

Rubrique : Expériences et pratiques exemplaires

DÉPLOYER UNE STRATÉGIE D'USAGE DES OUTILS D'IA GÉNÉRATIVE DANS UN ÉTABLISSEMENT D'ENSEIGNEMENT SUPÉRIEUR. LE CAS DE L'UNIVERSITÉ DE RENNES

Pierre BEUST

Professeur en Informatique et Directeur du SUPTICE
Université de Rennes, Campus de Beaulieu, France

Olivier WONG-HEE-KAM

Vice-président délégué au Numérique
Université de Rennes, France

Résumé

L'arrivée des outils d'intelligence artificielle générative transforme de manière importante les différents métiers de l'enseignement supérieur. Nous présentons dans cette contribution une démarche expérimentale en cours que nous menons à l'échelle de notre établissement, l'Université de Rennes. Notre démarche s'appuie sur une infrastructure locale permettant d'utiliser différents modèles d'intelligence artificielle générative. L'objectif est d'apporter un environnement de confiance pour nos collègues et de réfléchir avec eux à des éléments de cadrage. Notre problématique est de trouver le meilleur accompagnement possible de nos collègues face aux transformations potentielles de leurs pratiques professionnelles induites par l'intelligence artificielle.

Abstract

The generative artificial intelligence tools are significantly transforming the various professions involved in higher education. In this contribution, we present an experimental approach that we are currently implementing at our institution, the University of Rennes. Our approach is based on a local web platform that allows the use of several models of generative artificial intelligence. The aim is to create a trustworthy environment for our colleagues and to work with them to develop a framework. Our challenge is to find the best possible way of supporting our colleagues in the face of the potential transformations in their professional practices brought about by artificial intelligence.

Mots-clés

IA générative, Enseignement supérieur, Expérimentation

Key words

Generative AI, Higher Education, Experimental Approach

Pour citer cet article : BEUST P. et WONG-HEE-KAM O. (2025). « Déployer une stratégie d'usage des outils d'IA générative dans un établissement d'enseignement supérieur. Le cas de l'Université de Rennes », *Télescope SMP&OE*, vol. 22., décembre 2025, Dossier spécial.

1. Introduction

Cet article porte sur l'utilisation de l'Intelligence Artificielle (IA) dans une organisation publique, l'Université. Notre démarche relève d'une approche empirique. Nous décrivons comment nous mettons en place de manière expérimentale à l'échelle de notre université une démarche collective d'appropriation des concepts et des cas d'usage des outils d'Intelligence Artificielle générative (IAg). Cette expérimentation s'appuie sur un environnement de confiance, aussi bien dans les aspects techniques informatiques d'une plateforme déployée en local que dans les aspects de collégialité humaine autour de temps d'échanges, de partages, de discussions et de controverses. Notre problématique est de bien accompagner l'ensemble de nos collègues pour une intégration éclairée et raisonnée des outils de l'IAg dans leurs pratiques professionnelles.

Dans une première partie, nous présentons le contexte de notre institution publique,

l'Université de Rennes, notre problématique et notre ancrage épistémologique. Dans une deuxième partie, nous détaillons les actions engagées dans notre expérimentation locale sur les outils d'IAg appelée « RAGaRenn ». Nous décrirons ensuite quelques éléments d'analyse des coûts, des risques et des bénéfices. Enfin, nous donnerons quelques perspectives et éléments de conclusion.

2. Le contexte

Nous détaillons dans cette partie notre contexte local, notre problématique et notre ancrage parmi les travaux sur l'IA et la conduite du changement.

2.1. L'Université de Rennes

L'Université de Rennes est un établissement d'enseignement supérieur public français. Elle regroupe des composantes de formation pluridisciplinaires (sciences, droit, économie-gestion, santé, philosophie), quatre écoles d'ingénieurs¹, des instituts universitaires de technologie et trois écoles

¹ École Supérieure d'Ingénieurs de Rennes (ESIR), Institut National des Sciences Appliquées (INSA), École Nationale Supérieure des Sciences

Appliquées et de Technologie (ENSSAT), École Supérieure de Chimie de Rennes (ESCR)

supérieures². L'établissement est tutelle ou co-tutelle avec des organismes nationaux de 38 structures de recherche. L'Université de Rennes totalise plus de 37 000 étudiants, 4 500 personnels répartis sur 9 campus sur l'agglomération rennaise et la Bretagne. En 2024, l'Université de Rennes se positionne dans le top 401-500 du classement de Shanghai des meilleures universités du monde (seul établissement breton à figurer dans ce top 500). La signature scientifique de l'Université de Rennes s'articule autour des 3 principales transitions actuelles de la société : la transition environnementale, la transition en santé et la transition numérique.

C'est en lien direct avec cette signature et son excellence sur les sujets du numérique que l'université s'est lancée en 2024 dans une vaste expérimentation à l'échelle de l'établissement sur les outils d'IAg, que nous nommerons "les IAg" par la suite.

Notre établissement est bien identifié sur les questions d'IA. En témoignent notamment le projet lauréat de l'appel à manifestation d'intérêt "Compétences et Métiers d'Avenir" (CMA³), TIARe⁴ (Toute l'Intelligence Artificielle à Rennes) et le projet lauréat de l'appel à projets national IA Cluster intitulé Cluster SequoIA⁵.

2.2. Problématique

Les IAg impactent la plupart des activités de la société contemporaine, qu'il s'agisse des activités professionnelles (Alcantara *et al.*, 2020), sociales (Petit, 2021) ou encore culturelles (Ganascia, 2017). Face à l'arrivée "fulgurante" selon (Marty *et al.*, 2024) des outils d'IAg dans la société (principalement l'émergence de ChatGPT en 2022, suivie de beaucoup d'autres outils sur la génération de textes, d'images, de musique, de vidéos, ...), les établissements d'enseignements supérieurs réagissent de manières très variées. Certains, comme SciencesPo-Paris, ont prononcé une interdiction au recours à ChatGPT dans la production d'évaluations écrites ou orales⁶. Beaucoup d'autres adoptent des stratégies attentistes en observant et en suivant tant que possible les évolutions. Notre démarche se veut volontariste, expérimentale et progressive. Elle cherche la construction d'une position claire et étayée par la mise en action qui ne laisse pas les collègues dans la non-réponse. Cela rejoint l'avis du Russel Group (Russel Group, 2023) : *"plutôt que de chercher à interdire leur utilisation des IAg, les [élèves et] étudiants et le personnel doivent être*

² École Normale Supérieure (ENS), SciencesPo Rennes, École des Hautes Études en Santé Publique (EHESP)

³ Projet sur financement France 2030, pilotage Agence Nationale de la Recherche (ANR)

⁴ <https://intelligence-artificielle.univ-rennes.fr/tiare>

⁵ <https://actu.univ-rennes.fr/actualites/le-projet-sequoia-ia-cluster-confiance-et-securite-labellise-pole-dexcellence-en>

⁶ <https://www.sciencespo.fr/fr/actualites/sciencespo-fixe-des-regles-claires-sur-lutilisation-de-chat-gpt-par-les-etudiants/>

soutenus dans l'utilisation efficace, éthique et transparente de ces outils".

Notre problématique est de travailler à une encapacitation de nos collègues (on emploie aussi régulièrement l'anglicisme *empowerment*) au sens de Ganascia, (2017) ne les laissant pas sans moyens et accompagnement de l'établissement. Il est important d'avoir une forme de prise de conscience des retombées potentielles des IAg dans les métiers présents à l'université (enseignants, chercheurs, personnels de bibliothèque, personnels administratifs, ...) et de se projeter dans des cas d'usage intéressants. Mais cette démarche est compliquée si elle n'est pas prévue dans le temps de travail et sans préconisations claires. Nous voulons absolument éviter l'injonction contradictoire que ressentiraient les collègues que l'on résumerait de la sorte "transformez-vous, mais sans assistance". Nous rejoignons en cela l'avis de Christèle Lison (professeure à l'Université de Sherbrooke, présidente de l'Association internationale de pédagogie universitaire (AIPU international) : *"Pour moi, le piège pour un établissement serait de dire que le sujet relève de la gestion individuelle."*⁷

Pour les auteurs du présent article, ne pas être dans l'injonction contradictoire consiste à proposer un environnement de confiance.

C'est un environnement technique (avec des outils en interne à l'établissement) et humain (temps d'appropriation, accompagnement, partages) qui mise sur le collectif. Dans notre problématique, l'encapacitation n'est pas tant individuelle que collective. La transformation des postures professionnelles se joue dans des échanges à propos des cultures métiers.

Enfin, notre problématique cherche à construire avec l'ensemble de nos collègues une vision commune des impacts des IAg éclairée qui ne soit ni dans l'adoption irréfléchie de la technologie ni dans son refus systématique. Nous cherchons à nous inscrire dans une position mesurée entre le techno-solutionnisme et la techno méfiance.

2.3. Ancrage épistémologique

Notre ancrage épistémologique se situe aux intersections des travaux d'origines disciplinaires variées (management et éducation principalement) s'intéressant à la conduite du changement avec l'IA.

Dans le champ disciplinaire du management, de nombreux travaux insistent sur les changements importants dans les métiers avec l'IA. Beaucoup y voient des opportunités pour notamment affirmer les

⁷ <https://www.aefinfo.fr/depeche/709351-les-ia-peuvent-etre-une-aide-a-la-correction-sur-des->

[contenus-basiques-et-repetitifs-en-debut-de-licence-c-lison](#)

positions de leadership⁸, engendrer des gains de temps⁹, accroître l'efficacité organisationnelle¹⁰. Si certains y voient effectivement du positif, d'autres (40 % d'après une étude du Boston Consulting Group¹¹) redoutent qu'elle fasse disparaître des métiers. Le Fonds Monétaire International (FMI) considère que 60 % des emplois seront impactés d'une manière ou d'une autre par l'IA dans les prochaines années¹².

En affectant ainsi les organisations, en modifiant les métiers (ce qui ne veut pas dire nécessairement les faire disparaître), l'IA impacte le rapport au travail, ainsi que la manière dont les équipes collaborent. Face à cela, nous cherchons à déployer un environnement rassurant pour aborder ces questions. Nous sommes dans une approche empirico-inductive (Glaser et Strauss, 1967 ; Bryman, 2012), qui prône l'ancrage dans le terrain, et plus concrètement la réalisation d'observations et d'expériences. Nous avons aussi à engendrer de la confiance là où les IA sont généralement entraînées avec des données que l'on ne connaît pas et qui sont

généralement récupérées sans autorisation sur Internet (Chesterman, 2023).

À l'université, les personnels de la documentation, notamment, ont bien perçu les changements qui se jouent. Il faut former les personnels de bibliothèques (Hussain, 2023) pour former les usagers, ce qui est stratégique selon Paivandi et ses collaborateurs (2016). Il est important d'accompagner les usagers, de leur permettre de constituer une culture documentaire et informationnelle, de développer une posture de recherche, avec des approches réflexives et novatrices (Frisch, 2015).

