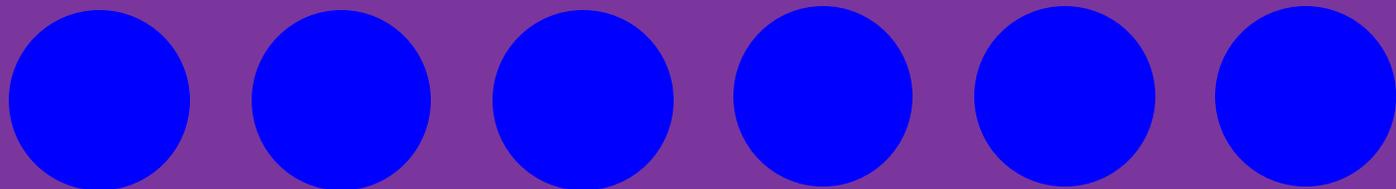


Accompagnement
des transformations

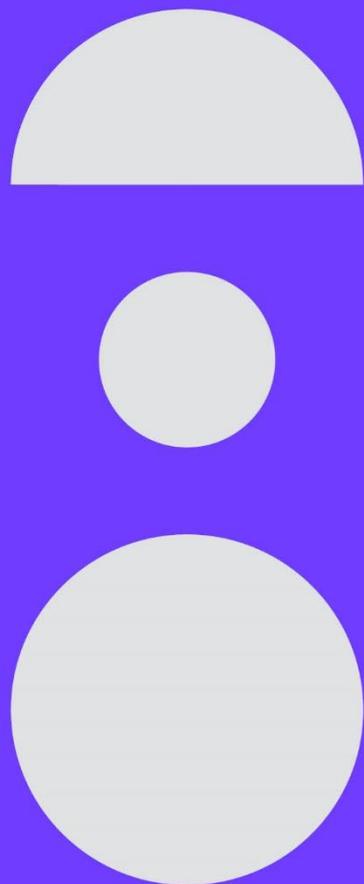


Transition numérique : focus sur l'IAG dans l'ESR

Cycle de webconférences - Retours d'expérience



Mars à juillet 2025



Webconférence #2

« Présentation d'outils internes d'Intelligence Artificielle générative »

Retours d'expérience

Centrale Supélec – Outil Aristote

Université de Rennes - Outil RagaRenn



CentraleSupélec





CentraleSupélec

CentraleSupélec

Intervenant :

Renaud Monnet, DSI-N



1. La genèse du projet : printemps 2023

Avril 2023 : 4 étudiants sont missionnés sur le thème "IA Générative et Education"

- ils choisissent la problématique : "comment aider l'étudiant à réviser ses cours et préparer ses examens avec l'IA"
- ils sont encadrés par le DSI Adjoint, en charge du Développement des Usages Numériques pour l'Enseignement

Juin 2023 : les 4 étudiants présentent à une conférence "IA et Education" le site qu'ils ont fabriqué

- l'étudiant dépose sur le site une vidéo (ex : le replay d'un cours)
- le site l'aide à réviser , en reformulant certaines notions et en proposant des quizz
- réactions : c'est important que l'enseignant puisse conserver la maîtrise sur le transcript de la vidéo et les quizz générés par l'IA

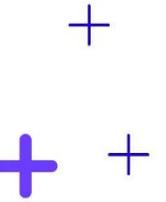
Juin 2023 : définition d'une posture "Souveraineté"

- des ateliers de partage de pratiques entre enseignants, animés par le pôle Innovation Pédagogique
- un amphi de vulgarisation sur le fonctionnement interne d'une IA Générative + les travaux menés par le laboratoire MICS pour des IA 100% Open Source
- une discussion en Comité de Direction sur les risques de fuite de données liés à Chat GPT
- un atelier personnels enseignants + non enseignants durant la Convention d'Encadrement Annuelle

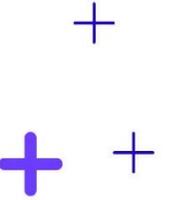
Problématique : comment utiliser l'IA Générative tout en garantissant le confinement des données ?

