

Déployer une IA de confiance

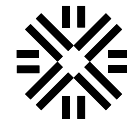
1^{er} Octobre 2025

Journée de la Donnée territoriale

Lycée agricole "Le gros Chêne" Pontivy

Pierre Beust – Pr. Université de Rennes

pierre.beust@univ-rennes.fr



**Université
de Rennes**

Plan

Les injonctions contradictoires
Un environnement de confiance
Conclusion

Plan

Les injonctions contradictoires

Un environnement de confiance

Conclusion

Les injonctions contradictoires

Beaucoup trop d'exemples ...

- Le temps
 - ❖ « Le train va passer vite, il faut monter dedans » mais « sans dire comment et à la place de quoi... »
- La formation
 - ❖ « Formez vous aux IA génératives ... » mais « débrouillez vous par vous-mêmes ... »
- L'outillage
 - ❖ « Utilisez les IA génératives ... » mais « précisément sans vous en dire plus ... »
- L'individu et le collectif
 - ❖ « Ca va changer tous vos métiers » mais « au final ça relève de chacun de nous ... »
- La transformation
 - ❖ « transformez vous ... » mais « continuez à faire vos tâches quotidiennes ... »

Les injonctions contradictoires

Quelles réponses ?

- Donner le temps de la réflexion, de l'acculturation, des échanges ...
- Chacun doit pouvoir s'y intéresser au moment qui est le sien
- Organiser des temps collectifs réguliers
- Prescrire des solutions techniques et en accompagner les usages
- Documenter la démarche (ressources d'auto-formation, guides d'usages)
- Reconnaître l'investissement des personnels
- Pousser l'idée d'une transformation « par petits pas » en amélioration continue
- Avoir une position d'établissement claire et lisible (ex. charte d'usage)

Plan

Les injonctions contradictoires

Un environnement de confiance

Conclusion

Un environnement de confiance

Les GAFAM ne sont pas un environnement de confiance

- Modèles économiques bi-faces : « Quand c'est gratuit, c'est vous le produit ! »
 - ❖ Des utilisateurs qui ne payent pas le service rendu
 - ❖ Des annonceurs qui payent pour un autre service que celui rendu
- Double usage des données des utilisateurs
 - ❖ Monétisation
 - ❖ Exploitation propre dans les phases de réentraînement du modèle
- Certains peuvent faire payer le service en contrepartie de la non exploitation des données
 - ❖ Cela peut être une solution (c'est mieux que rien)
 - ❖ Attention à la maîtrise de la trajectoire tarifaire

Un environnement de confiance

l'environnement technique

- Garder le contrôle sur les données
- Avoir un service le plus local possible pour limiter les communications
 - ❖ [Argument de sécurité et argument d'impact environnemental](#)
- Privilégier les solutions libres
- 3 possibilités :
 - ❖ [Logiciels d'inférence sur sa machine : llmstudio.ai, jan.ai](#)
 - ❖ [Serveurs d'inférence administrés localement](#)
 - ❖ [Services en ligne avec avec contrat de confidentialité](#)

Un environnement de confiance

l'environnement institutionnel

- Une volonté politique
- Une démarche expérimentale
- Prendre le temps qu'il faut
- S'inscrire dans des actions de mutualisation
- Viser le « juste nécessaire », pas plus ...

Un environnement de confiance

RAGaRenn : La « boîte à outils » de l'université de Rennes

- <https://ragarenn.eskemm-numerique.fr/>
- Demande de compte utilisateur (U. Rennes et extérieurs)
- FAQ : <https://projet-air.univ-rennes.fr/faq-ragarenn>
- Les « Cafés IA »
- La charte des usages : <https://personnel.univ-rennes.fr/charte-portant-sur-lutilisation-des-intelligences-artificielles-generatives-ia>
- Espace PédagogIA : <https://formation-permanente.univ-rennes.fr/course/view.php?id=96>
- Capsules vidéos :
 - ❖ Parlons IA (5-10 min) : https://video.univ-rennes1.fr/channels/#parlons-ia_wonrjan8t6
 - ❖ Causeries (45 min) : https://www.canal-u.tv/recherche?search_api_fulltext=ia+et+enseignement+sup%C3%A9rieur&op=Envoyer&f%5B0%5D=chaine%3A409

