



Commissariat Général au
Développement Durable



L'Intelligence Artificielle en santé-environnement, au-delà de la boîte noire

Applications, perspectives et enjeux

Actes du séminaire

21 novembre 2024

Pour écouter ou réécouter cette journée :
<https://www.youtube.com/playlist?list=PLI00sylAMv7S-RVFAieSqFpEq3hBORnbA>

Sommaire

Introduction	1
Master Class : Intelligence artificielle et Santé Environnement	2
I) IA générative et santé globale	2
II) L'IA pour la Santé Environnement : tour d'horizon et perspectives	3
III) Échanges	5
Session 1 : L'IA pour la collecte et le prétraitement de nouvelles données	6
I) Extraire pour collecter de nouvelles données environnementales et de santé	6
II) Modéliser pour proposer de nouvelles données environnementales et de santé	9
Capsule démo : nouveaux moteurs de recherche académiques basés sur de l'IA générative - dans quels cas les utiliser aujourd'hui et quelles sont leurs limites ?	13
Pitches : IA et santé-environnement dans les territoires	14
I) Predict AI'r: des données de téléphonie à une meilleure qualité de l'air grâce à l'IA	14
II) AMELIA : développer une cartographie dynamique air bruit mobilités	15
Session 2 : L'IA au service de l'analyse de données	16
I) Mobiliser le <i>machine learning</i> pour mesurer l'hétérogénéité de l'effet de la pollution de l'air sur le jeune enfant	16
II) Geospatial AI for One Health	16
III) Amélioration de scores de risque environnemental par machine learning informé et IA explicative	17
IV) L'IA en santé environnementale : cas d'étude avec la chaleur extrême	18
V) Échanges	19
Session 3 : L'IA pour répondre aux enjeux de restitution des résultats scientifiques et perceptions	22
I) L'IA générative au service de la démocratisation de l'accès aux données	22
II) Constituer un écosystème et mutualiser les corpus d'expertise pour élaborer une IA publique de confiance sur la transition écologique	22
III) Perception et acceptabilité de l'Intelligence Artificielle : l'apport des sciences comportementales	23
IV) Échanges	25
Table Ronde : Futur de l'IA en santé environnement - quels enjeux juridiques, éthiques et environnementaux	26
Annexes	33

Introduction

Caroline Semaille, Directrice générale de Santé publique France

« Je remercie les participants de leur présence, ainsi qu'Ecolab et la Direction Santé Environnement pour l'organisation de ce premier séminaire de Santé publique France dédié à l'IA. Ce séminaire est la suite logique d'un travail commun entre Santé publique France et Ecolab, qui a abouti à une synthèse d'acculturation à l'IA appliquée à la Santé-Environnement, disponible en ligne. Ce séminaire marque également une nouvelle étape pour notre Agence, dont les pratiques sont profondément transformées par l'IA. Santé publique France y voit un levier puissant pour répondre aux grands défis de santé publique. En tant que Directrice générale, je m'engage à faire de l'IA une priorité. Les opportunités sont nombreuses : améliorer la qualité des données environnementales, prédire des expositions ou renforcer notre capacité à anticiper des évènements sanitaires. L'essor des outils d'IA générative ouvre des perspectives inédites, tout en posant des questions cruciales sur la sécurité des données et l'impact environnemental auxquels Santé publique France est très sensible. Cette journée a été construite pour explorer les enjeux et les applications concrètes de l'IA dans nos études, les défis éthiques qu'elles soulèvent et les stratégies nationales à développer. Notre objectif est de construire des ponts entre les disciplines, les équipes et les institutions, afin de faire avancer les connaissances et les pratiques. Les travaux de cette journée nourriront une note d'orientation stratégique sur l'IA qui guidera nos projets en Santé-Environnement. »

Thomas Cottinet, Directeur de l'Ecolab (CGDD/SRI), Ministère de la Transition écologique

« Le sujet de l'IA en Santé Environnement est encore exploratoire. Le rapport évoqué par la Directrice générale est un bel exercice de vulgarisation. Au Ministère de la Transition Ecologique, la stratégie de déploiement de l'IA auprès des agents repose sur trois piliers :

- *L'IA au service des agents et de la modernisation publique : les outils tels que LIRIAe (outil d'aide à l'instruction des dossiers d'études d'impact environnemental) ou SOFIA (IA générative qui exploite de manière sécurisée toutes les productions du ministère, de l'ADEME et du CEREMA) requièrent un travail de formation et d'acculturation des agents ;*
- *L'IA au service des politiques publiques pour la transition écologique : Ecolab a la responsabilité de suivre les entreprises de la Greentech et a mis en place un dispositif de démonstrateurs territoriaux d'IA frugale (40 millions d'euros de projets) impliquant des consortiums de collectivités, chercheurs et start-up sur des sujets comme l'observation de la qualité de l'air ou le suivi du bruit*
- *L'empreinte environnementale de l'IA : Le coût environnemental de l'IA est très élevé, notamment celui de l'IA générative. L'Agence internationale de l'énergie a prédict un décuplement de la consommation énergétique d'ici 2030. Il est donc essentiel de bien réfléchir au recours à l'IA de confort et une utilisation frugale de l'IA. La France est le premier pays à avoir publié un référentiel général sur l'IA frugale, qui pourrait devenir en février 2025 une norme mondiale publiée lors du Sommet mondial de l'IA de Paris. »*