2. Le projet Aristote

- **Pour résoudre la problématique, l'approche utilisée est :**
 - ne pas construire de LLM : il en existe déjà beaucoup => utiliser des LLM Open Source existants
 - ne pas construire de nouveaux outils : les enseignants et étudiants sont déjà confrontés à de nombreux outils
 - ajouter de l'IA aux outils actuels, en proposant aux éditeurs de ces outils d'utiliser des API
- **Le projet consiste donc à construire des API :**
 - utilisables par les éditeurs d'outils éditeurs publics et privés)
 - dans une approche 100% Open Source le code est Open Source et les LLM sont Open Weight)
 - avec une exploitation sur des serveurs maîtrisés



3.L'organisation du projet



- La DSI-N est en charge de la recherche de financement :
 - 100K€ obtenus du FTAP (Fonds de Transformation de l'Action Publique)
 - et quelques financements complémentaires dans l'écosystème CentraleSupélec
 - et des discussions en cours avec les éditeurs de logiciels EdTech

- Les développeurs sont internes et externes :
 - les API sont développées en PHP Symfony
 - les couches basses ont fait l'objet d'itérations multiples impliquant des ingénieurs systèmes et des étudiants

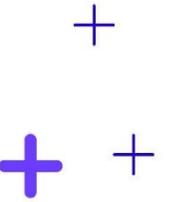
- Les retours produits par l'IA sont évalués par des enseignants et des étudiants :
 - un site Web permet de visualiser le résultat des API
 - toutes les vidéos stockées sur le serveur Ubcast de l'établissement ont été enrichies avec transcription + quizz

- Des enseignants ambassadeurs lancent régulièrement de nouvelles explorations (synthèse des mails reçus sur une boîte générique, analyse des données de scolarité)

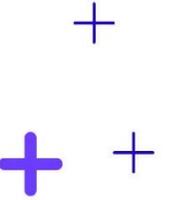


4. Les étapes franchies

- Outils officiellement raccordés à Aristote :
 - POD (depuis la version 3.7)
 - Whaller (outil collaboratif souverain)
 - Lutice (offre du BBB as a Service)
 - Ubcast plateforme vidéo)
 - Ideta (chabot)
- Discussions en cours avec plusieurs EdTEch



5.Travaux en cours : l'IA pour les Soft Skills



3 étudiants missionnés sur "l'IA pour aider les étudiants sur leurs Soft Skills"

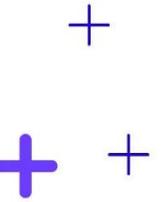
- il en ressort un nouveau Service Numérique en cours de test : Pitch Pilot
- l'étudiant se filme et dépose la vidéo dans Pitch Pilot
- Pitch Pilot lui donne des conseils, qui tiennent compte des objectifs qu'il s'est fixé et des consignes enregistrées par les enseignants

Ce travail amène des évolutions sur l'API Aristote :

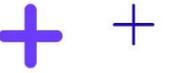
- analyse de l'image pour caractériser la posture (épaules, regard, ...)
- analyse du son pour identifier la dynamique
- analyse du fond du discours pour identifier sa structuration logique et la conformité aux objectifs

6.Travaux en cours : génération de contenus

- **3 étudiants missionnés sur le thème "Analyser arXiv et produire automatiquement un article sur un sujet donné"**
- **Ils ont utilisé Storm, un code Open Source publié par Stanford, et l'ont adapté pour des IA Open Source**
- **Ce travail amène des évolutions sur l'API Aristote :**
 - recherche de contenus sur Internet arXiv mais aussi Wikipedia
 - gestion d'agents qui échangent entre eux pour rédiger l'article
 - extension du principe à la génération de contenus de cours dans Moodle



7.Travaux en cours : l'analyse des copies d'examen ⁺



A l'occasion des examens, les copies sont le plus souvent scannées

=> Ce qui permet d'envisager l'utilisation d'une IA

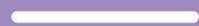
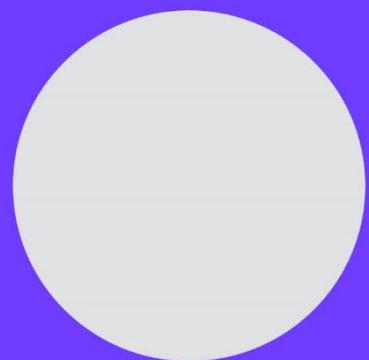
Cette IA doit être de confiance, les données ne doivent pas sortir.

L'objectif n'est pas de fixer une note, mais de fournir des feedbacks et des conseils.

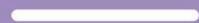
Université de Rennes

Intervenant :

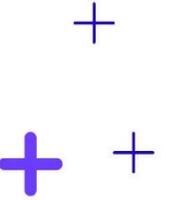
Olivier Wong-Hee-Kam, vice-président numérique



La genèse du projet



Expérimentation IA générative



Fin 2023 : interrogations & demandes fortes sur les outils d'IA générative – tests & veille

Mars 2024 : lancement d'une expérimentation sur les outils d'IA générative auprès des personnels, sous l'impulsion de David ALIS, Président de l'Université de Rennes.