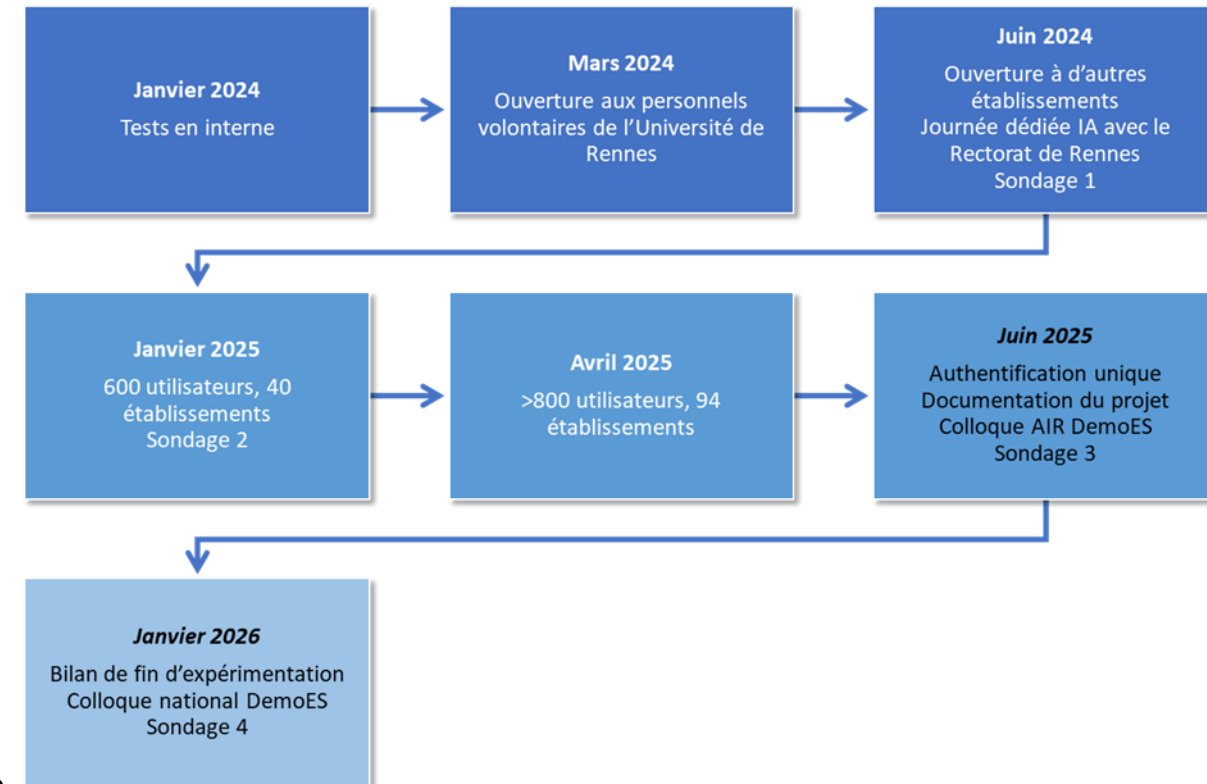
Un environnement de confiance

Le déploiement de RAGaRenn

- Circulaire du Président lançant l'expérimentation fin mars 2024
- Portage par le VP Numérique (Olivier Wong-Hee-Kam) + assistance du SUPTICE.