Master Class : Intelligence artificielle et Santé Environnement

I) IA générative et santé globale

Xavier Fresquet, Sorbonne Center for AI

De nombreux pays ont adopté une stratégie nationale dédiée à l'intelligence artificielle (IA), où la santé figure parmi les axes prioritaires. En France, la stratégie initiée par le rapport Villani en 2018 a conduit à la création de neuf Centres IA, lauréats de l'appel à projets IA Cluster lancé en 2023 et cofinancé par des entreprises. Le Centre IA de Sorbonne Université, bénéficiaire d'un financement de 70 millions d'euros, porte des projets de recherche stratégiques avec de nombreux partenaires, tout en explorant des thématiques émergentes via un programme transverse.

La recherche fondamentale sur l'IA n'est plus l'apanage du monde académique. Des entreprises comme OpenAI, Meta, Google DeepMind ou la licorne française Hugging Face jouent un rôle moteur. Depuis son émergence dans les années 1950, l'IA vise à modéliser la prise de décision humaine et à prédire des données manquantes avec précision. Son évolution inclut le machine learning (années 1980) et, plus récemment, le deep learning (années 2010), qui utilise des réseaux de neurones profonds pour apprendre automatiquement à partir de données massives. Aujourd'hui, l'IA générative, sous-domaine du deep learning, se distingue par sa capacité à créer de nouvelles données. Toutefois, son impact environnemental élevé impose une utilisation responsable. À cet effet, le gouvernement a mis en place Compar :IA, un outil permettant de comparer les performances de modèles d'IA et leur empreinte écologique (academie-medecine.fr).

Les fondamentaux de l'apprentissage par l'IA :

L'intelligence artificielle (IA) repose sur trois piliers fondamentaux : les données, la puissance de calcul et les talents. Elle se décline en plusieurs approches distinctes :

- IA basée sur les règles : Cette méthode, explicite et interprétable, modélise la prise de décision d'un expert en suivant des règles prédéfinies.
- Apprentissages automatiques (machine learning) et apprentissage profond (deep learning) :
 - Apprentissage supervisé : Cette approche repose sur des données étiquetées, permettant au modèle d'apprendre à partir d'exemples annotés.
 - Apprentissage non supervisé : Cette méthode identifie automatiquement des structures (clustering) à partir de vastes ensembles de données, sans nécessiter de supervision explicite.
 - Apprentissage par renforcement : Cette technique entraîne un agent virtuel qui apprend par essais et erreurs, trouvant des applications notamment dans la robotique et les véhicules autonomes.

Applications et enjeux :

Diagnostic et traitement assistés par IA : L'IA générative est utilisée pour analyser des données médicales complexes, facilitant ainsi le diagnostic et le choix des traitements.

Agents numériques de promotion de la santé : L'OMS a récemment lancé un agent numérique exploitant l'IA générative pour fournir des informations de santé précises et accessibles (who.int).

Explicabilité : Facile pour les modèles basés sur des règles, elle reste un défi pour les réseaux de neurones, bien que des outils permettent de retracer l'origine d'une erreur (academie-medecine.fr).

Biais et discriminations : Les modèles peuvent reproduire des biais présents dans les données d'entraînement, ce qui peut entraîner des discriminations, notamment en matière de genre ou d'origine ethnique.

Impact environnemental : La formation des grands modèles d'IA nécessite des ressources énergétiques colossales. Pour y répondre, OpenAI aurait acquis des parts dans une centrale nucléaire afin de garantir une alimentation énergétique stable et de haute capacité (<https://thebulletin.org/2024/12/ai-goes-nuclear/>).

IA générative et santé publique :

L'intelligence artificielle générative a révolutionné plusieurs disciplines scientifiques, en particulier la physique, la biologie et la chimie, contribuant à des découvertes majeures reconnues par l'attribution de prix Nobel en 2024. Cette technologie est déjà utilisée pour produire non seulement des textes et des images, mais aussi des vidéos à partir de simples descriptions textuelles. La rapidité avec laquelle les outils tels que les modèles de langage de grande taille (LLM), comme ChatGPT, ont été adoptés est spectaculaire, redéfinissant les modes de diffusion de l'information et posant de nouveaux défis quant à la véracité et la fiabilité des contenus générés.