Portage : vice-président numérique et directeurs de services (DSI, Suptice, DAJI)

Mars – juin 2024 : phase pilote

Jusqu'à fév. 2026 : phase consolidation

Expérimenter pour comprendre, apprendre et décider

- Ateliers participatifs autour des besoins et usages : acculturation sécurisée
- Déploiement de **RAGaRenn** : infrastructure de confiance, avec des briques open source
- Evaluer, mesurer, partager pour pouvoir décider : qualitatif (ateliers) et quantitatif (sondages, mesures)
- Ecriture d'un cadrage pour l'établissement : contribution à la proposition de charte DemoES



Le projet RAGaRenn



Ce projet a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du Programme d'Investissements d'Avenir portant la référence ANR-21-DMES-0001



RAG : utiliser des sources documentaires connues, validées pour améliorer les réponses



• Sans source

You
Comment utiliser ma carte Korrigo ?

OI Mixtral:Latest 27/03/2024 23:07

La carte Korrigo est une carte de transport utilisée dans certaines régions de France, notamment en Bretagne. Pour utiliser votre carte Korrigo, vous devrez suivre ces étapes :

1. Charger votre carte : Avant de pouvoir utiliser votre carte Korrigo, vous devez la charger avec des titres de transport valides. Vous pouvez le faire en ligne sur le site web de Korrigo, dans une borne de rechargement automatique ou dans un point de vente Korrigo.
2. Valider votre carte : Lorsque vous montez à bord d'un bus ou d'un train, vous devez valider votre carte en la passant sur le lecteur de cartes situé à proximité des portes d'entrée. La validation permet

• Avec source

You

 <https://docinfo.univ-...>
Document

Comment utiliser ma carte Korrigo ?

OI Mixtral:Latest

Pour utiliser votre carte Korrigo, vous devez l'associer à votre compte informatique Sésame sur un copieur Toshiba en passant votre carte et en saisissant vos identifiants et mot de passe Sésame. Si vous changez de carte, il faudra renouveler cette opération. Sans carte, vous pouvez toujours imprimer en saisissant vos identifiants Sésame à chaque utilisation d'un copieur.

La carte Korrigo vous permet également d'accéder à différents services universitaires tels que le service de photocopie et d'impression, l'inscription au SIUAPS et l'inscription aux BU. Lors de sa première création, la carte Korrigo est offerte, mais en cas de perte, vol ou détérioration, son renouvellement sera payant.

Périmètre du projet RAGaRenn

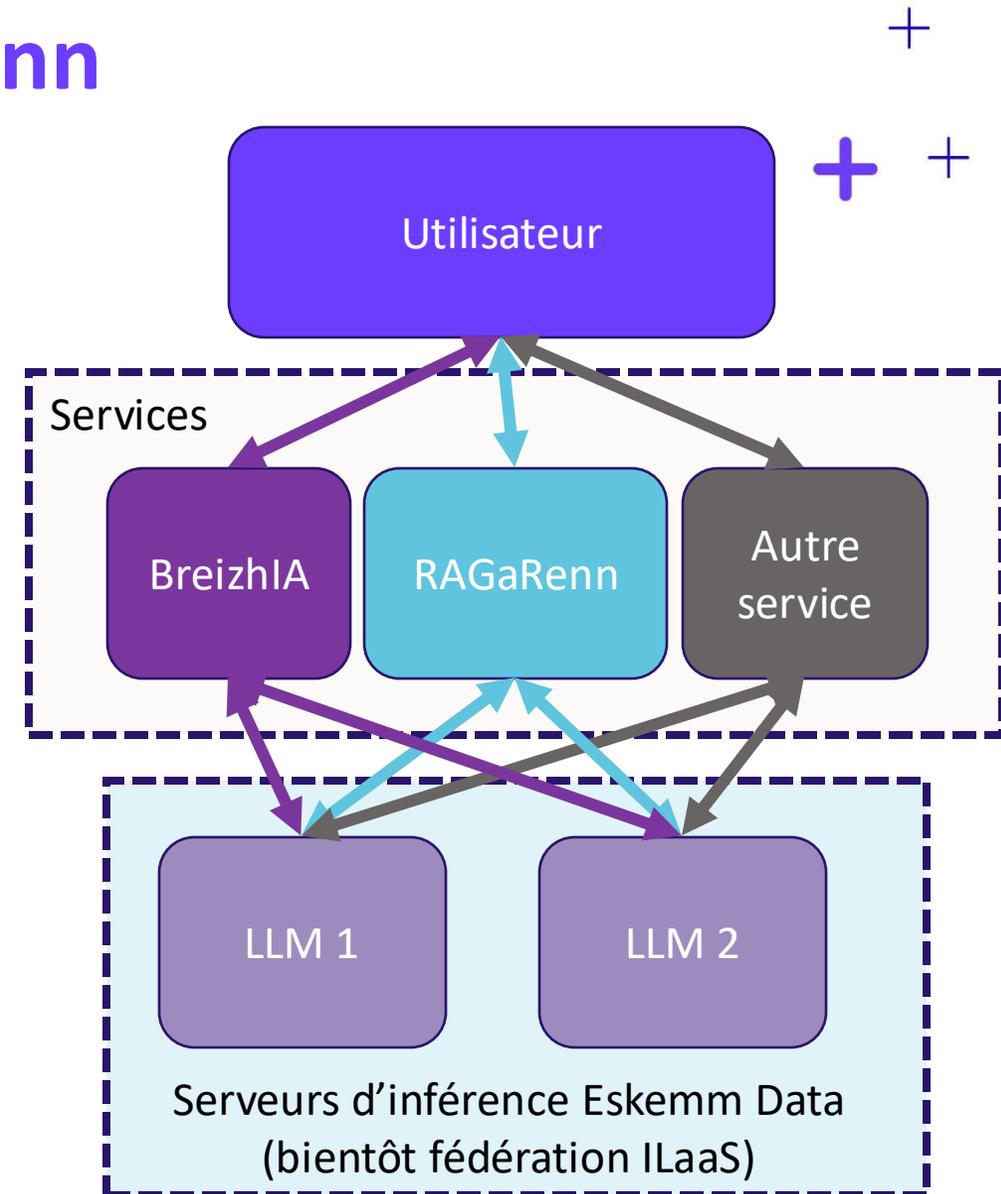
Principe : séparer l'accès à un LLM et l'accès à des services utilisant un LLM

