de questions, type de question voulu (QCM, Texte à trous, etc.), ainsi que le champ d'application des questions (ici le premier cours du programme). Précisez bien que, pour chaque item de réponse, il faut que l'IAG génère également un feedback afin que l'étudiant sache pourquoi la réponse était bonne ou fausse, ce qui permet de renforcer l'apprentissage : *« Je veux 5 questions de type QCM, 5 items de*

réponse à chaque fois, donnez la ou les bonnes réponses, et rédigez des feedbacks à donner aux étudiants en fonction de leur réussite ou de leur échec à chaque item de réponse »

- **Précisions**

Ici dans l'exemple, les précisions données portent sur la plateforme d'accueil des questions (Moodle), ainsi que sur la portée des questions (couvrir l'ensemble des objectifs). Vous pouvez apporter d'autres précisions (nombre de bonnes réponses maximum, longueur des questions et des réponses, etc.).

2.1.2. Importer des questions sur le Portail des études

Le Portail des études de l'ENS de Lyon est une plateforme Moodle, et il est possible d'importer des quizz en une seule fois dans son espace de cours. Pour cela, il faut générer un fichier .xml. Vous pouvez demander à une IAG de le faire pour vous en utilisant soit les questions de quizz générées par une IAG, soit vos propres questions, soit un mélange des deux.

Note : en testant sur plusieurs plateformes, il apparaît que NotebookLM génère un code xml qui n'est pas totalement utilisable en l'état. ChatGPT par contre y arrive « du premier coup », notamment pour des quizz simples comme des QCM. Plus le type de question est « compliqué » (particulièrement pour les questions « Cloze »), plus il existe des risques d'erreurs dans le code généré, qui empêche l'import. Néanmoins, en donnant à l'IAG le message d'erreur au moment de l'import, elle peut générer à nouveau un code xml corrigé.

Prompt : Je veux intégrer le quizz en pièce jointe sur un cours dans une plateforme Moodle. Génère-moi le code XML correspondant mais pour chaque proposition de réponse à chaque fois, supprime la lettre mise en premier caractère (A, B, C, D, E,).

Ensuite, explique moi comment intégrer ce code dans mon cours Moodle pour importer les questions.

[Ici vous pouvez visualiser le test généré](#) après l'import du code XML correspondant (le cours vous demandera de vous inscrire pour pouvoir le visualiser). La partie du prompt « explique-moi comment importer le quizz » vous permet d'avoir un tutoriel complet pour effectuer l'opération sur votre cours.

Pour obtenir ce résultat, le prompt est plus simple : il suffit de transposer un texte ou un document en langage XML, il n'est donc pas besoin de donner un rôle ou un contexte. Il est juste précisé de supprimer les lettres "A, B, C, ..." présentes dans le fichier source pour ne pas qu'elles apparaissent ensuite dans le test Moodle (double usage avec les numérotations de questions déjà présentes sur Moodle).

2.2. Relecture des évaluations

Que ce soit pour des évaluations formatives ou sommatives (évaluations « notées » qui vont permettre à l'étudiant de valider ou non votre cours), vous pouvez utiliser les IAG comme **un relecteur** afin de détecter **des ambiguïtés, des erreurs ou des différences entre votre évaluation et les objectifs du cours**.

En effet, comme dans la préparation du cours, il arrive que l'on soit victime du « syndrome de l'expert » dans les évaluations, en prenant pour acquises des notions que les étudiants ne maîtrisent pas forcément, en fonction de leur niveau. Il est aussi possible que les consignes soient plus ou moins ambiguës, une IAG pourra pointer des incohérences ou les différentes façons d'interpréter certains énoncés. Par ailleurs, les IAG pourront vérifier que l'évaluation correspond bien aux objectifs pédagogiques du cours ([alignement pédagogique](#)) ainsi qu'au niveau des étudiants.

Évidemment, les résultats générés par l'IAG devront toujours être vérifiés et contrôlés en deuxième lecture par l'enseignant, mais cela peut vous donner des indications et des pistes d'amélioration très rapidement.