En France, des startups innovantes de la French Tech se distinguent par leur contribution dans le domaine de l'IA appliquée à la santé publique. Parmi elles :

- **BioSerenity** : Spécialisée dans la santé connectée, elle développe des solutions intégrant des algorithmes d'IA pour le diagnostic et le suivi des maladies chroniques, notamment en neurologie et en cardiologie.
- **Withings** : Reconnu pour ses dispositifs de santé connectés, comme les montres et balances intelligentes, la société explore l'IA pour offrir des analyses prédictives et personnalisées à partir des données de santé des utilisateurs.
- **Bioptimus** : Cette startup se distingue par le développement d'un modèle de langage ouvert (LLM) spécifiquement conçu pour comprendre les mécanismes biologiques complexes et les interactions à différentes échelles de la santé humaine. Ce LLM vise à accélérer la recherche biomédicale et à optimiser les pratiques cliniques grâce à une meilleure analyse des données biologiques ([Les Échos](#)).

II) L'IA pour la Santé Environnement : tour d'horizon et perspectives

Laura Chaperon, Santé Publique France

La synthèse des travaux menés par Santé publique France et Ecolab a été publiée en juillet 2024, dans un objectif de vulgarisation de l'IA (concepts généraux et pour la santé environnement spécifiquement). L'IA est relativement discrète en Santé-Environnement, et cette synthèse s'est appuyée à la fois sur des recherches bibliographiques larges et sur des entretiens avec les acteurs français qui sont familiarisés avec les méthodes d'IA et/ou qui ont mené des travaux mobilisant l'IA dans le champ de la santé environnement. L'IA existe depuis les années 50 : même si elle a connu quelques hivers, elle a envahi notre quotidien. Elle cherche à mimer l'intelligence humaine, voire à la dépasser. Les enjeux autour de l'IA

sont importants, car le sujet peut en effrayer certains. Santé publique France a une forte volonté d'acculturation pour embarquer tous ses agents vers la connaissance et/l'utilisation de l'IA, un sujet prioritaire de l'Agence. Il est indispensable de disposer aujourd'hui d'outils permettant de synthétiser et traiter les énormes volumes de données en environnement et en santé.

En Santé-Environnement, la GéolA a la particularité d'utiliser des données environnementales géoréférencées. L'IA peut permettre un plus grand déploiement des outils de SIG (systèmes d'informations géographiques) et obtenir une simulation de données géographiques, de données manquantes, de scénarios du changement climatique ou encore aider les instances à prendre des décisions et gérer les risques.

Cette journée, qui reprend les grandes parties de la synthèse, est organisée en 4 thèmes :

- *La collecte des données* : l'objectif des outils IA est de permettre l'extraction de données dans des formats non ou peu exploitables vers des formats plus aisément mobilisables (notamment pour les données environnementales). Il existe de nombreux cas d'usage comme l'utilisation de données environnementales captées par les citoyens ou l'extraction de documents PDF pour alimenter des bases de données (OFB) ;
- *Le prétraitement et la consolidation de données* : il s'agit d'améliorer la qualité des données en utilisant l'IA pour l'imputation de valeurs manquantes ou permettre une meilleure exploitation croisée de bases de données de natures différentes (indicateurs de santé, données de mesures individuelles ou économiques...). Parmi les projets présentés aujourd'hui, on peut citer l'outil CosIA permettant à l'IGN de reconstituer des données d'occupation du sol ;
- *L'analyse de données* : il s'agit d'explorer, expliquer ou prévoir un phénomène. Les prévisions du changement climatique sont complexes, et l'INSPQ a mis en place un outil de prévision des effets de la chaleur extrême sur la santé ;
- *La valorisation et la communication des données et analyses* : une fois les données traitées, les résultats doivent être facilement exploitables et visualisables afin de les fournir aux décideurs et d'adapter le message au public. Plusieurs outils conversationnels sont particulièrement inspirants pour l'Agence : SARAH (OMS), Climate Q&A (base des rapports du GIEC), SOFIA (ADEME).

La dernière partie de la journée porte sur les grands enjeux de l'IA : fiabilité des données (qui garantit la fiabilité de l'outil), accessibilité des données (open source) versus protection des données (le domaine Santé-Environnement est à la croisée des données environnementales et de santé), vigilance quant aux fausses associations épidémiologiques. Seront également abordés les enjeux éthiques et juridiques (protection des données personnelles et des libertés, transparence et explicabilité des algorithmes des données, biais et discriminations algorithmiques, bouleversements du monde du travail) ainsi que les enjeux écologiques (le numérique est responsable de 10 % des émissions CO₂, le référentiel sur l'IA frugale pourrait être un moteur d'innovation).

Enfin, le rapport s'achève sur les perspectives de l'IA en Santé-Environnement. L'INRS a envisagé 4 scénarios d'ici 2035 dans sa prospective « L'intelligence artificielle au service de la santé et de la sécurité au travail, enjeux et perspectives à l'horizon 2035 » : une IA hors de contrôle, une IA encadrée réglementairement (IA Act européen, qui pourrait devenir un frein à l'innovation), une IA citoyenne (Open source, maîtrisée par tous) et enfin un nouvel hiver de l'IA.