Dans le champ de l'éducation, comme dans le management, les avis oscillent entre les positions résolument optimistes et les travaux plus pessimistes. Parmi les optimistes, certains mettent en avant la qualité de l'expérience éducative (Ayeni *et al.*, 2024) et d'autres pointent un accès plus rapide et plus direct à l'information (Crevier *et al.*, 2023 ; Kooli, 2023 ; Uzan *et al.*, 2024). Une autre étude (Li, 2023) se basant sur la théorie de l'action raisonnée développée par Fishbein et Ajzen (1975), et

⁸ <https://www.linkedin.com/pulse/lintelligence-artificielle-au-service-des-managers-r%C3%A9volution-ttg8e/>

⁹ <https://courriercadres.com/intelligence-artificielle-a-quoi-ressemblera-le-management-augmente/>

¹⁰ <https://www.scbs-education.com/intelligence-artificielle-et-management-une-revolution-en-marche/>

¹¹ <https://www.bcg.com/press/26june2024-ia-generative-au-travail-amie-ou-ennemie-des-salaries>

¹² <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379>

s'appuyant sur le modèle d'acceptation des technologies (Davis, 1989), montre que la perception de facilité d'utilisation des étudiants à l'égard des systèmes basés sur l'IA a un impact significatif et positif sur l'utilité perçue. Les pessimistes, quant à eux, rappellent que l'usage des IAg n'est pas sans risque. La fiabilité est souvent hautement discutable (Wu *et al.*, 2023) du fait des biais sociaux engendrés par des données d'entraînement (Weinberger *et al.*, 2020) et des phénomènes d'hallucination des IA (Khowaja *et al.*, 2024). L'IA peut ainsi contribuer à une forme de désinformation en raison du renvoi de données inappropriées ou décontextualisées (Currie, 2023). La formation des utilisateurs des IAg, tous profils confondus, est un point de vigilance particulièrement important. L'étude exposée dans (Lee *et al.*, 2025) soulève des questions inquiétantes sur le déploiement des outils d'IAg dans le service public, notamment avec la mise en évidence d'une baisse de la vigilance qui pourrait se développer du fait de la qualité « apparente » des productions de l'IAg et d'une diminution de la pensée critique. Il faut donc s'assurer, par des actions de formation (notamment sur les questions d'hallucinations inhérentes aux IAg), que les agents maintiennent un niveau élevé de pensée critique. Sinon, le

déploiement d'outils d'IAg pourrait s'avérer contre-productif.

Entre pessimisme et optimisme, la mesure et le réalisme semblent de bons ancrages. (Marty et Vasquez, 2024) soulignent l'importance de penser une intégration critique et responsable des IAg dans l'enseignement universitaire. Dans cette idée, nous accompagnons les enseignants à se questionner, se montrer lucides, objectifs et nuancés vis-à-vis de ces outils. Il convient de considérer à la fois les atouts et les limites des IAg afin de proposer aux étudiants une éducation aux IAg tout aussi critique et nuancée.

Il est primordial que les établissements d'enseignement s'inscrivent dans une transformation pédagogique incluant les outils d'IAg. Il en va du risque d'une double fracture numérique. Premièrement les étudiants s'emparent de plus en plus des outils d'IA dans l'objectif de gagner en performances (Lobet *et al.*, 2023)¹³.

Certains étudiants payent des outils à des coûts mensuels non négligeables (allant de 20 à 200 €/mois pour un accès personnalisé à ChatGPT). D'autres encore prennent le risque de diffuser des données personnelles importantes sans sécurisation. Deuxièmement, il est à craindre que l'usage

¹³ L'étude indique que 35 % des étudiants du panel affirment rédiger tout ou une partie de leurs devoirs grâce à des outils d'IAg et que 47 % des

étudiants affirment qu'ils ont de meilleures notes grâce à l'IAg.

des outils d'IAg ne soit pas aussi positif d'un étudiant à l'autre, s'ils ne sont pas accompagnés. Cela risque de rendre les « bons » étudiants encore meilleurs et les plus « faibles » encore plus en difficulté, car il faut savoir déceler les hallucinations, savoir « prompter » efficacement les systèmes utilisés, savoir comparer les productions et reformuler les propositions obtenues. Il faut intégrer les usages des IAg dans l'accompagnement des enseignants, dans leur pédagogie. Cela conditionne l'employabilité des étudiants. Ce n'est pas une compétence que les étudiants développeront par eux-mêmes en dehors de leurs enseignements reçus. L'usage des IA doit être une opportunité pour questionner les enseignements disciplinaires. Comme l'indique le rapport de l'UNESCO de 2024 : *« On peut craindre que les jeunes apprenants, parce qu'ils sont par définition moins expérimentés que les enseignants, n'acceptent sans le savoir et sans réflexion critique des résultats obtenus par l'IA générative qui sont superficiels, inexacts ou même nuisibles ».*

Pour intégrer en toute confiance les outils d'IAg dans la pédagogie des étudiants comme dans les usages des chercheurs, il faut en contrôler les impacts en matière de risques : risque de fuite de données, risque d'usages éthiquement non contrôlés, risque d'inflation injustifiée des impacts environnementaux, risque d'amplification

des pertes de souveraineté numérique, ... Pour cela il faut se doter d'infrastructures et d'offres de services émanant de la communauté de l'Enseignement Supérieur et la Recherche (ESR). La mutualisation des moyens et le partage des pratiques transformantes sont la bonne stratégie. Il faut se donner un écosystème où l'avance de certains acteurs de l'ESR ne soit pas vue comme un leadership concurrentiel, mais une transformation globale potentielle par essaimage et mutualisation de moyens sur l'ensemble de la communauté de l'ESR.

Localement, nous nous inscrivons dans l'héritage d'une vision rennaise de la transformation numérique telle que décrite par Pascal Plantard où la transformation *« s'appuie sur des logiques coopératives et systémiques qui encouragent le co-design participatif, l'engagement des acteurs et l'open-science »* (Plantard et al., 2022, p.21).

3. L'expérimentation RAGaRenn

L'expérimentation de l'outil d'Intelligence Artificielle RAGaRenn a été impulsée en mars 2024 par une circulaire interne du Président de l'Université de Rennes. L'objectif est d'établir un cadrage clair des usages des outils d'IAg prescrits et proscrits dans les différents métiers présents à l'université.

RAGaRenn comprend plusieurs actions :

- La volonté d'établir collectivement un cadrage des usages des outils d'IAg (pour le pilotage, la gouvernance, la recherche, la formation) et de le faire acter par nos instances.
- La mise à disposition d'une infrastructure locale¹⁴ d'inférence de *Large Language Models* (LLM) et de service(s) numérique(s) pour que les collègues puissent s'acculturer (en toute sécurité et contrôle) aux usages de LLM *open-weights* (à poids ouverts). Nous leur mettons à disposition plusieurs LLM pour multiplier les cas d'usage (des "gros", des "petits", des généralistes, des spécialisés, par exemple pour la génération de code).
- Des ateliers collaboratifs (en modalité hybride présence/distance) pour échanger sur les attentes, les craintes, les opportunités, les risques et partager autour de cas d'usages. Ces ateliers sont aussi, pour ceux qui en ont besoin, un temps de formation aux principes sous-jacents aux IAg et d'acculturation.
- Des journées scientifiques de partage aux échelles locale et nationale.
- Des services numériques déployés au-dessus de l'infrastructure locale d'inférence LLM pour des cas d'usages spécifiques : simulation d'entretiens d'embauche, tuteurs virtuels dans les espaces de cours, création assistée d'activités pédagogiques (notamment H5P pour intégration dans Moodle).
- Un lien fort vers la mutualisation inter-établissements, au travers du projet "ILaaS", coordonné par l'Université de Rennes et financé par le ministère de l'Enseignement Supérieur et de la Recherche (MESR).

Les différentes étapes du déploiement ainsi que la feuille de route envisagée sont résumées dans la frise chronologique ci-dessous (Figure 1).

¹⁴ <https://ragarenn.eskemm-numerique.fr/>

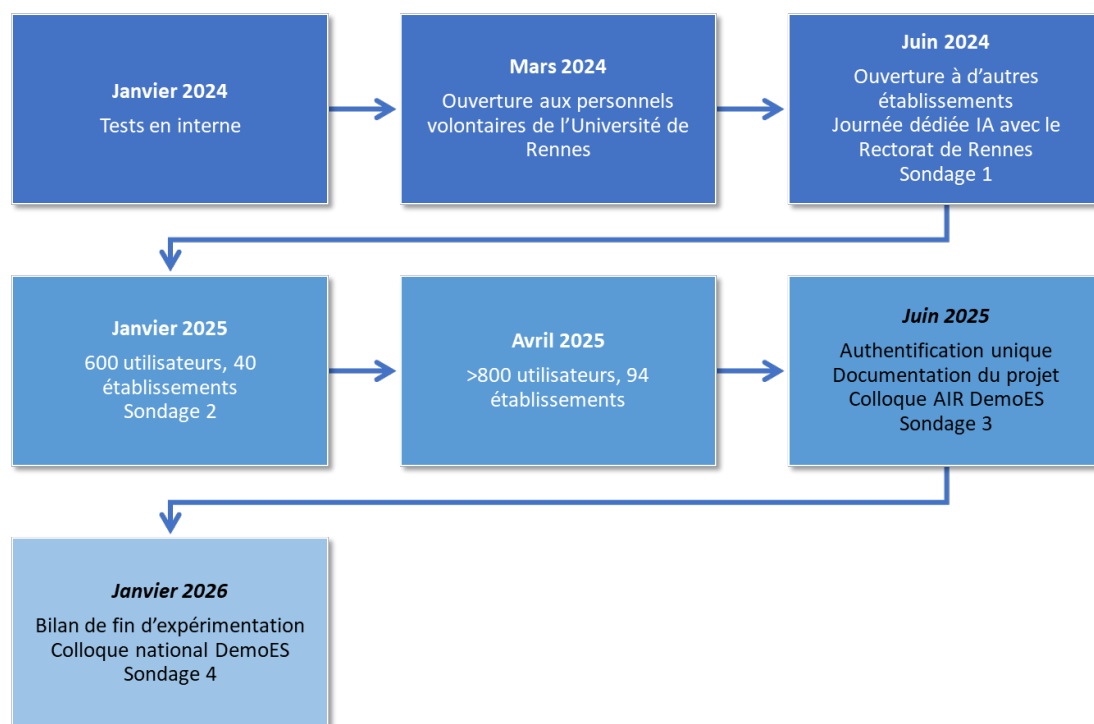


Figure 1. : Frise chronologique du déploiement de RAGaRenn

3.1. L'infrastructure locale

L'infrastructure locale est déployée grâce aux moyens de calcul et de stockage d'Eskemm Data, centre de données breton de l'enseignement supérieur labellisé faisant partie de la Fédération de centres de données régionaux construite par le ministère de l'Enseignement supérieur et de la recherche. L'Université de Rennes est un des membres fondateurs du GIP Eskemm Numérique qui assure le portage d'Eskemm Data : ce haut niveau de collaboration nous permet par exemple d'envisager de quantifier les

impacts environnementaux associés aux usages constatés, sujet qui est trop rarement abordé par des fournisseurs tiers.