Prompt : Je souhaite que tu relises l'évaluation en pièce jointe, destinée à des étudiants de niveau L3 en Licence d'Histoire. Ton rôle est de m'aider à améliorer la clarté et la cohérence des consignes, ainsi que de vérifier que l'évaluation est en adéquation avec les objectifs pédagogiques du cours qui sont : (Rappeler ici les objectifs)

Ton objectif est d'identifier les ambiguïtés, incohérences ou formulations trop longues, de vérifier l'adéquation entre le niveau des questions et le niveau des étudiants et de proposer des reformulations plus claires, sans modifier le fond.

Ne change pas le contenu disciplinaire, ne change pas la structure générale de l'évaluation sauf suggestion justifiée, et je ne veux pas de simplification excessive, il faut maintenir le niveau prévu.

Liste les points à améliorer (précisément), et proposes des reformulations si besoin.

- **Rôle et contexte**

Ici, pas besoin de rôle précis, il suffit d'indiquer à l'IAG que son rôle est de relire l'évaluation et de repréciser le niveau des étudiants évalués et les objectifs pédagogiques visés par l'évaluation.

- **Demande**

Dans la demande, il faut lister ce que l'on veut que l'IAG pointe en particulier dans l'évaluation choisie. Ici, on demande notamment de trouver les ambiguïtés possibles, les incohérences ou la syntaxe générale. On veut aussi savoir si l'évaluation couvre bien les objectifs pédagogiques du cours et si elle est adaptée au niveau des étudiants.

- **Contraintes**

Pour éviter toute modification non voulue, on pose des contraintes claires à l'IAG : pas de changement dans la structure de l'évaluation, pas de simplification trop grande, pas de changement dans le contenu disciplinaire (sauf si c'est justifié).

2.3. Situations fictives

Vous avez parfois besoin, pour vos évaluations, de créer des situations fictives (par exemple créer un cas d'études, un projet, une mise en situation historique, etc.). Ces situations, qui doivent être plausibles tout en étant inventées, peuvent prendre beaucoup de temps à imaginer.

Vous pouvez vous aider des IAG pour **créer des éléments fictifs pour initier une mise en pratique des connaissances** que les étudiants ont acquises lors de votre cours, créer un cas d'études ou une situation qui regroupe les différents éléments de votre cours. L'avantage, c'est que vous pouvez demander également à une IAG de créer plusieurs situations similaires, par exemple pour un travail de groupe où chaque groupe aurait à travailler sur une situation différente.

Prompt : Tu es un enseignant-chercheur en Histoire et un expert en conception d'évaluations universitaires. Crée une situation fictive mais historiquement plausible à utiliser comme support d'évaluation pour des étudiants de Licence 3 en Histoire.

Le scénario doit porter sur la Révolution industrielle. Il doit permettre d'évaluer la capacité des étudiants à contextualiser un phénomène historique, analyser des transformations économiques, sociales ou culturelles, interpréter un document fictif mais crédible, mobiliser des connaissances historiographiques, formuler une argumentation structurée.

La situation doit être réaliste et cohérente avec l'état des connaissances historiques. Le ton doit rester universitaire, sans dramatisation romanesque. Aucune donnée ne doit être manifestement fausse historiquement.

Je veux un résultat qui comprend : une courte contextualisation du lieu et de la période, la présentation d'un ou plusieurs acteurs, un document ou un dilemme historique fictif mais vraisemblable, une tension ou un problème précis que les étudiants devront analyser.

- **Rôle et contexte**

Ici on demande à l'IAG de se comporter comme un enseignant et un expert en évaluations pour maximiser la cohérence de la réponse générée, et on rappelle le niveau des étudiants à qui va s'adresser la mise en situation.