III) Échanges

- *Vous parlez de chercheurs étatsuniens et européens, mais qu'en est-il des chercheurs chinois ?*

Xavier Fresquet

La compétition est féroce dans le développement de l'IA et du *machine learning* : les grands pays qui développent ces outils sont d'abord la Chine puis les États-Unis. Le Canada est également un acteur majeur. La France et l'Europe essaient de rester dans la compétition, sans atteindre le niveau des deux leaders. La France essaie de développer un axe France-Canada-Japon.

RESSOURCES & REFERENCES

1. Outil Compar : IA : <https://www.comparia.beta.gouv.fr/>
2. L'intelligence artificielle au service de la santé et de la sécurité au travail, enjeux et perspectives à l'horizon 2035 : <https://www.inrs.fr/footer/actes-evenements/prospective-IA.html>
3. Synthèse Santé publique France :<https://www.santepubliquefrance.fr/docs/la-synthese-des-connaissances-et-perspectives-pour-la-sante-environnementale-a-sante-publique-france>
4. SARAH : <https://who-fr.digitalhero.cloud/landing/index.html>
5. Climate Q & A : <https://www.climateqa.com/>



Session 1 : L'IA pour la collecte et le prétraitement de nouvelles données

Matthieu Porte, coordinateur des activités IA à l'IGN, modérateur

I) Extraire pour collecter de nouvelles données environnementales et de santé

1) *Extraction de données structurées sur les services d'eau et d'assainissement à partir de PDFs (projet NARVAL)*

Geneviève Fleury, data scientist dans l'équipe « données et appui méthodologique » et Grégoire Etot, responsable de l'équipe « données et appui méthodologique » à l'Office Français de la Biodiversité (OFB)

Grégoire ETOT

Parmi les 5 grandes missions de l'OFB, la mission de « connaissance et expertise » s'appuie sur une Direction « surveillance, évaluation, données » auquel appartient l'équipe « données et appui méthodologique ». Elle est notamment en charge de la coordination des trois grands systèmes d'information fédérateurs et partenariaux : SI Biodiversité, SI Milieu Marin et SI Eau. Le projet NARVAL concerne SISPEA (SI des services publics d'eaux et d'assainissement), un des outils du SI Eau.

Geneviève FLEURY

L'objectif de NARVAL est d'enrichir la plateforme SISPEA avec un outil d'IA permettant de rechercher des données dans des PDFs. SISPEA centralise des données sur le « petit cycle de l'eau », créé par l'homme pour s'alimenter en eau potable. Chaque année, les 25 000 services d'eau et d'assainissement en France répartis sur 14 000 collectivités ont l'obligation de saisir sur SISPEA un certain nombre d'indicateurs codifiés renseignant sur le prix et la qualité des services. En pratique, seulement 50 % des services s'acquittent de cette obligation. L'idée est de récupérer les données manquantes qui figurent dans les rapports PDFs fournis annuellement par les collectivités, en développant un outil d'IA permettant d'en extraire les valeurs des indicateurs, de les faire figurer dans un tableau Excel et d'alimenter automatiquement SISPEA. Deux contraintes sont associées à ce projet : travailler avec un outil open source et frugal.

Quelles sont les principales difficultés ? Les indicateurs se trouvent dans le texte ou dans des tableaux et il n'existe pas de modèle harmonisé de PDFs. Ceux-ci doivent par ailleurs être segmentés. La valeur d'un indicateur n'est pas toujours indiquée : le modèle doit être en capacité de dire qu'il ne trouve pas la réponse. Quels sont les atouts ? De bonnes connaissances métier à exploiter, les unités, les bornes physiques et les statistiques des indicateurs connus.

Les 4 étapes successives de la démarche sont les suivantes : lecture du PDF (extraction du texte et des tableaux), segmentation du PDF (en identifiant pour chaque indicateur les pages pertinentes en posant plusieurs questions renvoyant à plusieurs pages), utilisation d'un modèle *transformer* de langage clé en main (Llama3.1-8b) qui permet de récupérer plusieurs réponses possibles, et enfin nettoyage et sélection des réponses en mobilisant les connaissances métiers. En sortie, on obtient un tableau Excel avec les différents indicateurs. L'étape d'évaluation compare les valeurs prédites par le modèle avec celles du PDF.

Dans le projet, l'évaluation a été réalisée sur un petit jeu de 15 PDFs d'assainissement collectif et 15 indicateurs. Le taux moyen de bonnes réponses est de 82 %, mais avec une grande variabilité d'un PDF à l'autre et d'un indicateur à l'autre. La typologie de ces erreurs a été analysée : un quart des erreurs sont irréductibles (qualité du PDF, réponse ambiguë...),

certaines réponses sont dans le PDF, mais le modèle ne les trouve pas, d'autres ne sont pas dans le PDF, mais le modèle renvoie un chiffre, enfin, la réponse est parfois dans le PDF, mais le modèle renvoie un mauvais chiffre. Les performances ont été progressivement améliorées en analysant les écarts. La mise en production sur 10 000 PDFs par an représenterait 100 kg de CO_{2,eq} émis. Plusieurs pistes d'amélioration ont été identifiées.