Techniquement, nous utilisons des cartes GPU Nvidia : 2 H100¹⁵ et 4 V100, qui sont utilisées par deux types de moteurs ou serveurs d'inférence : Ollama¹⁶ et vLLM¹⁷. Ces serveurs d'inférence sont utilisés par des services numériques. Le premier service que nous avons déployé auprès des personnels, Open WebUI¹⁸, permet d'exposer une interface permettant le « prompting » et le « RAG¹⁹ » à l'utilisateur (cf. Figure 2, Figure 3). L'ensemble est intégralement

¹⁵ <https://www.nvidia.com/fr-fr/data-center/h100/>

¹⁶ <https://ollama.com/>

¹⁷ <https://github.com/vllm-project/vllm>

¹⁸ <https://github.com/open-webui/open-webui>

¹⁹ Retrieval Augmented Generation ou Génération augmentée par la récupération d'informations

construit à l'aide de briques libres (*open-source*) aussi bien pour les outils de déploiement permettant de gérer les machines virtuelles et conteneurs, que pour les LLM installés, qui eux sont généralement à poids ouverts (*open weights*) téléchargés via la plateforme Hugging Face²⁰.

L'appellation "RAGaRenn" repose initialement sur l'approche de RAG que nous avons souhaité mettre en avant prioritairement, sans insister sur l'utilisation d'Open WebUI que nous n'étions pas certains alors de garder à court terme

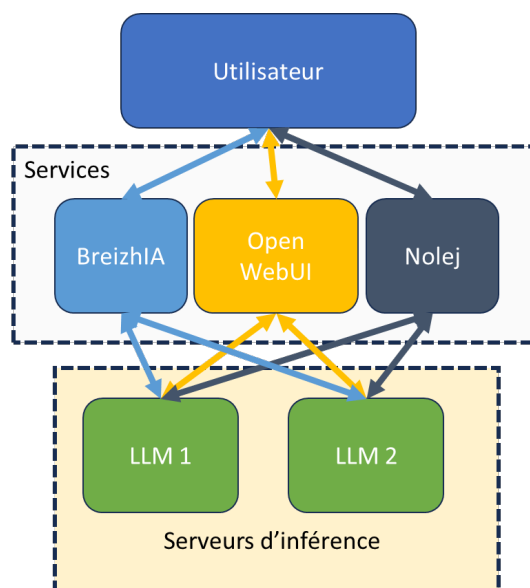


Figure 2. : Schéma de principe de RAGaRenn

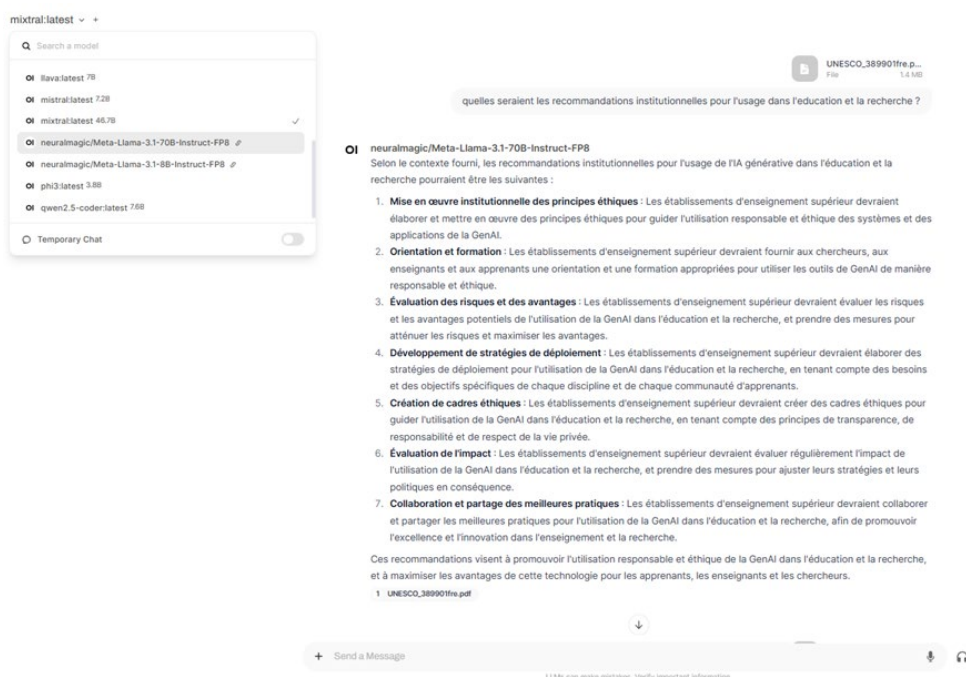


Figure 3. : l'interface utilisateur de RAGaRenn

²⁰ <https://huggingface.co/>

À ce stade, il nous semble important de préciser dans la suite de cette partie certaines notions à destination des lecteurs qui ne seraient pas familiers des concepts informatiques cités précédemment. Cet effort de pédagogie sur des concepts scientifiques et techniques pointus est nécessairement contraint et pourrait paraître trop réducteur pour des lecteurs plus experts.

La technique de **RAG ou Retrieval Augmented Generation** vise à exploiter la documentation interne en maîtrisant sa diffusion. Par rapport à un fonctionnement usuel où l'utilisateur fournit un prompt (contenu initial) à un LLM (Large Language Model) afin d'obtenir une réponse, le RAG introduit des étapes supplémentaires :

- Retrieval : l'outil cherche dans une base de documents, externe au LLM, des informations en lien avec le prompt initial saisi par l'utilisateur, puis sélectionne des extraits pertinents.
- Augmented Generation : le prompt initial est complété (augmented) par ces extraits, ce qui permet au LLM de générer une réponse plus précise, car ancrée dans des documents vérifiables.

Dans une telle approche de RAG, le LLM n'est pas modifié (pas de réentraînement ou *fine-tuning*) par les documents ou leurs extraits fournis, ce qui permet un cloisonnement des données entre utilisateurs d'une même infrastructure.

Un **LLM** (*Large Language Model* ou Grand Modèle de Langage) est un réseau de neurones profond de "grande taille" en raison du grand nombre de paramètres utilisés : il permet de générer le mot probable suivant à partir d'une séquence de mots fournie en entrée (requête ou *prompt*). Un réseau de neurones est un ensemble organisé de neurones artificiels interconnectés permettant la résolution de problèmes complexes. Un paramètre est généralement un poids associé à une connexion entre les neurones dans le réseau de neurones du modèle. Par exemple, en lisant le nom du modèle "Mistral Small 24B" on peut comprendre que Mistral (éditeur français) a publié un modèle nommé "Small" qui a 24 milliards de paramètres : le nombre devant B (qui signifie milliard de paramètres) représente la "taille" du modèle. Cette taille reste un premier indicateur des qualités intrinsèques du LLM, où les LLM de grande taille affichent des performances supérieures. De nouvelles avancées technologiques en 2025 ont toutefois remis en question cette approche de la performance, car des modèles récents plus légers égalent, voire dépassent, les performances de modèles antérieurs plus lourds.

Pour utiliser un LLM, il est courant d'utiliser une interface qui peut être locale sur un poste d'un utilisateur ou qui peut être mutualisée via un serveur web et rendue accessible à

plusieurs utilisateurs d'un même établissement, par exemple. Nous avons choisi **Open WebUI** qui est un logiciel libre (*open-source*) que nous avons rendu disponible sur serveur, et qui permet de rendre les mêmes services que ceux qu'on trouve couramment dans des outils grand public similaires. Début 2024 c'était un des seuls outils à proposer des fonctionnalités de RAG, plusieurs mois avant que des acteurs commerciaux ne proposent eux aussi du RAG dans leurs interfaces.

Au fil des demandes, nous avons activé plusieurs fonctionnalités qui sont proposées par ce projet communautaire : recherche web, mode vocal, analyse d'image en entrée, canevas de prévisualisation et interaction sur certains formats (html, svg,...), création de bots personnalisés (custom GPTs) disposant de réglages et de bases documentaires spécifiques. La particularité d'Open WebUI est d'avoir des liens historiques forts avec le serveur d'inférence **Ollama**. Tout comme **vLLM**, ce sont des outils *open-source* dédiés à l'exécution ou l'inférence de LLM. Ils permettent d'interagir avec des LLM par une interface de programmation applicative (ou API pour *Application Programming Interface*) qui est exploitée par des services numériques tels que Open WebUI.

Ces outils se différencient par leur orientation : Ollama, en lien avec Open WebUI, offre par exemple des facilités pour

déployer des LLM, mesurer et établir des statistiques. vLLM est plus orienté sur la vitesse d'inférence qui est mesurée en nombre de mots (ou *tokens*) par seconde : nous l'utilisons pour sa capacité à servir un LLM de façon plus productive à plus grande échelle.

Le choix d'un LLM est formulé selon les cas d'usages envisagés et selon les demandes des utilisateurs, en sélectionnant systématiquement des LLM à poids ouverts depuis *huggingface*. Après avoir testé de nombreux modèles durant la phase pilote, nous avons opté pour les propositions suivantes à partir de septembre 2024 :

- par défaut un « petit » LLM de type Llama 3.1 8B qui représentait à l'époque un compromis intéressant en le déployant sur vLLM : pour une qualité de réponse satisfaisante sur des sujets généralistes, sa taille relativement modeste nous a permis de servir un plus grand nombre d'utilisateurs en parallèle à des vitesses usuellement pratiquées par des acteurs commerciaux.
- pour des besoins plus poussés (qualité de réponse, contenu plus compliqué, etc.), nous avons déployé un modèle plus conséquent sur vLLM : Llama 3.1 70B, dont la vitesse moindre reste néanmoins dans les standards actuels.

D'autres LLM sont déployés suivant des cas d'usages spécifiques, sur Ollama :

- pour des besoins de programmation informatique, nous proposons différents LLM, qui comportent le terme « code » dans leur nom ;
- pour mieux tenir compte de spécificités linguistiques ou culturelles, nous proposons des modèles entraînés principalement sur des données en français (Lucie) ou en espagnol (Salamandra) dans une approche plus documentée, plus ouverte ;
- nous testons des SLM, Small Language Models, qui sont généralement de taille inférieure à 8B (Phi-3), mais généralement capables d'opérer des tâches précises avec un coût énergétique nettement réduit ;
- des modèles multimodaux, qui permettent par exemple d'analyser une image en entrée, sont testés lorsque leur licence l'autorise (Llava, gemma3, Mistral Small).