Approche agnostique

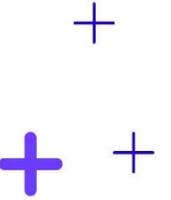
- Choix du/des LLM(s) utilisé(s) en open weights
- Compatible avec différentes infrastructures (DC labellisés, Méso, SecNumCloud, etc.)

Maitrise des flux de données, des impacts environnementaux & budgétaires

- Stockage et flux de données sécurisés
- Facturations indépendantes (services // LLM) et maîtrisées
- Mesure des usages réels & impacts associés



Organisation du projet



- **Pilotage** : vice-président numérique
- **Financement** : AIR DemoES, avec accord des tutelles pour réorienter une partie du budget
- **Infrastructure** : Eskemm Data, datacenter régional ESR labellisé par le MESR (COSIN)
- **Déploiement & développements** spécifiques : équipe de recherche Logica (IRISA)
François Bodin, Laurent Morin, Corentin Val, Benjamin De Zordo
- **Gestion** des utilisateurs : AIR DemoES
- **Cas d'usages & communications** spécifiques
 - administration : AIR DemoES
 - pédagogie : Suptice
 - recherche : SCD
- **Etudes** liées au passage en production : DGS, DAJI, DSI, Suptice
(authentification, sécurisation, aspects juridiques, budgétaires)

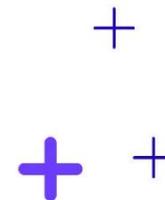




Le déploiement du projet



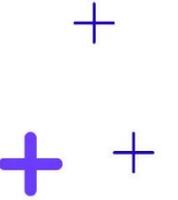
Le calendrier projet, les jalons



- Janvier 2024** Tests en interne
- Mars 2024 Ouverture aux personnels volontaires de l'Université de Rennes : 1 instance
- Juin 2024 280 utilisateurs
Ouverture à d'autres établissements
(instances mutualisées ou à la demande, accès API)
[Journée dédiée IA](#) (avec le Rectorat de Rennes) – Sondage 1
- Janvier 2025** 600 utilisateurs, 40 établissements – Sondage 2
- Avril 2025 >800 utilisateurs, 94 établissements
- Juin 2025 [Authentification unique \(fédération d'identité ESR\)](#)
[Ouverture de la documentation du projet – Sondage 3](#)
[Colloque projet AIR DemoES 12-13 juin \(Rennes\)](#) →
- Janvier 2026** [Bilan fin de l'expérimentation – Sondage 4](#)
[Colloque national DemoES](#)



Dispositif de déploiement



Echanges collectifs

- Acculturation, partage autour de cas d'usages : réfléchir aux opportunités et aux risques, partager réussites, doutes et échecs. Beaucoup plus de questionnements que de réticences
- S'appropriier les outils par la pratique, apprendre et tâtonner
- Apprendre, comprendre, expliquer le fonctionnement et la technologie
- Partager veille & prospective ; se coordonner (comité stratégique IA)

Chiffres & évolutions

- avril à décembre 2024 : 20+ sessions, 260 participants
- 2025 : 10+ sessions avant l'été (2-3 par mois) avec la labellisation Café IA au sens du CNUM