Un objectif :

- Faire des enseignants des exemples pour les étudiants de bonnes intégrations d'IAG dans la posture professionnelle



Un environnement de confiance

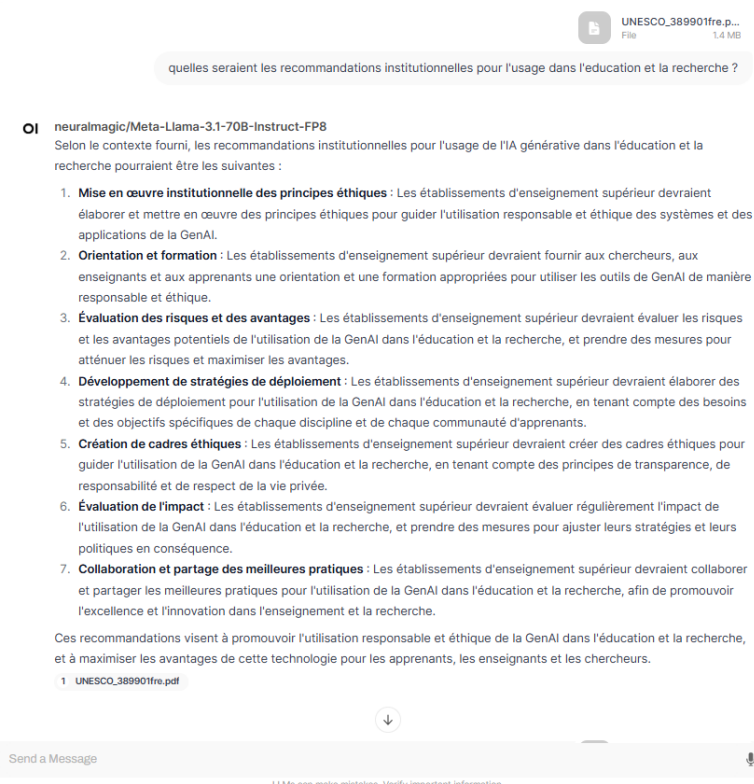
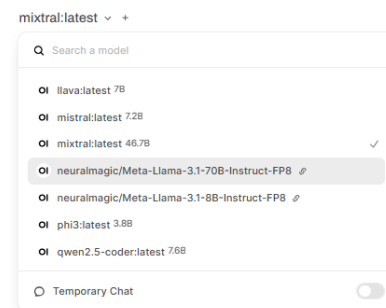
La dimension technique

- L'infrastructure RAGaRenn (différents LLM, RAG)

- ❖ Hébergée au datacenter Eskemm Data datacenter breton de l'ESR labellisé faisant partie de la fédération de datacenters régionaux construite par le MESR
- ❖ Cartes GPU NVidia : 2 H100 et 4 V100
- ❖ Deux types de serveurs d'inférence : Ollama (<https://ollama.com/>) vLLM (<https://github.com/vllm-project/vllm>).
- ❖ Chat et RAG : Open WebUI (<https://openwebui.com/>)
- ❖ LLM téléchargés via la plateforme Hugging Face <https://huggingface.co/>

- Usages de la plateforme

- ❖ Plus de 1000 collègues impliqués ...
- ❖ Des usages réguliers (1,8/outil./semaine)



Un environnement de confiance

La dimension technique : le Retrieval Augmented Génération (RAG)

- À ne pas confondre avec ...
 - ❖ L'entraînement des modèles
 - ❖ Le fine-tuning
- Une amélioration de la génération
 - ❖ Moins d'hallucinations, plus grande qualité de génération
 - ❖ Enrichissement du contexte
 - ❖ Sourcer la raison d'un passage généré



resume moi les propositions de ce document

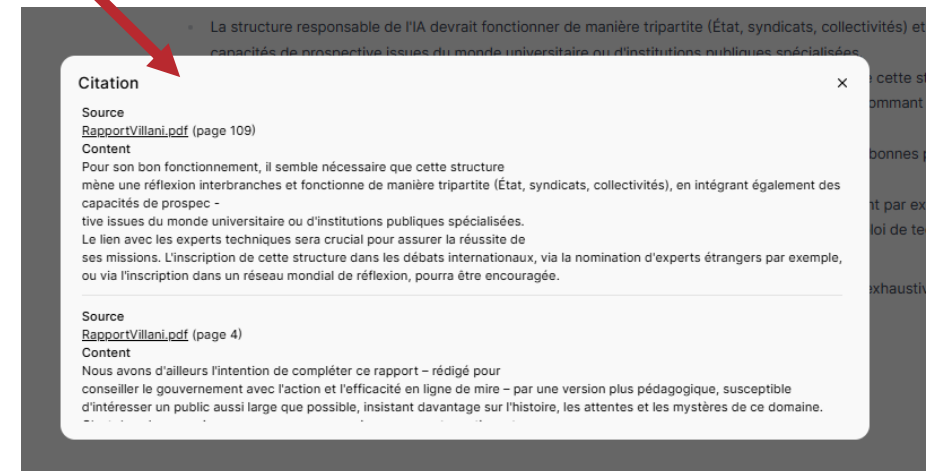
OI neuralmagic/Meta-Llama-3.1-70B-Instruct-FP8 3:41 PM