L'utilisation de LLM à poids ouverts est aussi un facteur décisif dans notre décision de déploiement d'une IA locale, qui reste conditionné par la régularité de publication des fournisseurs de LLM avec des licences

logicielles adaptées. Une telle approche nous permet de viser une certaine robustesse dans le déploiement de solutions d'IA locales, qui deviennent alors moins dépendantes d'un seul fournisseur monopolistique. Cette possibilité de choix nous permet de privilégier plusieurs LLM de moindre impact environnemental pour un usage donné. Ce choix possible correspond aussi à une certaine forme de souveraineté qui nous permet d'envisager une autre histoire que celle qui s'est répétée jusqu'à présent dans le domaine du numérique, par exemple sur les systèmes d'exploitation ou les suites bureautiques collaboratives : contrairement aux LLM fermés qui sont disponibles uniquement sur des plateformes propriétaires, à l'image des récents modèles d'OpenAI (via leur offre dédiée ou via une instance privée sur Microsoft Azure), les LLM à poids ouverts sont utilisables librement soit sur un poste de travail local pour les plus petits d'entre eux, soit sur un de nos serveurs pour les plus gros, dans une démarche de mutualisation à l'échelle de l'établissement.

Nous avons pu procéder à un tel déploiement grâce aux collègues chercheurs locaux (à l'IRISA²¹ et à l'INRIA²²) qui disposent des compétences adéquates pour ce faire, mais probablement de moins de temps que des

²¹ Institut de Recherche en Informatique et Systèmes Aléatoires implanté à Rennes.

²² Institut national de recherche en informatique et en automatique établi à Le Chesnay-Rocquencourt.

ingénieurs payés à temps plein dans des sociétés dont c'est l'activité première. Contrairement à de telles sociétés qui offrent des services au grand public, nous n'avons pas mis en place de mesures (technique ou organisationnelle) pour empêcher les contournements des garde-fous intégrés aux LLM (*model jailbreaking*). Comme nous privilégions des usages professionnels par des personnels, nous avons simplement averti les expérimentateurs que de tels systèmes de filtrage sont absents, ce qui permet d'ailleurs de mener des expérimentations intéressantes en recherche : dans un système grand public, il est difficile d'avoir des détails sur une procédure chirurgicale si elle est de nature à mettre en jeu la vie d'un patient, alors qu'il serait intéressant qu'un LLM biomédical dédié à la recherche puisse travailler sur de tels contenus. Actuellement, concernant RAGaRenn, un déploiement auprès de publics non avertis nécessiterait à minima une étude de sécurisation qui impliquerait probablement la mise en place de filtres entrants et sortants, et probablement avec une organisation dédiée en termes de suivis et de modération.

Concernant les services exploitant ces LLM, nous avons tenté au début trois approches en parallèle autour de la question d'un *chatbot* spécialisé sur nos données : développer de zéro une solution (en recourant à un prestataire spécialisé) ; utiliser une solution

commerciale d'une start-up ; et déployer un ensemble de briques *open-source*. Dans notre contexte, l'approche *open-source* s'est avérée de loin la plus fructueuse, car nous avons bénéficié d'avancées technologiques rapides d'une communauté mondiale, dont le rythme et la diversité sont nettement plus élevés que ce que nous avons pu développer par nous-mêmes ou en recourant à une offre dédiée d'une start-up dont l'adaptation a nécessité de nombreux allers-retours. Le temps investi a porté essentiellement sur les adaptations utiles à notre contexte (sécurisation et gestion de déploiements à la demande) et sur l'accompagnement des usages.

Les usages des expérimentateurs ont d'abord fait l'objet d'une analyse statistique en étudiant exclusivement les connexions sur RAGaRenn et le volume de données comptabilisé par les moteurs d'inférence, sans jamais accéder aux contenus. Sur plusieurs mois, le taux d'utilisation de la plateforme se situe à une moyenne de 1,8 utilisation par semaine, qui masque la disparité entre utilisateurs qui s'étale de plusieurs utilisations par jour à moins d'une fois par mois. Ces statistiques sont établies à la fin avril 2025, où la plateforme RAGaRenn compte plus de 800 utilisateurs, dont 50 % d'utilisateurs extérieurs à l'Université de Rennes, qui représentent 94 établissements tiers.

La satisfaction, en lien avec la vitesse perçue et la qualité des réponses obtenues (sur des documents ou non) a été évaluée auprès des expérimentateurs par deux enquêtes anonymes à 3 et à 8 mois après l'ouverture du service. Nous avons relevé un taux de satisfaction entre 70 et 80 %. Ces valeurs sont similaires à ce que nous constatons sur d'autres services numériques proposés par l'établissement. La baisse de satisfaction pourrait provenir de la difficulté récente de RAGaRenn à analyser des pages web externes, car certains sites ont récemment durci leur politique d'accès à des contenus par des robots. Volontairement nous n'avons pas mis en place de mesure technique de contournement afin de respecter la volonté des sites tiers : certains utilisateurs voient maintenant échouer certaines extractions de contenus de sites alors que ce n'était pas le cas début 2024. Ce point fait écho à des débats d'actualité sur les notions de propriété intellectuelle et d'usage raisonnable de contenus publics à l'ère de l'IA²³.

Les coûts financiers associés à l'utilisation de LLM ont aussi fait l'objet d'une première étude en juin 2024, dont le périmètre est volontairement restreint : nous avons étudié le coût de l'inférence seule attribuée aux

moments de fonctionnement d'un GPU pour une tâche d'inférence. Le coût complet de cette inférence est relié à la consommation des fluides (électricité, climatisation recourant à un circuit d'eau dédié), à l'amortissement du matériel (malheureusement rapide) et aux ressources humaines nécessaires pour la gestion de ce matériel. Pour des LLM identiques à vitesse égale, nous sommes parvenus à un coût au token de notre infrastructure similaire, voire inférieur, comparé à d'autres fournisseurs "grand public" en juin 2024.

Il est important de préciser que d'autres coûts ne sont pas intégrés. Par exemple, la mise en place de l'expérimentation a nécessité un investissement humain qui n'a pas été évalué précisément et qui n'a donc pas été répercuté ici : or, tout service à vocation commerciale intègre dans le coût à l'utilisateur une part de financement de sa partie Recherche & Développement. Autre exemple : nous avons constaté durant les 3 premiers mois un taux d'utilisation moyen d'environ 2 %, ce qui signifie que, pour une période de 24 h la plateforme est réellement utilisée 28 min. et 48s. Le reste du temps, la consommation en veille, qui est loin d'être négligeable, ne sert à rien ni personne si ce n'est réchauffer la planète... Usuellement pour un service commercial, même si le taux d'utilisation

²³<https://www.forbes.com/sites/virginieberger/2025/03/15/the-ai-copyright-battle-why-openai-and-google-are-pushing-for-fair-use/>

devrait être nettement supérieur en raison du plus grand nombre d'utilisateurs : le coût pour le client final intègre de telles pertes dues à l'inactivité des serveurs.

Dans un souci de sobriété numérique, nous avons néanmoins cherché à optimiser ce taux d'utilisation en lien avec les usages réels, en étudiant les plages horaires où la solution était sollicitée : les statistiques obtenues nous ont conduits à éteindre les serveurs de 22 h à 6 h, plage pendant laquelle ils étaient très faiblement utilisés, voire utilisés exclusivement par les auteurs du présent article qui se sont organisés autrement.

Au vu de l'impact environnemental potentiel, malgré une forte mutualisation de nos usages sur nos propres serveurs, nous nous sommes rapprochés d'autres établissements qui partageaient des préoccupations similaires au printemps 2024 : en particulier les collègues de l'école d'ingénieurs CentraleSupélec Paris-Saclay ont publié à cette période un composant *open-source* Aristote-dispatcher²⁴ qui agit comme une “gare de triage” pour distribuer des requêtes d'inférence LLM entre serveurs. La finalité de ce composant est d'améliorer le taux d'utilisation des serveurs, tout en répartissant la charge liée à un pic de sollicitations sur plusieurs

serveurs. Nous avons échangé ensemble, et convenu de lancer et coordonner une démarche de mutualisation nationale au travers du projet ILaaS²⁵ porté par les universités de Lille, Reims Champagne-Ardenne, Paris 1 Panthéon-Sorbonne, Rennes (coordinateur) et la grande école CentraleSupélec Paris-Saclay. L'objectif est de créer une fédération ESR mutualisant des serveurs d'IAg ouvertes (contrairement aux approches propriétaires) en s'appuyant sur la fédération de datacenters régionaux, voire à terme sur des mésocentres et/ou des centres de calcul nationaux. La mutualisation vise une meilleure sobriété numérique en mixant les usages (formation, recherche, administration) pour maximiser l'utilisation des moyens de calcul existants. Le projet ILaaS s'inscrit dans une démarche de résilience, de confiance partagée, et de soutenabilité sur plusieurs plans (budgétaire, environnemental, éthique et usages) : pour un établissement utilisant déjà des serveurs existants pour proposer de tels services, rejoindre la fédération offre de nouvelles possibilités sans pénaliser ses propres usages locaux que la “gare de triage” Aristote-Dispatcher va servir prioritairement, en utilisant les moments de disponibilité pour des requêtes émanant d'autres établissements.

²⁴ <https://github.com/CentraleSupélec/aristote-dispatcher>

²⁵ <https://www.ilaas.fr/>

3.2. L'accompagnement des collègues

Lors d'une transformation des pratiques, l'accompagnement des collègues est toujours une action de toute première importance. L'arrivée des IA_g dans nos postures professionnelles est évidemment une transformation importante. Elle a, qui plus est par rapport à d'autres transformations (par exemple, l'approche par compétences en pédagogie ou encore la protection des données personnelles dans les processus administratifs), un caractère extrêmement rapide. Cela change en partie les méthodes d'accompagnement mises en œuvre dans les services d'appui au sein de nos institutions. Il est souvent question de monter en compétences en même temps que les équipes accompagnées.

Dans notre expérimentation, nous avons cherché à accompagner les collègues impliqués dans tous les métiers de l'université : la gouvernance, le pilotage, l'administration, la recherche et la pédagogie. Dans tous les domaines de métiers, il ressort une grande curiosité des collègues pour les outils d'IA_g. Ils comprennent bien qu'une nouvelle transformation des pratiques en appui sur le numérique est en train de se jouer. La majeure partie des collègues y voit une opportunité pour passer moins de temps sur

des tâches à faible plus-value faisant en sorte de pouvoir se dégager du temps. Ce temps à retrouver est une attente pour le réinvestir dans des actions de cœur de métiers : prendre le temps de la réflexion, mieux s'investir dans l'accompagnement des étudiants, par exemple.