Communications

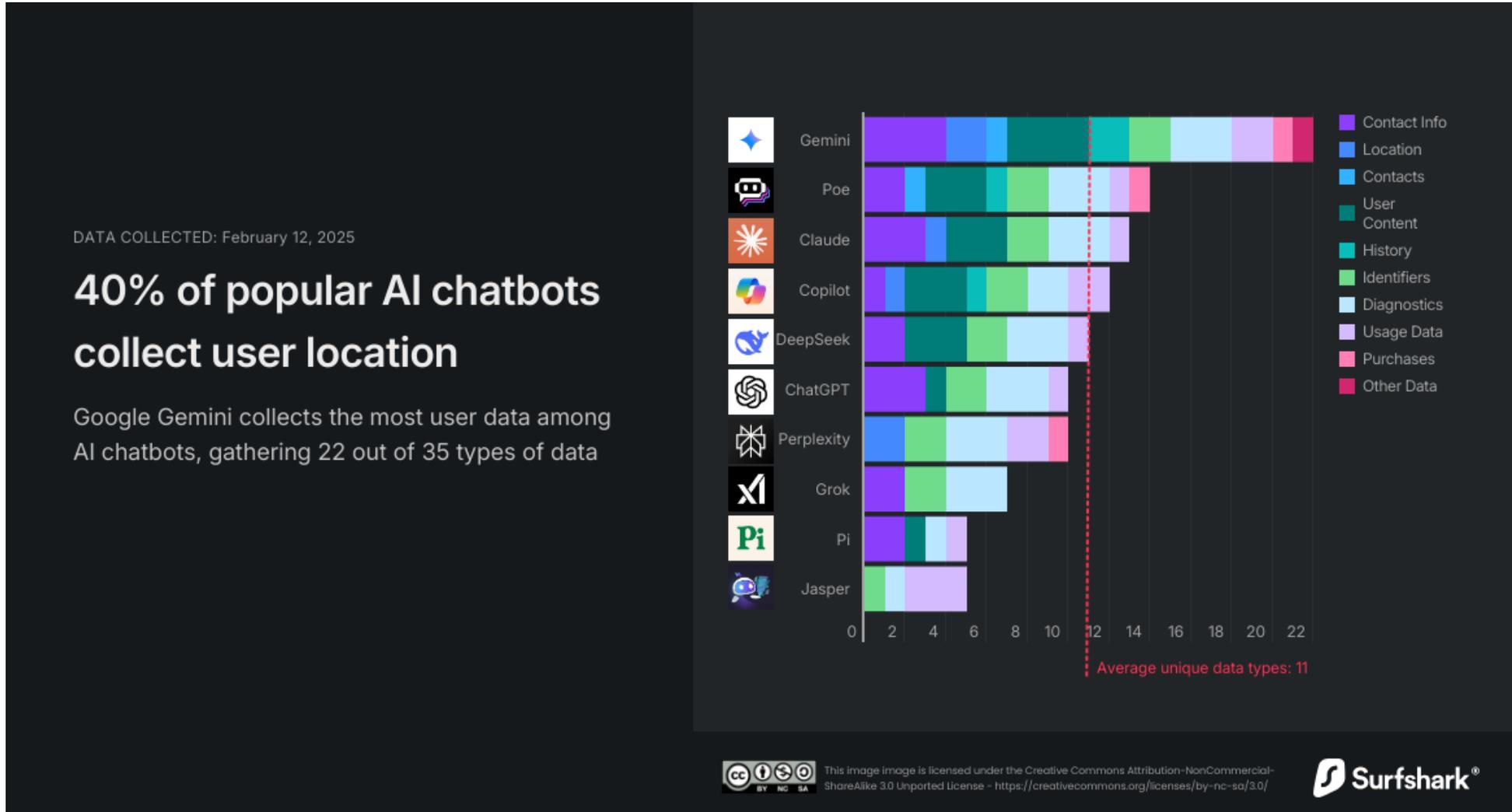
- 30 interventions (conférences, webinaires, tables rondes,...)
- Articles en réflexion / préparation



.....



Services gratuits d'IA générative : collecte de données



[Article 30% of popular AI chatbots share data with third parties, 18 fév. 2025](https://surfshark.com/research/chart/ai-chatbots-privacy)

[https://surfshark.com/research/chart/ai-chatbots-privacy...](https://surfshark.com/research/chart/ai-chatbots-privacy)

Etude Ipsos « L'usage de l'IA par les français » fév. 2025

Selon vous, quels sont les principaux risques liés à l'utilisation des IA génératives ?