- **Demande**

La demande va comporter le thème (ici il est resté très général mais il est possible, voire recommandé, de l'affiner en fonction de vos besoins) ainsi que les objectifs de cette mise en situation : une base pour une évaluation (comme dans l'exemple), une illustration dans un cours, un point de départ pour un travail de recherches, etc. Il faut également limiter l'IA dans le ton et les éléments qu'elle va pouvoir utiliser dans sa réponse (une demande de scénario peut vite déboucher sur une situation trop romanesque, d'autant plus si le sujet comme ici est un sujet courant de romans). Demander une tension ou un problème clairement identifié sert de base pour l'utilisation qui va être faite de la situation fictive.

- **Format de sortie**

Enfin, il faut préciser quel type de mise en situation vous souhaitez : si vous voulez des documents fictifs comme des chiffres, des cartes, des lettres, rapports, extrait d'articles de presse, ou tout autre type de documents.

3. Évaluer les étudiants à l'ère des IA

La conception même des évaluations est désormais fortement impactée par la généralisation des IA : il faut prendre en compte le fait que les étudiants ont accès à ces outils et qu'ils peuvent s'en servir.

Il est donc important de poser un cadre clair pour vos évaluations et de définir très précisément auprès de vos étudiants les utilisations que vous acceptez ou pas, et d'adapter vos évaluations en fonction de cela.

Vous pouvez par exemple consulter le travail de Jean-François Van de Poël qui a créé un modèle "[Positionnement et Réflexion sur les Activités d'Évaluation en Contexte d'Intelligence Artificielle \(PRAX-IA\)](#)". Ce modèle offre un guide destiné aux enseignants pour évaluer les risques, opportunités et implications pédagogiques de l'usage de l'IA dans les pratiques d'évaluation.

3.1. Poser un cadre

Vous pouvez vous aider de la [grille d'échelle de l'évaluation](#) proposée par Perkins, Roe & Furze qui propose cinq niveaux d'acceptation de l'utilisation des IA dans les évaluations, du plus fermé (sans IA) au plus ouvert (exploration avec l'IA).

Il appartient à chaque enseignant de choisir le niveau d'utilisation des IA qu'il accepte, et ce niveau peut être différent d'une évaluation à l'autre, d'un niveau d'étudiants à l'autre. Le principal est d'être très transparent dans :

- Ce que vous acceptez et ce que vous refusez
- Comment les étudiants doivent, le cas échéant, indiquer ce qu'ils ont fait avec l'aide des IA, voire pourquoi ils les ont utilisés, et comment ils ont intégré cela à leurs devoirs.

Il est aussi important de noter qu'une partie des étudiants **ne veut pas utiliser les IAG** : il faut donc faire attention, si vous en autorisez l'utilisation, qu'il est possible de faire l'évaluation sans IAG (par exemple, ne pas demander à un étudiant de générer un texte via ChatGPT pour ensuite le corriger).

3.2. Évaluations résilientes aux IAG

Si vous choisissez de réaliser des évaluations dans lesquelles l'emploi des IAG n'est pas (ou très peu) admis, vous devrez choisir **des modalités qui restent résilientes à leur utilisation** si vous voulez vous assurer que les étudiants respectent ce choix. En effet, bien qu'il existe à l'ENS de Lyon un logiciel qui détecte la probabilité pour un texte d'être généré par IAG, ce type d'outil n'est pas entièrement fiable et ne permet en aucun cas de prouver que l'étudiant s'est servi des IAG pour rédiger son devoir.

En plus de l'évaluation « sur table », qui est à l'évidence une évaluation qui limite presque entièrement (sauf « triche » classique) l'utilisation d'un outil tiers pour s'aider à rédiger, il existe d'autres modalités d'évaluation qui peuvent, par leur nature, limiter au minimum l'apport que les IAG peuvent amener aux étudiants.