Nous avons organisé entre avril et décembre 2024 plus d'une vingtaine d'ateliers d'échanges (environ 260 collègues réunis en tout et en moyenne 12 par séance). Ces ateliers étaient souvent des temps d'appropriation des outils et d'acculturation aux notions clés. Depuis janvier 2025, nous nous sommes rapprochés du Conseil national du numérique (CNNUM)²⁶ et nos ateliers sont depuis reconnus comme des actions labellisées « Café IA » au sens du CNUM²⁶. Cela contribue à donner une visibilité plus grande et une reconnaissance de la démarche par un tiers de confiance national extérieur à notre établissement. C'est un aspect particulièrement important à nos yeux que de positionner une méthode de management participatif d'une institution publique locale en alignement avec des préconisations de l'État français.

Nous avons décliné des ateliers dans les 3 grandes familles de métiers : Administration, Recherche et Pédagogie. Les ateliers ont été des temps de partage de

²⁶ <https://cnnumerique.fr/annonce/cafe-ia-pour-un-dispositif-collectif-dappropriation-de-la-technologie>

cas d'usages, mais aussi de partage de doutes, de questionnements, de ressentis de risques et d'opportunités. Les ateliers ont montré beaucoup plus de questionnements et d'intérêts que de réticences.

Nous avons établi et mené 3 types d'ateliers afin de recueillir et exploiter efficacement les retours dans une perspective d'analyse des risques et opportunités :

Le premier type d'atelier est dédié à l'acculturation et au recueil de cas d'usages. Il est organisé en deux temps : la première partie est consacrée à la prise en main de RAGaRenn et aux échanges concernant l'outil en termes de possibilités techniques, mais aussi plus généralement sur les différents concepts liés à l'IAg et aux conséquences associées, par exemple sur la fiabilité des contenus produits ou la date de coupure de l'entraînement d'un LLM qui définit sa limite temporelle en termes de "connaissances" embarquées dans le LLM. La deuxième partie est dédiée à un échange en petits groupes sur leurs différents cas d'usages avérés ou potentiels, qu'ils doivent décrire dans un canevas pré-établi sur papier avec des notes autocollantes ou sur Klaxoon (Figure 4).

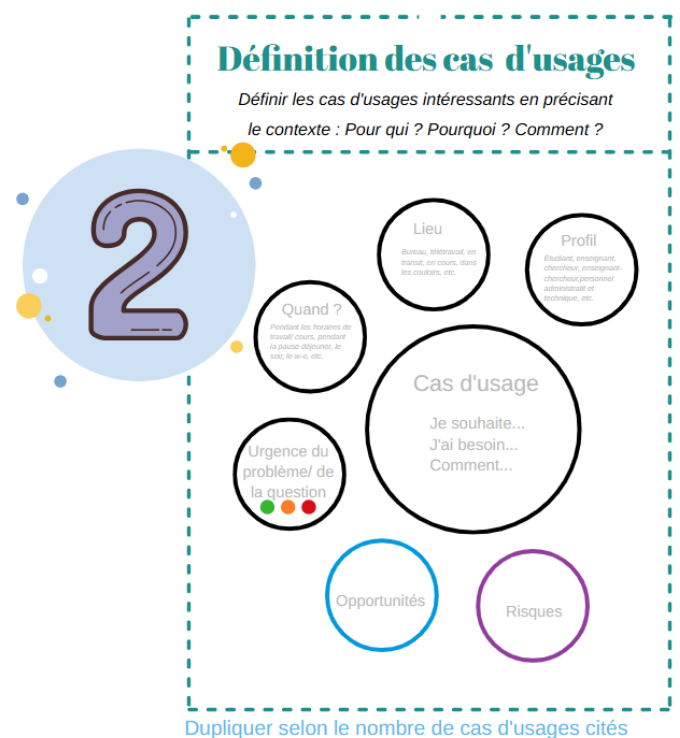


Figure 4. : Canevas permettant de décrire un cas d'usage

Le deuxième type d'atelier vise à étudier plus précisément les risques et opportunités, et fait suite au précédent : l'objectif est de passer en revue les différents cas d'usages produits au sein des groupes, et d'en dégager l'ensemble des risques et opportunités associés, en procédant à des regroupements si besoin. Par exemple, les risques de "fausses informations produites" ou de "dépendance à l'outil" se retrouvent dans plusieurs cas d'usages, mais sont reportés une seule fois dans la synthèse globale, qui est consolidée à l'échelle de l'ensemble des ateliers du premier type.

Le troisième type d'atelier vise à prioriser les risques et opportunités identifiées, tout en se projetant sur les mesures à mettre en place pour mitiger les risques tout en étant

capables de saisir les opportunités. Pour faciliter les échanges autour de cette priorisation avec des participants de profils variés sur un temps restreint, nous avons choisi d'utiliser des matrices de positionnement qui ont pour avantage de placer les sujets relativement les uns aux autres, et visualiser les évolutions possibles.

Dans toutes les actions d'accompagnement sur les outils d'IAg, nous avons souhaité partir de là où en sont les collègues et les écouter sans jugements. Sur les questions techniques, les différences de maîtrise des uns et des autres sont très importantes ainsi que les préjugés. Le meilleur cadre pour ces discussions s'est avéré être le partage des cas d'usages et des pratiques.

Dans les ateliers d'échanges que nous avons animés, la plus grande partie des collègues réunis avaient en général des idées de cas d'usages qu'ils aimeraient mettre en œuvre, mais pas souvent des véritables partages d'expériences avec des retours objectivés. C'est normal, la place des IAg dans nos métiers n'est pas encore trouvée. Le processus est en cours et demande un temps pour réfléchir à ce qu'on veut faire, comment on peut le faire et avec quelle plus-value.

Nous avons organisé des ateliers thématiques sur les métiers de la Recherche, de l'Administration et de la Pédagogie.

Plusieurs idées de cas d'usages y ont été débattues :

Dans les activités de recherche, les principaux cas d'usages discutés ont été l'assistance à la production écrite (correction/amélioration/reformulation de textes, proposition de plans, tâches rédactionnelles de peu d'intérêt), l'aide à la créativité (stimuler l'idéation, questionner un sujet, explorer une question) et l'aide à l'analyse de vastes corpus documentaires et de grandes quantités de données.

Dans les tâches administratives, les attentes en matière d'usages ont évoqué la création de robots conversationnels (accompagnement des usagers, réponses de premier niveau, réponses en horaires décalés, réponses personnalisées complémentaires aux FAQ), la création simplifiée de codes informatiques pour automatiser des tâches, l'annotation de contenus pour en faciliter l'accessibilité, l'aide à la synthèse de documents (par ex., de nature technique ou juridique) et l'assistance au pilotage de la formation.

Dans le domaine de la pédagogie, les principales questions évoquées concernaient l'évaluation des étudiants (évaluations formatives ou sommatives) avec des modalités variées (études de cas, quizz, jeux de mots), la conception des évaluations dans des contextes où les étudiants peuvent utiliser des IA, la conception d'outils

favorisant l'ancrage mémoriel (notamment avec la production de *flashcards*²⁷), les robots conversationnels dans des espaces de cours en ligne et les outils-auteur pour l'enseignant ou l'ingénieur de pédagogie pour accompagner la production de syllabus, de scénarisation pédagogique, de contenus pour aller plus loin.

Aujourd'hui nous avons quelques exemples de questionnement des IA dans les postures professionnelles qui vont plus loin que ces premières idées. Des travaux sont menés et questionnent la plus-value des IA dans les différentes catégories de métiers, comme le montrent les exemples suivants :

En Recherche : Nous accompagnons le travail de recherche d'une doctorante de l'Université de Caen qui intègre des modèles d'IA dans des approches hybrides Sémantique-LLM pour le peuplement automatique d'ontologies à partir de textes d'annonces (Sahbi *et al.*, 2025).

Pour le pilotage de la formation : nous travaillons avec les responsables des formations de notre établissement à préparer l'évaluation et l'accréditation des formations par l'État²⁸. Pour cela, il faut faire une

analyse de la compatibilité des maquettes de formation avec les référentiels attendus²⁹. Nous utilisons RAGaRenn pour produire automatiquement des premières mises en correspondance dans le but de faciliter le travail des responsables de formation. Par ailleurs nous collaborons avec des collègues de l'Université Gustave Eiffel en région parisienne pour l'utilisation de l'IA dans la préparation semi-automatisée de notes sectorielles (notamment à partir de données en *open data* sur le marché de l'emploi et les référentiels de compétences métiers) pour éclairer le positionnement et l'attractivité de formations.

En matière de pédagogie : C'est le domaine où les cas d'usages concrets sont les plus nombreux. Certains enseignants, par exemple, recommandent aux étudiants de s'aider de l'IA dans la rédaction de leurs travaux sous réserve qu'ils le mentionnent et indiquent en annexe les prompts qu'ils ont réalisés et ce qu'ils en ont pensé (c'est une façon de questionner l'esprit critique). D'autres collègues, par exemple en Informatique, intègrent la génération de code dans leur pédagogie pour questionner les étudiants sur leurs compétences de futurs

²⁷ Pour plus de détails sur l'usage des *flashcards* pour la mémorisation, cf. conférence en ligne de Erik Gustafsson, Université de Franche-Comté : <https://tube-numerique-educatif.apps.education.fr/w/qusbpGBKQWB3seYyWhpbC>

²⁸ Le Haut Conseil à l'Évaluation de la Recherche et de l'Enseignement Supérieur (HCERES) et la

Direction Générale de l'Enseignement Supérieur et de l'Insertion Professionnelle (DGESIP).

²⁹ Le Registre National de la Certification Professionnelle (RNCP) et le GreenComp (https://joint-research-centre.ec.europa.eu/greengcomp-european-sustainability-competence-framework_en)

développeurs. À l'Institut Universitaire de Technologie de Saint-Brieuc, les étudiants de la licence professionnelle CPA (Commercialisation des Produits Alimentaires) s'initient à l'art du prompt professionnel et testent l'infrastructure RAGaRenn lors d'un projet tuteuré dans le cadre de leur alternance avec des entreprises du secteur alimentaire. Enfin, à titre d'application plus techniquement pointue, nous réalisons actuellement un projet de simulation d'entretiens de recrutement (avec notre service de formation continue et d'apprentissage). L'étudiant, qui joue le rôle de candidat, discute avec un avatar virtuel, qui joue le rôle de recruteur. Les énoncés de l'avatar sont produits via une couche applicative spécifique à la tâche qui interroge notre infrastructure LLM pour la formulation des questions et des réponses.

3.3. L'objectif d'un cadrage institutionnel

Un objectif majeur de l'expérimentation RAGaRenn est de produire un cadrage institutionnel sur le déploiement et les usages des outils d'IAg à l'échelle de notre établissement. Notre démarche est de considérer que ce cadrage doit venir d'une contribution collective d'un maximum de nos collègues et que ces collègues doivent expérimenter dans un environnement de confiance pour s'approprier des outils, s'acculturer à des techniques, explorer les possibilités.