2) *SursynDC : traitement automatique du langage et technique de machine learning pour la surveillance syndromique des décès*

Johnny Platon, data scientist et data engineer et Anne Fouillet, épidémiologiste sur la surveillance sanitaire des urgences et des décès, Direction DATA de Santé publique France

Anne FOUILLET

Nos travaux ont commencé en 2016 dans l'objectif de faire une analyse réactive des causes médicales de décès à partir des certificats électroniques de décès. Ce projet s'inscrit dans le cadre du système de surveillance syndromique SurSaUD (Surveillance Sanitaire des Urgences et des Décès) afin de réaliser une collecte automatisée, sans sélection et quotidienne de 4 sources des données de morbidité et de mortalité, dans l'objectif de détecter et surveiller de manière réactive des variations inhabituelles ou attendues des indicateurs. Une centaine d'indicateurs (regroupements syndromiques ou diagnostics médicaux) a été identifiée. Les données se trouvent dans la partie médicale du certificat électronique de décès et sont renseignées par les médecins. Une première partie décrit le processus morbide et une deuxième mentionne des comorbidités et des causes associées. Avec un délai variable, les données sont codées par le CépiDc. Trois méthodes de classification automatique des causes de décès sous forme de texte libre ont été appliquées : une méthode traditionnelle (à base de règles), une méthode par apprentissage automatique (SVM) et une méthode conjointe (SVM + traditionnelle). Les performances de chaque méthode ont été évaluées.

Johnny PLATON

Avant d'appliquer les méthodes décrites, un prétraitement des données textuelles est nécessaire afin de les normaliser (passage en minuscule, suppression des accents...).

- La méthode à base de règles s'appuie sur le dictionnaire du CépiDc qui répertorie l'ensemble des expressions des causes de décès retrouvées au moins une fois dans les certificats et des tables de standardisation de texte. Un regroupement syndromique est associé à chacune des 150 000 expressions identifiées. Cette méthode s'articule en 4 étapes : standardisation des causes de décès à l'aide des règles basées sur des expressions régulières ; segmentation des textes à l'aide de séparateurs ; utilisation d'un correcteur orthographique ; projection du dictionnaire.
- La méthode par apprentissage automatique (supervisé - méthode SVM) cherche à classer des objets (champs du certificat) dans des classes (regroupement syndromique) grâce à un ensemble d'apprentissages. Deux phases d'apprentissage ont été réalisées : construction d'un ensemble de vecteurs et prédiction (un échantillon de 3500 certificats a été annoté manuellement et utilisé pour entraîner les modèles et 1000 certificats pour faire la prédiction et évaluer la performance des modèles).
- De très bonnes performances ont été obtenues avec la méthode « combinée » en termes de précision et de rappel. L'échantillon a été construit afin d'avoir les regroupements syndromiques les plus variés possible et des regroupements syndromiques assez rares ou qui se ressemblent.

Les perspectives de ce projet sont d'obtenir une classification des causes de décès en regroupements syndromiques à partir des certificats électroniques de décès reçus chaque jour à Santé publique France, puis un calcul d'indicateurs de surveillance par regroupement syndromique. Il s'agira ensuite d'intégrer les indicateurs dans des bulletins épidémiologiques SurSaUD. Sur le plan méthodologique, il sera intéressant d'explorer et d'évaluer d'autres

méthodes de *deep learning* en comparant les performances à celles observées avec ces premiers modèles.

Anne FOUILLET

Historiquement le CépiDc codait avec le codage automatique IRIS/muse (à base de règles) et le codage manuel. De premiers travaux ont exploité des méthodes de *deep learning* pour coder automatiquement : après évaluation, ces travaux ont été mis en production pour le codage automatique d'une partie des causes de décès. Ces différents modèles plus ou moins complexes aboutissent à 4 prédictions : le CépiDc a mis en place un « surmodèle » permettant de choisir le meilleur codage. Sur les données codées lors de la campagne 2021, le CépiDc obtient une performance de 96 % de résultats. La performance est de 97 % pour un regroupement de codes. Santé publique France souhaite s'inspirer des travaux du CépiDc pour améliorer son exploitation des causes de décès pour la surveillance réactive de la mortalité.

3) *Échanges*

Grégoire REY (Inserm, France Cohortes, ancien directeur du CépiDc)

L'équipe de l'OFB a-t-elle prévu d'utiliser les scores de confiance donnés par ces modèles pour identifier ce qu'il fallait reprendre manuellement ? Concernant la 2^{ème} présentation, le CépiDc fait des prédictions à l'échelle individuelle, alors que votre objectif est de détecter des événements imprévus. Les mêmes outils ne permettent pas d'obtenir les mêmes performances et il est préférable de ne pas faire de « surajustement » par rapport au passé.

Geneviève FLEURY

Le modèle LLaMA que nous utilisons ne donne pas de score de confiance. D'autres modèles qui en fournissent pourraient être utilisés.

