

Rubrique multimédia (version numérique) de Repères-IREM n°143

Gérard Kuntz (g.kuntz@libertysurf.fr)

La rubrique multimédia est confiée dans ce numéro 143 de Repères-IREM à Grégory Quiquempois et Philippe Goémé, enseignants et formateurs dans le domaine de l'IA à l'INSPÉ de Créteil. Ils présentent leur philosophie et leurs approches dans l'article qui suit.

L'IA à l'INSPÉ de Créteil

gregory.quiquempois@u-pec.fr et philippe.goeme@u-pec.fr

Comme beaucoup, en particulier dans le monde de la formation et de l'éducation, nous avons été percutés en tant que formateurs INSPÉ, par l'avènement de l'IA dans le paysage « grand public », en novembre 22. Certes, les travaux et les réflexions existaient bien avant, mais ce qui est apparu pour nous, à ce moment-là, c'est, non seulement ses capacités, mais aussi son extrême facilité d'accès et d'usage. Nous sommes passés en quelques semaines, d'un domaine pointu, réservé à quelques spécialistes ultra compétents, à un outil très facilement accessible à tous. Cette prise en main, extrêmement simple au premier abord, n'est pas sans conséquences dans notre domaine, la formation d'enseignants.

Nous sommes, dans ce cadre, confrontés à une triple contrainte : Il nous faut intégrer l'IA en tant qu'enseignant du supérieur, avec ce que cela implique en termes de ressources pour nos étudiants, mais aussi de possibilités d'échapper à ce qui fait sens dans un enseignement universitaire : productions d'écrits non assimilés, jeux de référence non maîtrisés, vacuité des travaux rendus... D'autres parts, comme nous formons de futurs enseignants, la formation à l'IA, en toute congruence, doit aussi leur permettre tant une maîtrise technique qu'une élaboration de ce qu'ils pourront en faire, ou non, face à une classe.

Enfin, il faudra qu'ils intègrent le cadre général d'usage construit au fil de l'eau par les diverses institutions, dont l'Éducation Nationale et notre université de rattachement toujours en retard (et cela est inévitable), sur les possibilités offertes par cet outil. Si ces cadres d'usages se construisent petit à petit (un cadre vient d'être diffusé par l'Éducation Nationale en juin 2025), cela est plus compliqué pour les universités, même si le travail est lancé.

C'est donc à cette intersection, que nous avons imaginé les scénarii de formation qui suivent. Il n'est pas question de nier les dimensions politique, économique, écologique, philosophique qui sont opposées à l'usage de cette technologie. Ces problématiques sont réelles, et doivent être intégrées, mais l'objet de cet article est de rendre compte du cheminement que nous avons suivi, pendant presque deux ans pour intégrer l'IA à nos formations et les différentes étapes que nous avons suivies.

Ces étapes correspondent aux différents stades par lesquels nous sommes passés, de l'évidence de prompts immédiats jusqu'au travail de personnalisation de différents agents IA, via des méta prompts précis et l'import de corpus sélectionnés par nos soins.

Un effet Waouh...

La première étape a consisté à apprivoiser, dans le cadre de formation, et dans un premier temps dans une démarche de simple partage, les possibilités offertes par les IA à disposition du grand public. L'objectif était de découvrir, et de faire découvrir les potentialités de ce nouvel outil, immédiatement disponible et dont la prise en main, au premier abord, est très simple. Les propositions de transposition d'un document dans un autre support, l'usage facile des fonctions orales, les propositions de traduction (nous avons mené une conférence en islandais en traduction

directe), la possibilité d'obtenir des séances ou des séquences, des corrections de copies... nous avons donc organisé des séances de démonstration et de manipulation (le distanciel s'y prête très bien), avec l'élaboration de tutoriels accessibles à tous et une forme de « veille » en mettant à disposition un certain nombre de ressources sous forme d'articles régulièrement actualisées. Cette étape indispensable, a provoqué un ensemble de questions légitimes quant aux limites de l'outil, malgré ses possibilités indéniables. La deuxième étape a consisté à tenter de supprimer, autant que faire se peut, les hallucinations et les biais.

Waouh...mais so what ?

Rapidement donc, les biais et hallucinations, bien connus des utilisateurs (imprécision des réponses, faiblesse des sources, réponses stéréotypées en lien avec les données d'entraînement, erreurs manifestes...) ont été mises en avant par les utilisateurs testeurs.

L'usage de photos générées par IA a pu servir de façon efficace à illustrer ces biais d'usages : une hallucination est visible, une manipulation de réalité est patente, une incompréhension est visible, une impossibilité révélée. De plus, la limitation d'accès à travers les offres gratuites a été rapidement un vrai frein. Il a donc été nécessaire de passer à une phase supplémentaire consistant à agir directement sur l'IA elle-même.

Problème : personnaliser les IA ne répond pas au problème de la limitation !

Dans le même temps, nous avons donc testé les différentes propositions d'IA afin de trouver celles, dans la multitude des offres, qui permettaient de répondre au mieux à nos besoins. Nous sommes donc passés progressivement de ChatGPT à Mistral puis à Perplexity, pour des raisons techniques, mais aussi juridiques (un accord potentiel de l'[UPEC](#) avec Mistral).