D'après le contexte fourni, les propositions de ce document semblent se concentrer sur l'éthique de l'intelligence artificielle (IA) et les réflexions à mener pour son développement responsable. Voici quelques-unes des propositions mentionnées :

- La structure responsable de l'IA devrait fonctionner de manière tripartite (État, syndicats, collectivités) et intégrer des capacités de prospective issues du monde universitaire ou d'institutions publiques spécialisées.
- Il est crucial d'établir un lien avec les experts techniques pour assurer la réussite des missions de cette structure.
- L'inscription de cette structure dans les débats internationaux est encouragée, par exemple en nommant des experts étrangers ou en s'intégrant dans un réseau mondial de réflexion.
- La France devrait jouer un rôle de premier plan dans la définition de réglementations ou de guides de bonnes pratiques au niveau international pour les systèmes d'armes développés, acquis ou adoptés.
- La France pourrait étudier des pistes qui s'intéressent à des solutions technologiques déterminant par exemple le niveau d'interaction humain-machine nécessaire dans le développement, le déploiement et l'emploi de technologies émergentes.

Il est important de noter que ces propositions sont issues du contexte fourni et peuvent ne pas être exhaustives ou détaillées.

1 RapportVillani.pdf



Citation

Source
RapportVillani.pdf (page 109)
Contenu
Pour son bon fonctionnement, il semble nécessaire que cette structure mène une réflexion interbranches et fonctionne de manière tripartite (État, syndicats, collectivités), en intégrant également des capacités de prospective issues du monde universitaire ou d'institutions publiques spécialisées. Le lien avec les experts techniques sera crucial pour assurer la réussite de ses missions. L'inscription de cette structure dans les débats internationaux, via la nomination d'experts étrangers par exemple, ou via l'inscription dans un réseau mondial de réflexion, pourra être encouragée.

Source
RapportVillani.pdf (page 4)
Contenu
Nous avons d'ailleurs l'intention de compléter ce rapport – rédigé pour conseiller le gouvernement avec l'action et l'efficacité en ligne de mire – par une version plus pédagogique, susceptible d'intéresser un public aussi large que possible, insistant davantage sur l'histoire, les attentes et les mystères de ce domaine.

Un environnement de confiance

La dimension collégiale

- Des ateliers d'acculturation, de partage d'usage, de réflexions (risques, bénéfices, vigilances ...)
 - ❖ Environ 2 ateliers proposés par mois
 - ❖ Beaucoup plus de questionnements que de réticences
 - ❖ labellisées « Café IA » au sens du CNUM (<https://cnnumerique.fr/annonce/cafe-ia-pour-un-dispositif-collectif-dappropriation-de-la-technologie>)



Un environnement de confiance

L'encapacitation

- C'est donner davantage de pouvoir à des individus ou à des groupes pour agir sur les conditions sociales, économiques, politiques ou écologiques auxquelles ils sont confrontés.
- On emploie aussi régulièrement l'anglicisme *empowerment* ou la notion d'autonomisation
- Plus de détails ...
 - ❖ <https://fr.wikipedia.org/wiki/Empowerment>
 - ❖ Jean-Gabriel Ganascia, 2017 : *IA : Vers une domination programmée ?* (Le Cavalier Bleu ed.)

Un environnement de confiance

Quels déterminants de l'encapacitation ?