Anne FOUILLET

Les objectifs sont effectivement différents. Le CépiDc code chaque décès et attribue un code CIM-10. Notre objectif est de classer ces causes dans des champs plus larges et d'avoir une surveillance réactive sur les regroupements. L'identification de variations inhabituelles dans ces indicateurs nécessite peut-être de recourir à d'autres méthodes.

Milena SUAREZ CASTILLO (DREES)

La DREES est confrontée à des enjeux proches de ceux de l'OFB (complétion de valeurs manquantes dans des enquêtes, pré-remplissage de questionnaires). Il est possible de renvoyer la donnée structurée pour obtenir la réponse non sous forme de conversation, mais de schéma de données. Le nettoyage de la réponse que vous avez effectué est-il complexe ?

Geneviève FLEURY

Nous n'avons pas testé la méthode que vous évoquez, car nous interrogeons le modèle page par page : nous lui demandons de renvoyer un nombre, pas un tableau structuré (par souci de temps de calcul et de frugalité).

Grégoire ETOT

Nous n'avons pas accès à des data center puissants, mais nous pouvons néanmoins utiliser l'IA grâce à nos expertises thématiques qui permettent de réduire exponentiellement les coûts de calcul et environnementaux. Dans nos métiers, l'avenir de l'IA serait de la coupler avec une expertise thématique et de l'utiliser de manière ciblée sur une partie de la

chaîne. Nous n'utilisons l'IA ni dans les étapes amont (découpage du document) ni aval (élimination des bornes min-max).

Jean-Yves CHATELIER (INERIS)

Concernant l'annotation de documents, qui est chronophage, l'OFB a-t-il recours à des systèmes d'augmentation de données ? De même, l'équipe de Santé publique France pourrait-elle envisager des comptes-rendus automatiquement générés ?

Grégoire ETOT

Nous n'avons pas utilisé de systèmes d'augmentation de données dans ce projet, mais nous avons mené une expérimentation de ce type pour développer un outil de reconnaissance des espèces de poissons à partir de photos.

Anne FOUILLET

Nous avons fait l'annotation manuellement en 2018. Des techniques actuelles permettraient d'annoter plus systématiquement en volume et feraient gagner en performance. Elles sont à évaluer sur nos données.

II) Modéliser pour proposer de nouvelles données environnementales et de santé

1) *CoSIA : construire les cartes de couverture du sol par Intelligence Artificielle*

Eva Bookjans, Data scientist à l'IGN

Les cartes CoSIA distinguent 15 classes de couverture de sols (y inclus différents types de végétation). Elles ont été développées pour la production de l'OCS GE (occupation de sol à grande échelle) qui permet notamment de mesurer l'artificialisation des sols dans le contexte de la loi climat et résilience. L'IGN a utilisé l'IA pour accélérer la production de cartes de couverture de sol et réduire les coûts d'annotation d'images. L'objectif est de détecter la couverture de sol à partir d'images aériennes (résolution de 20 cm x 20 cm). Le modèle retenu est celui de la segmentation sémantique (chaque pixel est attribué à une classe) basé sur un apprentissage supervisé avec la création d'un jeu d'annotations manuelles puis l'entraînement d'un modèle, en prenant en compte le verrou de la généralisation spatio-temporelle. Le premier test réalisé en 2022 sur le Gers a été satisfaisant, mais ne pouvait pas être étendu au territoire français dans sa globalité. Les classes ont été améliorées en intégrant d'autres annotations, et le jeu de données représentait en octobre 2023 2 623 km². Des modèles de réseau neuronal pré-entraînés (ResNet) ont également été utilisés. Une nouvelle architecture d'apprentissage profond (issue d'un défi scientifique FLAIR basé sur l'attention et qui prend davantage en compte le contexte spatial) a permis d'améliorer les métriques.

CoSIA a été développé pour des utilisateurs de produits vecteurs, qui prévoient des étapes de simplification à partir des pixels en sortie brute des modèles IA. La production de CoSIA est entièrement automatique et ce qui se trouve sous la canopée n'est pas corrigé. Des erreurs de confusion entre classes demeurent également (culture versus pelouse...). Les données sont disponibles en Open Data (CoSIA.ign.fr).

Les applications de CoSIA sont nombreuses : mesure de l'imperméabilité, le foncier, la nature en ville (ex. : modélisation des îlots de chaleur), la surface agricole utile, les haies et bocages, détection de corridors écologiques... CoSIA permet également de détecter les changements. Le projet GEO-K-PHYTO de l'IGN cartographie les vignes et les vergers, afin de produire des indicateurs spatialisés d'exposition indirecte aux pesticides et d'identifier de potentiels liens avec certains types de cancer.