Une première réponse : garantir la qualité de la source en important nos propres corpus.

La première étape a consisté à créer et implanter nos propres corpus de référence, en lien avec la formation de futurs enseignants, afin d'éviter les références douteuses pouvant être trouvée par l'IA dans son balayage de sources sans contrôle. La première IA ainsi personnalisée a été alimentée par tous les textes fondamentaux qui régissent l'action des enseignants, du code de l'éducation à la déclaration des droits de l'Homme, de la circulaire de rentrée à l'intégralité des ressources « Laïcité et Valeurs de la République » du site Eduscol en passant par les droits et devoirs des élèves et des fonctionnaires. Au fur et à mesure de nos tests et de la découverte des limites de nombre de fichiers importables dans les différentes IA, nous avons procédé à une fusion des documents source (.docx). Ces fichiers consolidés ont ensuite été convertis en formats plus légers et compatibles (Markdown) de manière à optimiser leur ingestion par le modèle et à garantir une alimentation cohérente et structurée de l'IA. La configuration des agents permettant de ne pas autoriser l'accès à Internet pour répondre à nos prompts nous a ainsi assuré d'une recherche exclusive sur le corpus dédié.

Une deuxième étape : le méta-prompt pour guider la réponse dans notre cadre professionnel

« Prompter » est à priori très simple, puisque nous sommes dans un modèle « conversationnel ». Cependant, pour gagner en précision dans les réponses attendues, il est vite apparu qu'il était indispensable de retravailler la première réponse obtenue. Il nous a donc semblé évident qu'il était nécessaire d'éviter ces étapes de tâtonnement en pré-programmant le prompt afin que l'IA nous pose de lui-même une série de questions, sur la base d'une dichotomie de choix, afin de guider la construction de précision des demandes pour obtenir la réponse la plus adaptée (Annexe 1)

Phase 3 : Passer en local

Si la personnalisation des IA grand public nous a permis de réduire considérablement les risques d'hallucinations des IA, nous restons limités par deux contraintes juridiques majeures, à savoir le respect primordial à nos yeux du RGPD et du droit d'auteur des éléments que nous pouvons importer. En l'attente d'une solution souveraine, la création d'IA en local (sans connexion à Internet) nous a semblé être, à l'heure actuelle, la seule solution légale pour passer outre ces

limitations.

LM studio a été le premier logiciel utilisé pour télécharger, gérer et lancer différents modèles de langage open source (comme [LLaMA](#), [Mistral](#), [Falcon](#), etc.), avec une interface utilisateur proche de ChatGPT. Si les premiers résultats sont satisfaisants, l'import des corpus demeure fastidieux ... AnythingLM, récemment découvert nous permet d'aller encore plus loin, car il offre les mêmes possibilités que ceux précédemment cités, mais en plus il permet de mettre en place des systèmes de type RAG ([Retrieval-Augmented Generation](#)). Concrètement, une fois différents fichiers importés (.doc, .pdf, base de données, etc..), il va les vectoriser (via un [moteur d'embedding](#)) et les stocker dans une base vectorielle.

Ainsi lorsqu'une requête est posée, l'outil effectue une recherche contextuelle (retrieval) dans ces documents, puis injecte les extraits pertinents dans le prompt du modèle (augmentation du contexte). Le modèle de langage (LLM) peut alors générer une réponse enrichie, fiable et contextualisée à partir des données fournies.

Il nous reste à résoudre la question des données à importer pour mettre en place le RAG, l'IA utilisée en local n'ayant pas accès à Internet pour consulter par exemple les programmes officiels d'enseignement.

A l'heure où nous écrivons ces lignes, ce sont plus de 12000 PDF du site Eduscol pour assister les enseignants du cycle 1 au cycle 4 qui ont été répertoriés, téléchargés et fusionnés pour alimenter la base du RAG qui pourra ensuite être enrichie par chacun sans limitation juridique.

Conclusion.

S'approprier l'IA et ses possibilités ne peut se résumer à un apprentissage de techniques sédimentées. Parce que les possibilités offertes évoluent quasi quotidiennement, il est impossible de s'appuyer sur une base technique fixe et transposable. Utiliser l'IA demande donc des formes de mises à jour permanentes, impliquant une posture particulière, attentive aux évolutions aujourd'hui très rapides du système. Cela est vrai sur le plan technique, mais aussi sur le plan juridique, politique, philosophique, économique .

Post-scriptum de G. Kuntz

Il est important de prolonger cet article qui dit l'essentiel, par un document signé par les mêmes auteurs, plus vaste et très détaillé, capable aussi d'épouser les évolutions fulgurantes des IA. Il est accessible sur le site de l'INSPÉ de Créteil. Il accompagnera utilement celles et ceux qui voudront se former sérieusement à ces outils, pour les utiliser de façon experte dans le cadre très encadré et régulé de l'Éducation Nationale. Nous vous le recommandons vivement.

Vous pourrez télécharger le PDF de ce document de 98 pages [à cette adresse](#).