- La stimulation
 - ❖ Adhérer à un besoin de transformation
- La connaissance et compréhension
 - ❖ En matière d'IA, de technologie, les acquis individuels peuvent être très hétérogènes
- L'autonomie
 - ❖ Être capable d'utiliser seul (mais pas sans les autres ...)
 - ❖ Valoriser une démarche empirique
- Les ressources et supports
 - ❖ Documenter, indexer et diffuser des ressources ...
- Le mouvement collectif
 - ❖ Un groupe c'est beaucoup plus qu'une somme d'individus

Un environnement de confiance

Cadrage des usages

- Construction collaborative
 - ❖ À l'échelle de l'établissement
 - ❖ À l'échelle nationale : action dans le cadre du défi "L'IA au service de la performance éducative" à l'IA Summit (fev. 25) par les projets France 2030 DemoES et EdTech France
3-4 pages adaptables, compatible réglementations et chartes informatique locales
<https://www.demoes-pro3.fr/proposition-charte-ia-esr/>
- Des établissements déjà organisés :
 - ❖ U. Orléans : https://www.univ-orleans.fr/upload/public/2024-06/dircom_2024_charte_IA_UO_R.pdf
 - ❖ U. de Liège
 - ❖ U. Laval
 - ❖ ...

Un environnement de confiance

- Cadrer ce n'est pas limiter les usages ... c'est les sécuriser !
- 4 principes fondamentaux :
 - ❖ la curiosité
 - ❖ la transparence
 - ❖ la précaution
 - ❖ la parcimonie
- Fixer des limites : obligations et interdictions
- Voté fin juin 2025 par différentes instances (CAC - CSA - CA) et a revue par les référents déontologue et intégrité scientifique

Plan

Les injonctions contradictoires

Un environnement de confiance

Conclusion

Conclusion

Quelques principes à rappeler

- On ne prompte pas une IA si ...
 - ❖ on arrive très bien à faire sans ...
 - ❖ on a pas bien compris la question/le travail à faire ...
 - ❖ on ne réfléchit pas ...
 - ❖ on veut faire faire à sa place ...
 - ❖ on ne veut pas rendre clair qu'on a utilisé une IA dans son travail ...
 - ❖ on veut faire « comme tout le monde » ...
 - ❖ ...
- Surtout, on ne prompte pas une IA si ...
 - ❖ **on est pas capable d'évaluer les résultats ...**

Conclusion

Discuter des valeurs ajoutées mais aussi des risques

- Risques de fuites de données stratégiques
 - ❖ « J'ai demandé à ChatGPT de rédiger la conclusion de mon projet de recherche ... »
- Risques de délégation de compétences non contrôlée, de perte de compétences
 - ❖ Cf. étude 2025 du MIT
Kosmyna, N., Hauptmann, E., Tong Yuan, Y., Situ, J., Liao, X.-H., Beresnitzky, A.-V., Braunstein, I., Maes P. (2025). Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task,
<https://doi.org/10.48550/arXiv.2506.08872>
- Risques en matière de fracture numérique
 - ❖ Des bons outils pour savent s'en servir, une fragilité supplémentaire pour les plus fragiles
- Risques environnementaux
 - ❖ Importance d'une prise de conscience pour le « juste nécessaire »
- Risques socio-politiques
 - ❖ Attention à la question de la souveraineté numérique (quelle place pour les européens dans l'IA ?)

Conclusion

Quelles places pour les IA dans nos postures professionnelles ?

- Les outils ne doivent pas être pensés pour faire les choses à la place des gens.
- Ne croyons pas trop à l'idée qu'on va « gagner du temps » ...
- L'objectif, c'est de faire que les gens eux-mêmes fassent
 - ❖ Différemment (si besoin),
 - ❖ Mieux (si possible),
 - ❖ comme ils le veulent (de toute façon) ...

Conclusion

Les questions au centre des évolutions à venir ...

- La sobriété des usages
- La confiance en la technologie
- Le contrôle (qui fait quoi et qui assume une responsabilité)
- La place de l'humain

La balle est aussi dans notre camp

- Si l'on fait nous même de notre travail quelque chose de remplaçable, il ne faudra pas s'étonner d'être remplacé ?
- Ce n'est pas la dernière évolution technologique (il y en aura d'autres ...)
- Mais ce n'est pas qu'une évolution technologique ...