Les modèles entraînés et diffusés sont spécifiques aux caractéristiques du produit BD Ortho et au territoire de France métropolitaine. Ils doivent être réentraînés pour traiter d'autres modalités, territoires ou tâches. Par exemple, le modèle CoSIA peut être entraîné sur des images satellites ou des images aériennes historiques en utilisant du transfert d'apprentissage. En perspective, les nouvelles architectures (modèle lauréat du défi FLAIR) seront accessibles sur Hugging Face début 2025, le jeu de données complet et multimodal sera diffusé et CoSIA sera utilisée sur d'anciens millésimes de BD Ortho.

2) *Gérer les ressources et les risques du sol et du sous-sol : apport et potentiels de l'IA au BRGM*

Théophile Lohier, ingénieur en traitement de données BRGM

Les activités du BRGM (Bureau de Recherche Géologique et Minière) sont structurées autour de 5 grandes thématiques métiers - géologie, gestion des eaux souterraines, gestion des risques et aménagement du territoire, ressources minérales et économie circulaire, transition énergétique et espace souterrain (géothermie) – et d'un 6^{ème} axe autour de la donnée et des services et infrastructures numériques à mettre en place pour partager les recherches. Le BRGM n'a pas d'équipe dédiée à la recherche en IA, mais compte une dizaine d'experts en IA qui accompagnent les experts thématiques dans le choix et la mise en œuvre des algorithmes d'IA et *machine learning*.

Concernant les prévisions de séries temporelles (une des 4 grandes catégories d'activité en *machine learning*), le BRGM travaille notamment sur les données temporelles de données piézométriques permettant de suivre le niveau des nappes phréatiques. Quelques outils permettent déjà de réaliser des projections du comportement des nappes et l'objectif est de construire des méthodes plus parcimonieuses avec de l'IA frugale. Le projet Previzo s'appuie ainsi sur une acquisition *in situ* de données sur des cours d'eau de surface (Cher) ; sur des données du réseau piézométrique du BRGM et enfin sur des images satellites et données météorologiques (Copernicus). Pour construire des modèles ne nécessitant pas de consommer de gros volumes de données ni d'une grande puissance de calcul pour s'entraîner et calibrer les paramètres. Le projet vise à entraîner un modèle par type ou par famille de séries piézométriques et de faire des estimations pour faciliter la prévision et la gestion des nappes.

En matière d'exploration minière et d'étude de sites pollués, la recherche de modèles parcimonieux est liée à un souci d'économie. Des techniques d'IA *in situ* sont développées, grâce à un laboratoire mobile et une série de capteurs permettant de réaliser de premières analyses sur des données très volumineuses et coûteuses à traiter (méthodes de réductions spatiales, spectrales et de sélections de variables qui permettent d'identifier sur des carottes de forage des groupes de pixels au comportement similaire (approche non supervisée) et de caractériser la nature et le volume des minéraux qui constituent l'assemblage (approche supervisée).

Le BRGM réalise également des cartographies plus génériques et prédictives qui s'intéressent au sous-sol et s'appuient sur des approches semi-supervisées ou supervisées. Les sources d'informations géolocalisées sont nombreuses, mais les points d'échantillonnage sont rares. Le projet PEPR Terre Numérique vise à identifier des formations géologiques de surface à partir de modèles numériques de terrain. Le projet interne DROP vise à déterminer le potentiel de présence d'une ressource minérale dans le Massif armoricain (carte de favorabilité) pour faciliter la prospection. Le projet européen ISLANDR, à partir de mesures ponctuelles et de co-variables (géologie, présence d'industries, de villes...) vise à déterminer des probabilités d'occurrence de pollution.

Le BRGM et l'ADEME commencent à travailler sur le traitement du langage naturel pour répondre à un fort besoin d'extraction de rapports de forage, scans et bases documentaires. Le premier outil trouve les coordonnées du forage, mais pas le code BSS (identifiant du forage).

Quelles sont les perspectives ? Le BRGM souhaite monter en puissance sur l'utilisation de l'IA générative et des modèles de langage pour répondre aux besoins des experts. Le BRGM souhaite également développer le traitement massif de données géophysiques ou de télédétection et produire des jumeaux numériques facilitant le suivi et l'adaptation aux enjeux liés aux géosciences (ressources en eau, gestion du littoral, aménagement du territoire).



3) Échanges

Question à distance

Le projet CoSIA a-t-il envisagé l'utilisation du LiDAR (*Light Detection and Ranging* ou télédétection par laser) pour voir sous la canopée ?

Eva BOOKJANS

Il n'est pas prévu de faire évoluer CoSIA pour le moment. L'IGN a acheté un LiDAR et réalise des MNT (modèles numériques de terrain) et des MNS (modèles numériques de surface) pour détecter la couverture au sol. Cette base de données, ainsi que BD TOPO, pourront être croisées avec CoSIA.

Matthieu PORTE

La relation est presque inverse. CoSIA permet de guider la classification des nuages de points LiDAR HD en donnant un indice de la végétation et en le croisant avec la donnée 3D.