Base = 155 utilisateurs professionnels / 328 utilisateurs privés
[L'usage de l'intelligence artificielle par les Français](#), février 2025

Pendant la ruée vers l'or, vendez des pelles et des pioches⁺



NVIDIA Corp
NASDAQ : NVDA

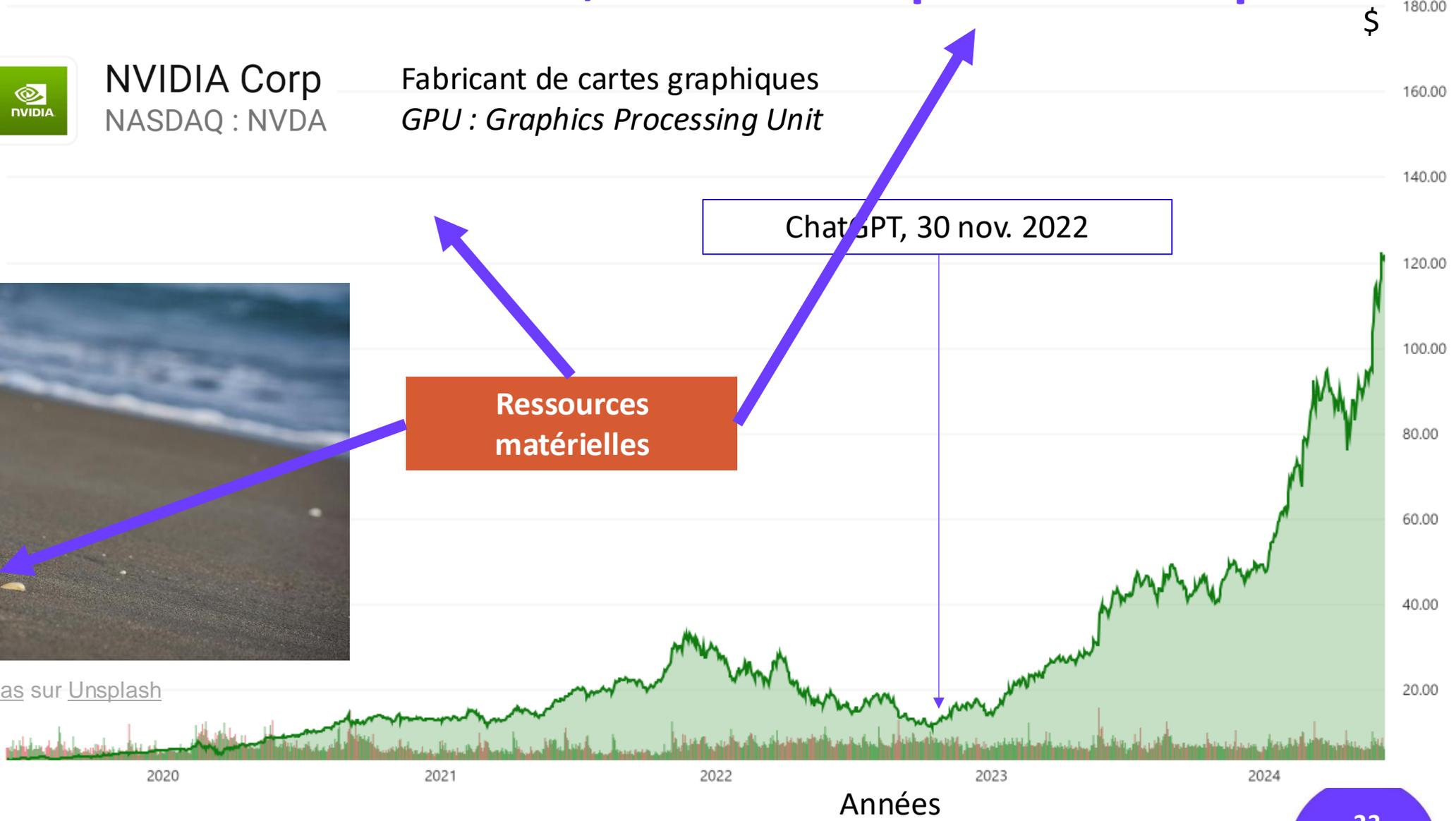
Fabricant de cartes graphiques
GPU : Graphics Processing Unit