Nous n'ignorons pas que la mise en place d'un environnement de confiance se heurte à certaines formes de réticences institutionnelles, plus ou moins conscientes. Certains collègues demandent par exemple à utiliser des outils dont ils entendent parler juste pour ne pas se sentir à côté des usages grand public (notamment ceux des étudiants). D'autres se disent que si l'université pousse l'usage de ses outils locaux, c'est pour mieux contrôler ces usages. En tant que fournisseurs d'accès de services numériques nous avons effectivement le devoir de pouvoir remonter des journaux de connexions et des données d'utilisateurs en cas de commission rogatoire. Mais cela ne veut pas dire qu'on le fait dans le fonctionnement normal de l'établissement et les utilisateurs ont d'ailleurs la possibilité de ne pas conserver dans leur historique les prompts et résultats obtenus. La mise en place d'un cadrage institutionnel est pour nous un enjeu de sécurisation et de confiance avec nos collègues.

Dans l'organisation réglementaire de l'établissement, ce cadrage prendra la forme d'une charte annexée à celle qui encadre les usages du numérique. Celle-ci prévoyant déjà des points tels que le respect des règles existantes concernant les droits de propriété intellectuelle (liés à l'utilisation de contenus tiers ou non), le traitement des données à caractère personnel, le devoir de réserve, de

discrétion, de neutralité et le secret professionnel ou encore l'information pour s'assurer de la compréhension et du recueil du consentement des utilisateurs.

Nous avons discuté des démarches existantes comparables à la nôtre (par exemple, les démarches de l'Université de Liège³⁰, de l'Illinois State University³¹, du Cégep du Vieux Montréal³², de l'Université Laval³³, de l'Université d'Orléans³⁴) et des complémentarités. Nous avons contribué à un travail de consolidation en lien avec d'autres établissements au sein de groupes de travail du réseau inter-DemoES. Nous avons formulé des principes directeurs sous forme synthétique (devoirs & interdictions), qui sont publiés sous la forme d'une proposition de charte et de guides sur le site du projet DemoES Pro3 : <https://www.demoes-pro3.fr/proposition-charte-ia-esr/>

Au titre des "Devoirs", nous avons par exemple identifié l'utilisation de l'IAg avec discernement et parcimonie pour des cas d'usages à bilan d'impacts positif, l'utilisation des outils d'IAg validés par l'établissement pour des besoins professionnels pertinents, la priorisation de l'« *open-source* », le contrôle par l'humain

des productions de l'IAg. En tant qu'étudiant, dans le cadre d'une évaluation, lorsque l'usage est autorisé, mentionner l'utilisation de l'IAg dans les passages concernés. En tant qu'enseignant, communiquer explicitement les consignes d'usages de l'IAg auprès des étudiants durant les activités pédagogiques, notamment celles donnant lieu à évaluation.

À titre d'interdiction, nous avons ciblé la non-utilisation en substitution d'un geste enseignant, sans supervision ni vérification, pour des activités telles que la correction des copies, la formulation de retours personnalisés (*feedbacks*) aux étudiants. Nous avons aussi alerté à ne pas mettre implicitement les étudiants dans des situations d'addiction numérique à des outils qui privilégient l'immédiateté de la réponse à son bien-fondé.

Nous souhaitons que notre charte repose sur quatre principes fondamentaux : la curiosité, la transparence, la précaution et la parcimonie. La curiosité encourage l'exploration et l'appropriation ainsi que la montée en compétences des collègues au moyen de l'expérience d'usage. De ce point de vue, une charte ne se veut pas seulement limitative, mais cherche à maintenir une

³⁰ https://www.student.uliege.be/cms/c_19230399/

³¹ <https://prodev.illinoisstate.edu/pedagogy/ai/>

³² <https://www.cvm.qc.ca/wp-content/uploads/Charte-sur-l-IA-mai-2023-vf.pdf>

³³ <https://www.enseigner.ulaval.ca/numerique/intelligence-artificielle-generative/principes-directeurs>

³⁴ <https://www.univ-orleans.fr/fr/univ/universite/reglements/documents-reglementaires>

cohérence entre curiosité et sécurité. La transparence garantit que les processus et les décisions prises par les outils d'IAg sont autant que possible compréhensibles et accessibles à tous. Le principe de précaution assure que les risques sont anticipés et gérés de manière proactive. Enfin, la parcimonie, en lien avec la volonté de sobriété numérique, vise à utiliser les ressources de manière judicieuse et responsable, en minimisant particulièrement l'impact environnemental préoccupant de l'IAg.

Dans le contexte anxiogène sur la disparition des métiers avec l'IA que nous connaissons, nous profitons de l'écriture collective de la Charte sur l'IA pour affirmer quelques valeurs. Par exemple, le non-remplacement de l'humain par la machine. L'IA doit être pensée pour soutenir les activités humaines, c'est un rôle *d'adjuvant* au sens de Luc Julia (Julia, 2019) pour compléter, prolonger et démultiplier l'action humaine. Une métaphore explicative dans le domaine de la santé humaine est celle de la différence entre une orthèse qui soutient une partie du corps et une prothèse qui remplace une partie du corps. L'IAg que nous appelons de nos vœux doit agir en orthèse, non en prothèse (Lardellier et Carré, 2025). Cela a pour conséquence que l'utilisation d'une IAg doit se voir comme une assistance réversible

(comme on déciderait de se passer d'une orthèse). On doit se garantir à nous-mêmes de toujours pouvoir faire le choix de ne plus utiliser l'IAg et prévoir dans le pire des cas un mode dégradé des services métier sans assistance d'IAg.

La charte est en cours d'écriture et de finalisation. Nous souhaitons la présenter et la faire valider par les conseils décisionnels de notre établissement à l'été 2025. Une fois actée par les instances, elle sera rendue publique sur le site de l'établissement <https://www.univ-rennes.fr/>.

3.4. Partages de pratiques au-delà de l'établissement

Nous avons organisé, le 12 juin 2024, une journée de conférences avec des intervenants d'autres établissements particulièrement identifiés - autour de la thématique « Outils d'IAg : stratégie & expérimentation » (programme et captations vidéo disponibles en ligne³⁵). Cette journée était coorganisée avec la Direction régionale académique du numérique dans l'éducation (DRANE) et nos collègues de l'Académie de Rennes (Région Bretagne). Elle a regroupé plus de 150 personnes.

En juin 2025, nous avons organiser un colloque d'audience nationale sur les environnements physiques et numériques

³⁵ <https://suptice.univ-rennes.fr/evenements/outils-dia-generative-strategie-experimentation>

pour enseigner et apprendre et réfléchir sur les transformations pédagogiques au sein de l'enseignement supérieur. Parmi les thématiques débattues, nous avons abordé tout particulièrement les outils d'IA dans la formation universitaire, dans les technologies immersives, les pédagogies interactives et dans les environnements d'apprentissage capacitant.

L'activité d'essaimage autour de notre expérimentation en cours passe également par du travail de recherche de visibilité et de recherche de fonds. Dernièrement, nous avons contribué à une réponse française à l'appel à projets européen "AI Factory"³⁶ pour proposer le développement d'une action d'ampleur nationale sur l'IA et les technologies pour l'éducation. Notre proposition a été retenue dans la réponse française acceptée par la Commission européenne.

Dans le cadre du Sommet pour l'action sur l'intelligence artificielle³⁷, les projets France 2030 DemoES et EdTech France s'allient pour concrétiser un défi intitulé "L'IA au service de la performance éducative", avec le soutien du Secrétariat Général Pour l'Investissement (SGPI), de l'Agence Nationale de la Recherche (ANR), et de

Mistral AI. Ce défi valorise l'excellence de la recherche et de l'entrepreneuriat français ainsi que la force de la collaboration des écosystèmes pour mettre l'IA au service de l'enseignement. Nous participons à ce défi dans l'idée de proposer un RAG pour les établissements d'enseignement supérieur basé sur la solution RAGaRenn. Nous mettons en avant notre infrastructure qui se veut plus frugale, plus éthique, avec des règles algorithmiques plus appropriées aux usages d'enseignement et à la finalité d'intérêt général, dans le respect des libertés et l'ambition de diffusion des savoirs.

4. L'analyse collective des coûts, des risques et des bénéfices

Les participants des ateliers ont travaillé en petits groupes pour identifier des éléments d'opportunité et de risque. Nous leur avons demandé de remplir des matrices inspirées de la matrice d'Eisenhower (PMBOK, 2021) selon les dimensions "importance et difficulté" pour les opportunités et les dimensions "probabilité et gravité" pour l'évaluation des risques, dans une approche inspirée de la méthode AMDEC³⁸ (cf. Figure 5). La première est une matrice "Opportunités" construite sur les axes "Difficulté" (va de "faible" à "élevée") et

³⁶ <https://digital-strategy.ec.europa.eu/fr/news/eu-boosts-european-ai-developers-ai-factories-call-proposals>

³⁷ <https://www.elysee.fr/en/sommet-pour-l-action-sur-l-ia>

³⁸ https://fr.wikipedia.org/wiki/Analyse_des_modes_de_d%C3%A9faillance_de_leurs_effets_et_de_leur_criticit%C3%A9

“Importance” (va de "moins important" à "plus important"). La matrice est divisée en quatre quadrants, chacun représentant un type d'opportunité : Distraction (basse importance, faible difficulté), Grande Valeur

(basse importance, forte difficulté), Luxe (forte importance, faible difficulté), Stratégique (forte importance, forte difficulté).