3.2.1. Évaluations orales

Lors d'une évaluation orale, l'étudiant doit répondre en temps réel devant un enseignant ou un jury, justifier ses choix et ajuster son discours face aux questions ou relances. En plus d'être un type d'évaluation intéressant par le fait qu'il permet de vérifier que l'étudiant peut mobiliser une grande partie des notions fondamentales d'un cours donné, et donc de fournir des analyses complètes argumentées, les évaluations orales sont totalement résilientes aux IAG : même en cas d'utilisation d'un outil d'IAG lors de la phase de préparation de l'oral, une fois devant le jury l'étudiant doit **s'adapter aux questions directes et construire un argumentaire** qui doit refléter les connaissances et les compétences qu'il a acquises.

Cette forme d'évaluation met en avant trois dimensions :

- **Réactivité** : l'étudiant doit analyser rapidement une question, organiser ses idées et formuler une réponse cohérente.
- **Justification** : il ne suffit pas de donner une réponse correcte ; il faut expliquer pourquoi cette réponse est pertinente et comment elle s'insère dans le cadre étudié.
- **Adaptation** : l'enseignant peut poser des relances, demander des précisions ou présenter des objections, obligeant l'étudiant à ajuster son raisonnement en temps réel.

Les limites de ce genre d'évaluation est qu'il est chronophage, logistiquement plus compliqué (convocation d'un jury, grille de notation harmonisée entre les différents étudiants, multiplication des sujets qui se doivent d'être distincts d'un étudiant à l'autre, etc.), et qu'il ne permet pas, si c'est cela qu'on souhaite, de vérifier finement la

connaissance de tous les concepts ou notions prévues dans un programme de cours, puisqu'il s'agit surtout d'une évaluation globale.

3.2.2. Évaluations processuelles

Les **évaluations processuelles** se distinguent des évaluations classiques qui se concentrent sur le produit final (dissertation, rapport ou synthèse). L'objectif est de **suivre et évaluer le cheminement intellectuel de l'étudiant**, sa capacité à planifier, expérimenter, ajuster et réfléchir sur sa propre démarche.

Autrement dit, l'évaluation s'intéresse aux **stratégies employées, aux difficultés rencontrées et aux décisions méthodologiques** plutôt qu'à un résultat final parfaitement rédigé. Cela implique de rendre visibles les étapes de la pensée : hypothèses initiales, analyses intermédiaires, erreurs et ajustements successifs.

Les IAG peuvent produire un texte final cohérent, mais elles sont incapables de reproduire le parcours intellectuel propre à chaque étudiant, avec ses hésitations, expérimentations et ajustements successifs. Ce processus de réflexion, lié à l'expérience personnelle de recherche et aux choix contextuels, est unique et ne peut être simulé par une IAG. En se concentrant sur la démarche plutôt que sur le résultat final, vous rendez l'évaluation plus difficilement délégable à une IAG.

Ce type d'évaluation repose sur :

- **La documentation du processus** : l'étudiant consigne ses réflexions et décisions au fur et à mesure de l'avancement de son travail.
- **La réflexion sur les choix méthodologiques** : l'accent est mis sur la justification des méthodes employées et sur les ajustements effectués en fonction des obstacles rencontrés.
- **L'apprentissage visible** : l'enseignant évalue non seulement le résultat final, mais la progression, la capacité d'adaptation et la maturité méthodologique.

La limite de ce type d'évaluation est qu'il peut donner l'impression de doubler la charge de travail pour l'étudiant, qui doit à la fois rendre un devoir et documenter son parcours jusqu'à la rédaction finale. De plus, si la notation est axée sur le processus plutôt que sur le résultat, le risque est que l'étudiant se disperse en se concentrant plus sur la rédaction d'un carnet de bord que sur le résultat attendu. Enfin, rédiger un journal, formaliser ses choix et ses réflexions requiert une méthodologie particulière qui ne peut pas forcément être demandée à tous les étudiants, notamment ceux qui sont encore en début de cursus universitaire.

3.3. Adapter ses évaluations pour intégrer les IAG

A partir du moment où l'on accepte le fait que certains étudiants utiliseront des IAG pour produire des textes, résumer des informations ou organiser des idées, l'objectif peut alors ne pas être de « criminaliser » l'usage, mais de prévenir le contournement et de repenser l'évaluation pour qu'elle conserve sa valeur pédagogique.