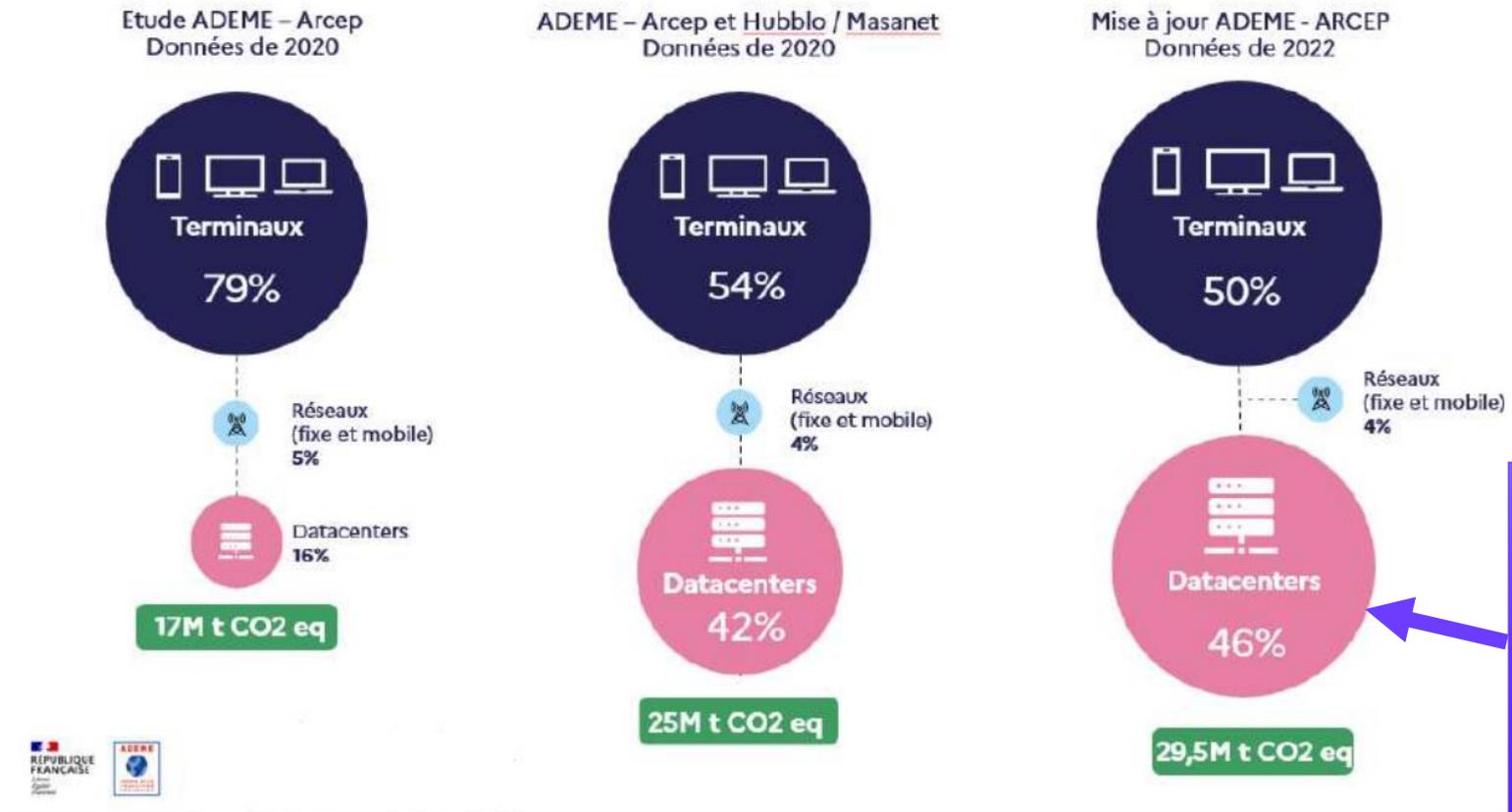
Photo de Alejandro Alas sur Unsplash

Ressources matérielles

ChatGPT, 30 nov. 2022



Impact environnemental du numérique en France



En prenant en compte les datacenters hébergés à l'étranger mais ! datacenters IA récents non pris en compte

Figure 4 – Comparaison des résultats avec les études précédentes – par tier

Limites & approches

IA générative : 10 à 1000 fois (texte à génération d'image) plus impactant qu'une recherche web
 Charge smartphone à 100% = 2 images générées

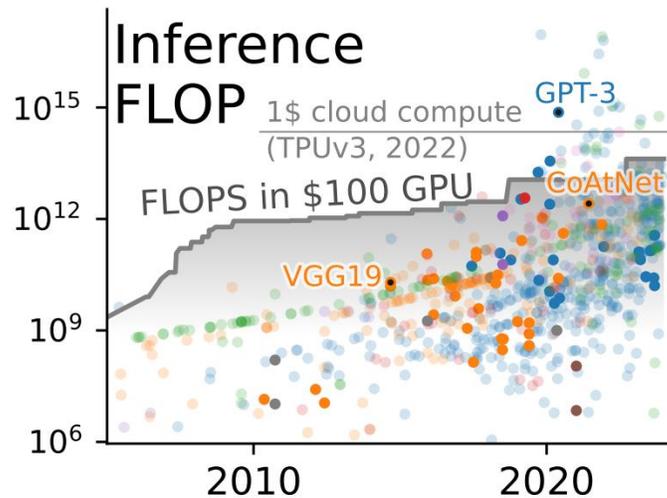


Figure 3. The cost of a single inference is growing faster than compute is improving