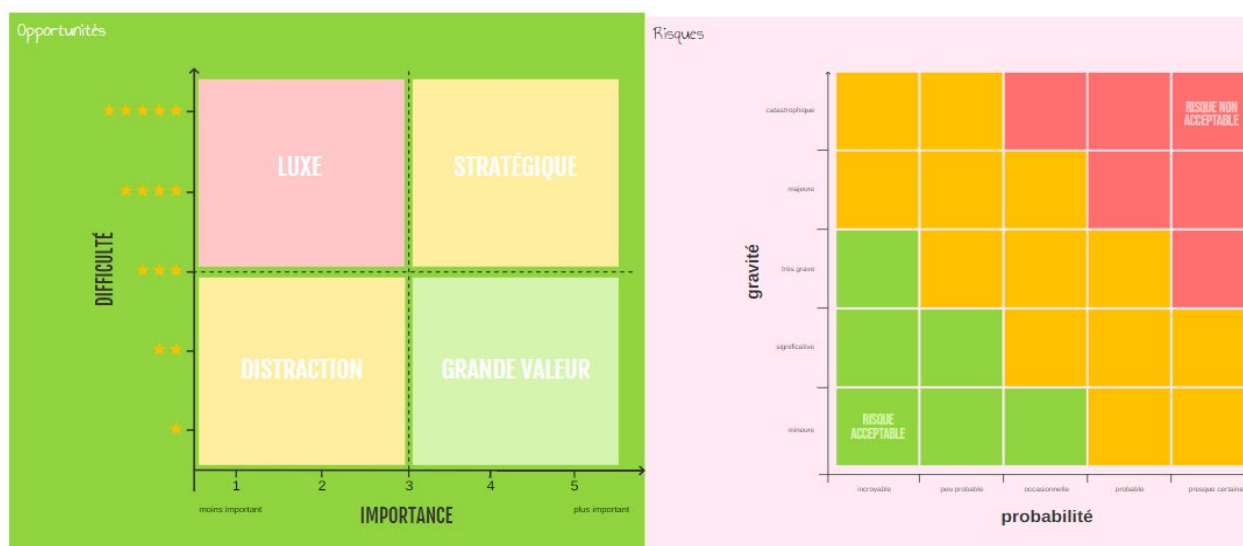


Figure 5 : matrices utilisées en ateliers

La seconde est une matrice “Risques” construite sur les axes “Gravité” (va de "catastrophique" à "négligeable") et “Probabilité” (va de "improbable" à "très probable"). La matrice est divisée en zones de risques : non acceptable, à surveiller, acceptable.

La Figure 6 donne un aperçu de la variabilité des opportunités et risques qui ont été discutés. Parmi ceux-ci, les opportunités qui reviennent souvent sont l’enrichissement des contenus, le gain de temps, la possibilité d’une assistance d’ingénierie pédagogique, le renouvellement de la pédagogie. Les

risques qui ressortent le plus sont : l’utilisation de l’IA pour un devoir/évaluation maison (comment identifier les cas de fraude/plagiat ?), la tendance des étudiants à utiliser sans retenue des services externes (comme ChatGPT, Gemini, Claude ou autres ...), l’impact environnemental du numérique, la diffusion de données à caractère personnel, la fuite d’informations stratégiques, la perte de souveraineté numérique, la responsabilité peu claire des contenus créés et enfin les fausses informations produites. De manière transversale aux catégories “Recherche”,

“Pédagogie” et “Administration” que nous avons incité à suivre dans notre méthodologie de recueil, ces risques peuvent être analysés en plusieurs méta-thèmes inter-catégories : la sécurité des données et des processus, les impacts sociaux et environnementaux, le rôle des universités dans la construction des connaissances,

l’accompagnement des parcours d’apprentissage.

À ce stade, ces méta-thèmes demandent encore à être stabilisés et pourraient être inspirants pour revisiter notre méthodologie de constitution de données empiriques.

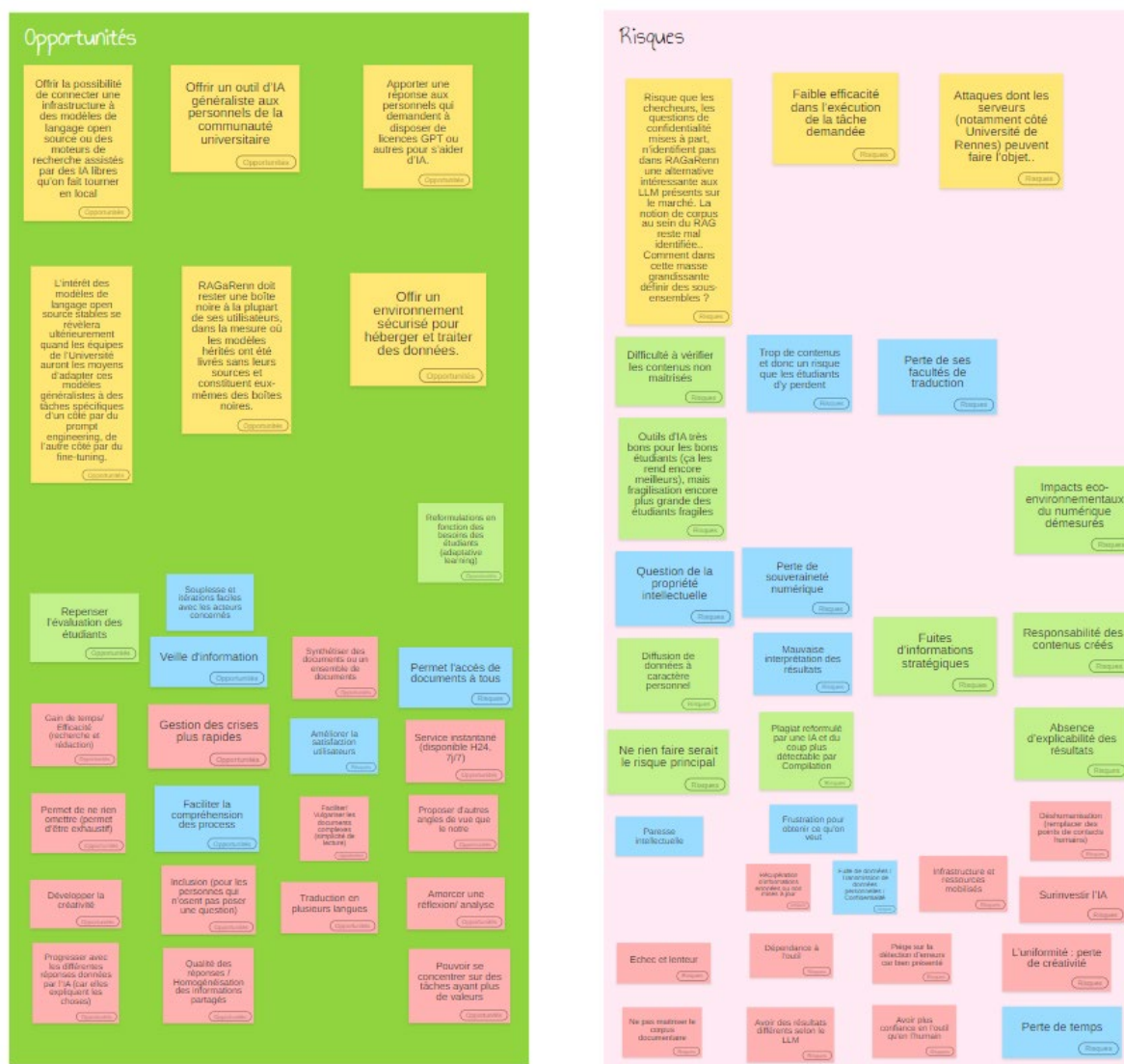


Figure 6 : Opportunités et risques, résultats issus des ateliers

Légende des étiquettes : jaune = Recherche ; vert = Pédagogie ; rose = Administratif

Quand des risques étaient jugés non acceptables, car très probables et très graves, la réflexion consistait à identifier des pistes de réduction du risque le faisant revenir dans une zone d'acceptabilité. Par exemple, le risque de fausse information donnée aux étudiants par hallucination de l'IAg est très probable (c'est inhérent au fonctionnement des LLMs) et risqué s'ils ne sont pas à même de percevoir l'hallucination. La méthode de réduction de risque consisterait à indiquer explicitement aux étudiants que les productions issues des IAg doivent bien être rediscutées avec l'enseignant.

Il ressort des échanges qui ont eu lieu lors des ateliers avec les collègues participants que nous avons tous à cœur de faire de notre université une organisation flexible qui aborde sereinement la conduite du changement. Beaucoup des discussions ont concerné les modifications des postures professionnelles. De ce point de vue, en mettant en avant notre adaptabilité et notre flexibilité dans nos modes de fonctionnement internes, nous nous apparentons clairement à une organisation de culture de type "Clan" au sens du modèle de Cameron et Quinn (2005). La culture de clan, qui se caractérise par un lieu de travail convivial où les gens partagent beaucoup d'eux-mêmes, est également en phase avec

notre vision de notre rôle en tant qu'animateur de la communauté. Nous ne cherchons pas à nous présenter comme des décideurs, mais plutôt à être perçus comme des mentors, des personnes de confiance et en même temps des porte-paroles des remontées de terrain. Nous attachons beaucoup d'importance à l'engagement collectif, au climat de travail, à l'attention portée aux collègues, à la construction de consensus.

5. Perspectives

Pour 2025, l'expérimentation continue avec la volonté d'inciter de nouveaux cas d'usages chez les collègues et les partager en ateliers.

Les fonctionnalités de la plateforme vont être augmentées avec l'acquisition de deux nouveaux services installés en couche au-dessus de l'infrastructure de LLM RAGaRenn :

- Le service Nolej³⁹ : outil auteur pour accompagner l'enseignant et/ou l'ingénieur de pédagogie à produire automatiquement des activités en ligne sous forme de modules H5P à partir des documents fournis par l'enseignant (textes, diaporama, vidéo ...) ;
- Le service K2R2⁴⁰ : couche applicative installée au-dessus de RAGaRenn pour

³⁹ <https://www.nolej.io/>

⁴⁰ <https://www.e-charlemagne.com/index.php/assistant-k2r2/>

produire des tuteurs virtuels dans des espaces de cours Moodle où l'interaction entre le tuteur et l'utilisateur étudiant est contrôlée pour rester dans la tâche prévue.

Un autre objectif de l'année à venir est de développer nos actions de formation sur l'appropriation des outils d'IAg. Nous mettons en place un espace en ligne complémentaire aux ateliers pour s'acculturer (avec plusieurs ressources en ligne pour l'autoformation, notamment une collection de capsules vidéo courtes intitulée *Parlons IA*⁴¹) et partager des cas d'usages. Nous poursuivons la formation des personnels d'accompagnement à la transformation pédagogique. De septembre à novembre derniers, une dizaine d'ingénieurs pédagogiques du SUPTICE (service d'appui à la pédagogie de l'Université de Rennes) ont suivi un parcours de formation. Cela leur était essentiel pour les sécuriser dans leur posture professionnelle, car les formations d'ingénierie pédagogique intègrent peu de connaissances et compétences sur les IA. Ces formations vont s'enrichir d'autres sessions avec la prolongation de l'expérimentation.

Nous déclinons aussi des actions de formation destinées aux enseignants, notamment dans le cadre des formations obligatoires à la pédagogie pour les nouveaux enseignants-chercheurs. Nous travaillons également à déployer des actions de formation professionnelle en dehors de l'université (à destination des collègues de l'Académie de Rennes (Région Bretagne) ou à destination des entreprises, par exemple la formation « Utilisation des IA génératives - LLM en entreprise »⁴²). L'idée est de développer sur mesures des Actions de Formation en Situation de Travail (AFEST⁴³) dédiées aux outils d'IAg pour lesquels nous apportons notre expertise.

6. Conclusions

Notre expérimentation est encore en cours. Beaucoup de choses ont déjà été faites et d'autres restent encore à venir, notamment sur l'analyse des transformations, le cadrage local, la mutualisation, la formation ... Nous percevons que les retours de nos collègues sont très positifs sur la démarche engagée.