3.3.1. Déplacer l'évaluation du produit vers le processus

Les IAG sont très utiles pour générer des textes finis, mais elles ne peuvent pas reproduire le cheminement intellectuel d'un étudiant. Comme on l'a vu juste avant, évaluer autant le processus que le résultat final incite l'étudiant à porter un regard réflexif sur ce qu'il fait, et demande un haut niveau de réflexion pour porter un regard critique et réfléchi sur sa méthodologie et son utilisation des IAG. **Quelles informations ont été générées par l'IA ? Quelles modifications ou choix personnels ont été apportés ? Pourquoi avoir fait générer une certaine partie du contenu par IAG ? Qu'est-ce que l'IAG a apporté à sa réflexion, ou au contraire qu'est-ce qui l'a limité ?**

3.3.2. Favoriser les évaluations interactives et en contexte

L'idée ici est de créer des situations où la pensée de l'étudiant doit s'exprimer en contexte, et où l'IAG ne peut pas fournir de réponses « toutes faites » : cas pratiques ou études de scénarios, conduite de projet en conditions réelles, choix argumenté parmi plusieurs options, etc. En contextualisant les évaluations à des données « locales », c'est-à-dire qui ne sont pas accessibles à des IAG, on réduit leur utilisation à la simple mise en forme ou rédaction. S'il faut interviewer des professionnels, observer une situation en réel, créer et expérimenter un protocole, etc., **l'IAG ne pourra être utilisée qu'à des fins de « mise en forme », et la réflexion ou la mise en pratique de connaissances ne pourra dépendre que de l'étudiant seul.**

3.3.3. Encadrer l'usage des IAG

Il est essentiel de définir clairement dans quels contextes les IAG peuvent être utilisées (organisation, brouillons, recherche documentaire) et ce qui doit rester personnel (analyse critique, justification, raisonnement). En formant les étudiants à l'usage critique des IAG (vérifier les sources, expliciter les choix, comparer les options, ...), on s'assure alors que les étudiants vont pouvoir utiliser les IAG en toute connaissance de cause, et en sachant leurs limites et leurs atouts. **L'objectif est de transformer l'IAG en outil pédagogique plutôt qu'en menace.**

En résumé, l'évaluation s'adapte aux IAG dès lors qu'elle met l'accent sur :

- la pensée située et contextualisée,
- le processus et la progression,
- la réflexion et la justification personnelle,

- et la créativité ou personnalisation du travail.

L'IA n'est alors plus un facteur de triche, mais un outil que l'on peut **intégrer dans un cadre pédagogique réfléchi**.

4. Conclusion

La dissémination importante des IAG dans les environnements numériques, et leur démocratisation express, imposent désormais de **réfléchir à leur impact dans l'évaluation des étudiants**, pour assurer un résultat valable et représentatif du niveau des étudiants, notamment dans le contexte universitaire dans lequel, dans la grande majorité du cursus, la "note finale" fait foi.

Il existe plusieurs angles d'approche : soit on choisit d'empêcher l'utilisation des IAG dans le processus d'évaluation, en proposant des formats et des modalités qui ne permettent pas (ou seulement à la marge) leur utilisation pour produire un devoir, soit on choisit de les intégrer en acceptant leur utilisation sous des règles déclarées et en imposant de justifier de leur utilisation.

Dans tous les cas, il est nécessaire de les **prendre en compte**, et de **communiquer le plus précisément possible** avec ses étudiants à leur sujet.

C'est désormais la fin de ce dossier pédagogique consacré aux IAG dans l'enseignement supérieur. Si vous souhaitez apporter des ajouts, faire part de votre expérience dans le domaine, ou simplement partager des ressources ou des avis sur le sujet, vous pouvez me contacter : emilie.leroy@ens-lyon.fr