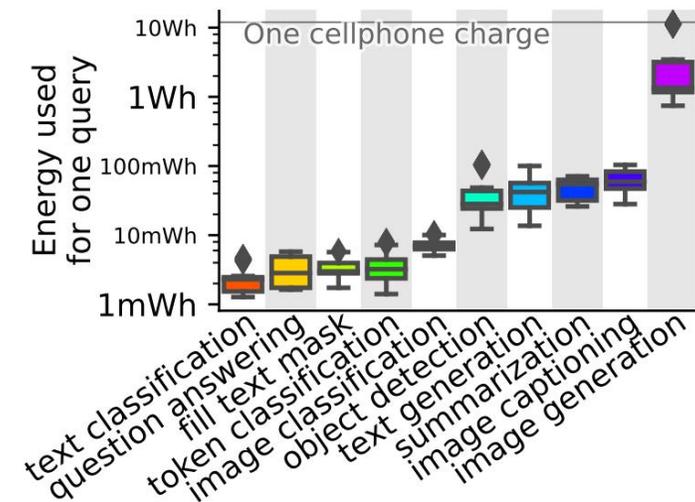
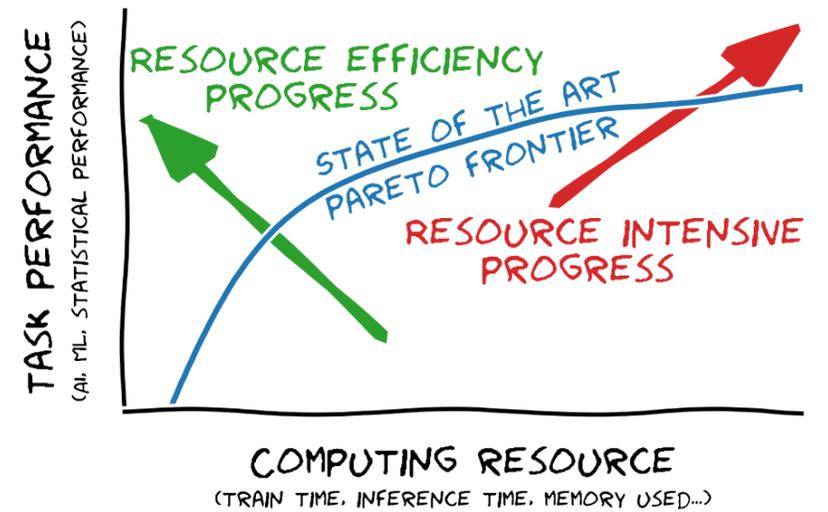
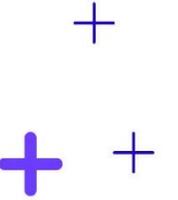


Figure 4. A single inference uses more energy for models with broad purposes. Data from Luccioni et al. (2024).

Varoquaux, G., Luccioni, A. S., & Whittaker, M. (mars 2025). Hype, Sustainability, and the Price of the Bigger-is-Better Paradigm in AI. <http://arxiv.org/abs/2409.14160>

Réflexions stratégiques : utile & nécessaire



Utilité & sobriété numérique

- IA : tendance de fond
- IA générative : viser utile, utilisable, utilisé

Stratégie « juste nécessaire » à établir

- Collaborer : communautés internes, autres acteurs publics
- Mutualisation : fédération infrastructures ESR (projet ILaaS)
- Déterminer les cas d'usages pertinents
- Cadrage : charte & guide de recommandations (travaux ESR)

Agir & construire ensemble

- Expérimenter solutions et pratiques : vous aussi, c'est possible : <https://ragarenn.eskemm-numerique.fr/>
- Partage retours, cas d'usages, réflexions : expérimenter et échanger





Echanges sur les autres projets en cours

Un point sur le projet ILaaS

Un projet qui vise à proposer un accès API à des capacités d'IA générative (inférence LLM) pour des services numériques « consommateurs » en fédérant des serveurs GPU grâce à Aristote-dispatcher.

=> Une nouvelle fédération de serveurs de l'ESR pour favoriser résilience, confiance, souveraineté, sobriété numérique en IA générative

=> Une recherche d'optimisation de l'utilisation, d'amélioration de la résilience, et d'évaluation de l'impact environnemental afin de viser des usages vertueux

Présentation lors des ESUP-DAYS le 1er avril

<https://www.esup-portail.org/esup-days>



Support et replay disponibles sur le portail :

<https://videos.esup-portail.org/esup-days/esupdays39/video/1775-esupdays39-01042025-matin/>



Cycle de webconférences

"Transition numérique : focus sur l'IA dans l'ESR"

Prochain rendez-vous :

15 mai 17h/18h - Dispositifs d'accompagnement des enseignants et des étudiants [Je m'inscris](#)

Revoir la 1^{ère} webconférence :

14 mars – Charte IA [Accéder au replay](#)