Nous communiquons le plus possible pour faire partager et essaimer notre expertise collective : nous avons dernièrement présenté RAGaRenn dans des instances très variées, telles qu'un réseau d'entreprises

⁴¹ https://video.univ-rennes1.fr/channels/#parlons-ia_wonrjan8t6

⁴² <https://intelligence-artificielle.univ-rennes.fr/formation-continue-IA>

⁴³ <https://www.anact.fr/quest-ce-quune-action-de-formation-en-situation-de-travail-afest>

dans les technologies éducatives, l'assemblée générale des directions de systèmes informatiques d'universités, le conseil économique, social et environnemental de Bretagne ... Nous avons aussi plusieurs collègues d'autres établissements universitaires qui nous demandent de partager notre expérience : Université de Caen Normandie, Université de Bordeaux, Université de la Réunion, Université Gustave Eiffel, par exemple.

Clairement, il nous semble que le risque principal aujourd'hui sur l'impact des outils d'IAg dans l'enseignement supérieur et la recherche publique serait de miser sur une posture attentiste. La mise en action permet de se donner les moyens d'une transformation choisie de nos pratiques professionnelles, et non subie en réponse à des pressions exogènes (étudiants, société, etc.). Notre démarche permet d'aborder de manière éclairée et transversale plusieurs aspects : éthiques, techniques, culturels, réglementaires, institutionnels, politiques et environnementaux. La dynamique collective que nous avons mise en place nous apparaît essentielle pour réellement se donner un environnement de confiance. Même une infrastructure technique locale (telle que nous la défendons plutôt que de recourir à des outils extérieurs dont on serait captifs) pourrait encore être perçue avec une certaine méfiance vis-à-vis de l'institution par certains collègues si elle n'était pas

accompagnée par une dynamique collective (c'est la culture de clan au sens de Cameron et Quinn), un partage de la culture métier.

Sur le plan de nos pratiques professionnelles, nous identifions plusieurs questions complémentaires auxquelles notre démarche cherche à apporter quelques premières réponses : Comment former à l'IA nos étudiants ? Comment former nos étudiants avec l'IA ? Comment transformer nos métiers et nos institutions avec l'IA ? Comment mettre en avant nos universités comme des structures résilientes aux changements de la société ? Comment faire de nos établissements des laboratoires du changement au service de la société ?

Notice biographique

Pierre BEUST est professeur d'informatique à l'Université de Rennes. Il est également directeur du SUPTICE, service d'appui à la pédagogie de l'Université de Rennes. Il enseigne à l'ISTIC (UFR d'Informatique et Électronique) et effectue ses recherches dans l'équipe SHADOC du laboratoire IRISA. Il travaille sur les environnements numériques pour la pédagogie (LMS, learning analytics, IA pour enseigner et apprendre, par exemple). Il participe à l'expérimentation RAGaRenn : expérimentation de grande ampleur sur les usages des outils d'intelligence artificielle à l'échelle de l'Université de Rennes dans tous les métiers : gouvernance, pilotage,

administration, recherche et, évidemment, pédagogie.

Olivier WONG-HEE-KAM est vice-président en charge du numérique pour l'Université de Rennes et porteur du projet AIR, lauréat de l'AMI Démonstrateur numérique dans l'Enseignement Supérieur (DemoES). En tant que président de l'association VP-NUM qui regroupe les vice-présidents numériques de l'enseignement supérieur, il s'investit sur les enjeux liés à la mutualisation, à la

cybersécurité, et à l'IA. Il coordonne et contribue à différentes expérimentations en IA générative, sur Rennes (RAGaRenn) et à l'échelle nationale (ILaaS, POC FVE Amue/DGESIP), afin d'alimenter la vision stratégique sur l'IA générative par des actions et des réflexions sur les usages, leurs impacts et leur cadrage, tout en répondant aux enjeux de mutualisation sur les services numériques et les infrastructures de calcul.

Bibliographie

ALCANTARA, C., CHAREST, F. et LAVIGNE, A. (2020). IA et communication des organisations. *Communication & management*, vol. 17, n°2, 5-8.

AYENI, O. O., AL HAMAD, N. M., CHISOM, O. N., OSAWARU, B. and ADEWUSI, O. E. (2024). AI in education: A review of personalized learning and educational technology. *GSC Advanced Research and Reviews*, 18(2), 261-271.

BRYMAN, A. (2012). *Social Research Methods*. Oxford University Press.

CAMERON, K. S. et QUINN, R. E. (2005). *Diagnosing and changing organizational culture: Based on the competing values framework*. John Wiley & Sons. <https://www.researchgate.net/profile/Tahani-Aldahdouh-2/post/Is-the-any-test-or-questionnaire-to-measure-the-organizational-climate-anyone-can-suggest/attachment/59d61f45c49f478072e97966/AS%3A271750183489537%401441801700739/download/Diagnosing+and+changing+organizational+culture+Based+on+the+competing+values+framework.pdf>

CHESTERMAN, S. (2023). Good Models Borrow, Great Models Steal: Intellectual Property Rights and Generative AI (SSRN Scholarly Paper 4590006). <https://doi.org/10.2139/ssrn.4590006>

CREVIER, S., SAMUEL, F. S. L. and TISSERAND, M. (2023). *Compétence numérique des étudiantes et étudiants: expérimentations et perceptions*. Dans Actes du 42e Colloque de l'AQPC : "Le collégial en effervescence", les 7, 8 et 9 juin 2023. Association québécoise de pédagogie collégiale (AQPC).

CURRIE, G. M. (2023, September). Academic integrity and artificial intelligence: is ChatGPT hype, hero or heresy? In *Seminars in Nuclear Medicine*, 53(5), 719-730. WB Saunders.

DAVIS, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 1989, 13, 319-340.

FISHBEIN, M. et AJZEN, I. (1975). *Belief, Attitude, Intention and Behavior : An Introduction to Theory and Research*. Reading, MA : Addison-Wesley

FRISCH, M. (2015, 9 octobre). Didactique de l'Information-Documentation, des modèles, des concepts et un prototype de matrice curriculaire dynamique. Canal-U. <https://doi.org/10.60527/88cw-p695>

GANASCIA, J.G (2017). *IA : Vers une domination programmée ?* Le Cavalier Bleu.

GLASER, B.G. et STRAUSS, A.L. (1967). *The discovery of grounded theory: strategies for qualitative research*. Aldine.

HUSSAIN, A. (2023). *Use of artificial intelligence in the library services : prospects and challenges*. *Library Hi Tech News*, 40(2), 15-17.

JULIA, L. (2019). L'intelligence artificielle n'existe pas. First.

KHOWAJA, S. A., KHUWAJA, P., DEV, K., WANG, W. and NKENYEREYE, L. (2024). Chatgpt needs spade (sustainability, privacy, digital divide, and ethics) evaluation : A review. *Cognitive Computation*, 1-23.

KOOLI, C. (2023). Chatbots in education and research : A critical examination of ethical implications and solutions. *Sustainability*, 15 (7), 5614. <https://doi.org/10.3390/su15075614>

LARDELLIER, P., CARRE, E. (2025). *L'IA dans l'enseignement supérieur : les leçons de Platon et du mythe de Theuth*. <https://theconversation.com/lia-dans-lenseignement-superieur-les-lecons-de-platon-et-du-mythe-de-theuth-244894>

LEE (Hank), H.-P., SARKAR, A., TANKELEVITCH, L., DROSOS, I., RINTEL, S., BANKS, R., and WILSON. N. (2025). The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects from a Survey of Knowledge Workers. In *CHI Conference on Human Factors in Computing Systems* (CHI '25), April 26-May 01, 2025, Yokohama, Japan. ACM, New York, NY, USA, 23 pages.

https://www.microsoft.com/en-us/research/wp-content/uploads/2025/01/lee_2025_ai_critical_thinking_survey.pdf

LI, K. (2023). Determinants of college students' actual use of AI-based systems : An extension of the technology acceptance model. *Sustainability*, 15(6), 5221.

LOBET, M., HONET, A., WATHELET, V. (2023). *ChatGPT, l'usage par les étudiants de première année à l'université*. <https://shs.hal.science/halshs-04190226v1/document>

MARTY, S., VASQUEZ, K. (2024). Intégration des IAs génératives à l'université. *Revue COSSI*, (13). https://doi.org/10.34745/numerev_1961

PAIVANDI, S. (coord.) (2016). *Articulation et complémentarité des équipes pédagogiques et des services de documentation au cœur de la transformation pédagogique*. Rapport à la DGESIP réalisé par le LISEC, 86 p. <https://urlz.fr/q0ql>

PETIT, L. (2021). Les SHS et les SIC au défi de l'IA. *Communication, technologies et développement*, 10|2021. <https://doi.org/10.4000/ctd.4475>

PLANTARD, P. (2022), Collectif DESIR (sous la direction de P. Plantard), *Transformations pédagogique et numérique dans l'enseignement supérieur*, Presses de Mines, <https://www.pressesdesmines.com/produit/transformation-pedagogiques-et-numeriques-dans-lenseignement-superieur/>

PMBOK (2021). *Project Management Body of Knowledge*, 7ème édition du guide PMBOK, <https://www.theprojectgroup.com/blog/fr/7eme-edition-guide-pmbok/>

RUSSELL GROUP (2023). *Russell Group principles on the use of generative AI tools in education*. Cambridge, Russell Group.

SAHBI A.N.E., ALEC C., BEUST P. (2025). Semantic vs. LLM-based approach: A case study of KOnPoTe vs. Claude for ontology population from French advertisements, *Data & Knowledge Engineering (DKE)*, Vol. 156, 102392, ISSN 0169-023X

<https://doi.org/10.1016/j.datak.2024.102392>

UNESCO (2024), *Orientations pour l'intelligence artificielle générative dans l'éducation et la recherche*, <https://doi.org/10.54675/HBCX3851>

UZAN, G., POULAIN, P., REBY, F. et TALY, A. (2024). *Évaluation de l'utilisation d'un robot conversationnel comme assistant en accessibilité pour la lecture des articles scientifiques*. Dans *Conférence Handicap 2024 – 13ème édition Recherches pluridisciplinaires pour l'autonomie des personnes en situation de handicap*, Jun 2024, Paris, France. fihal-04419718v2f

WEINBERGER, E., JANIZEK, J. and LEE, S. I. (2020). Learning deep attribution priors based on prior knowledge. *Advances in Neural Information Processing Systems*, 33, 14034-14045.

WU, T., HE, S., LIU, J., SUN, S., LIU, K. and HAN, Q.-L. (2023). A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development. *IEEE/CAA Journal of Automatica Sinica*, 10(5), 1122-1136. <https://doi.org/10.1109/JAS.2023.123618>