De la salle

Vous utilisez beaucoup de données existantes : au regard des objectifs de frugalité soulevés par chacun des intervenants, ne serait-il pas pertinent de demander aux agriculteurs de déclarer certaines informations (ex. type de cultures semées) qui permettraient de préciser les modèles IA ? L'IGN peut-il descendre au niveau du type de culture ? De même, serait-il envisageable de proposer aux collectivités des modèles de rapport pour obtenir des rapports harmonisés sur les indicateurs d'assainissement d'eau ? Enfin, le BRGM a-t-il utilisé des méthodes statistiques classiques pour la prédition ?

Eva BOOKJANS

Les images aériennes utilisées par l'IGN permettent de détecter facilement certaines cultures comme les vignes. Les déclarations de cultures (RPG) existent dans le cadre des subventions de l'UE à l'agriculture. L'IA permet de détecter les cultures pour vérifier si la réglementation est respectée : l'IGN utilise les déclarations RPG pour réaliser des annotations et entraîner les modèles, en utilisant des images satellites pendant une saison. Des projets sentinelles sont lancés sur le sujet, avec des radars permettant de voir sous la couche nuageuse.

Théophile LOHIER

Concernant l'exploitation des PDFs, l'idée est d'exploiter les données historiques. Le BRGM travaille sur des outils permettant de standardiser la manière dont la donnée remonte, mais il est difficile d'imposer un format de document aux compagnies de forage. Le BRGM n'utilise pas de méthodes statistiques comme le krigage (très supervisée) par manque de données pour la calibration et la construction de variogrammes.

Geneviève FLEURY

L'OFB souhaite également exploiter les données historiques. Nous avons effectivement proposé un modèle, mais il est peu utilisé.

De la salle

D'autres pays ou consortium internationaux ont-ils appliqué le même genre de modèle, ce qui permettrait à l'IGN ou au BRGM d'exploiter leurs bases de données ?

Eva BOOKJANS

Certains organismes sont intéressés pour utiliser nos modèles (Swisstopo). Le Laboratoire en sciences et technologies de l'information géographique pour la ville et les territoires durables (LASTIG) de l'IGN (unité mixte de recherche entre l'IGN, l'Université Gustave Eiffel et l'École d'ingénieurs de la ville de Paris (EIVP)) participe aux consortiums. La première augmentation a été réalisée avec le ResNet (réseau neuronal artificiel), qui propose des modèles en ligne accessibles à tous, parfois pré-entraînés. Dans le cadre du défi FLAIR (Challenge IA et occupation du sol), l'IGN a engagé des experts pour l'aider à identifier de nouvelles solutions.

Matthieu PORTE

En termes d'IA pour la télédétection, il existe très peu de grands jeux de données de ce type (grand volume, imagerie haute résolution et annotation manuelle) au niveau mondial. Les jeux de données existants relèvent d'une échelle pré *deep learning* et ne suffisent pas à déterminer l'apport des méthodes d'apprentissage profond. Les transferts depuis d'autres modèles sont néanmoins examinés (modèles fondation). L'annotation est le volet le plus coûteux.

Eva BOOKJANS

Elle est en grande partie réalisée par des sous-traitants. Les photo-interprètes internes réalisent des contrôles qualité. Des contrôles et des croisements de données ont été effectués sur le modèle IA relatif à l'artificialisation des sols. Les images satellites n'offrent pas une résolution aussi fine.

Matthieu PORTE

Sur certains sujets, il est nécessaire d'avoir une expertise très pointue de lecture des images aériennes (ex : classification des essences forestières).

RESSOURCES & REFERENCES

1. Plateforme de partage des modèles d'IA : [Hugging Face – The AI community building the future.](https://huggingface.co/)
2. Outil CoSIA : <https://cosia.ign.fr/>
3. L'Atlas « Cartographier l'Anthropocène » 2024 - spécial IA : <https://www.ign.fr/espace-presse/ign-publie-son-atlas-2024-cartographier-lanthropocene-sur-la-thematique-de-lintelligence-artificielle-au-service-de-la-transition-ecologique>
4. Le laboratoire LASTIG : <https://ensg.eu/fr/lastig>
5. Défi FLAIR : https://ignf.github.io/FLAIR/index_fr.html



Capsule démo : nouveaux moteurs de recherche académiques basés sur de l'IA générative - dans quels cas les utiliser aujourd'hui et quelles sont leurs limites ?

Laetitia Haroutunian, responsable de la documentation, Santé publique France

La recherche de données bibliographiques scientifiques est une activité centrale des documentalistes et des experts en santé publique. Actuellement, un travail d'évaluation est en cours au sein du service de documentation de Santé publique France afin d'examiner la pertinence de l'apport de l'IA pour cette activité.

Les fonctionnalités offertes par les éditeurs de bases de données ont évolué au gré du perfectionnement des algorithmes de classement et de classification. Ainsi depuis 2010, de nouvelles fonctionnalités apparaissent chez les principaux acteurs, telles que l'aide à l'extraction des données, la classification automatique, la traduction, les réseaux de citations.

L'IA générative généraliste a fait son apparition en 2020, et de nouveaux modèles entraînés sur des données académiques ont vu le jour. Plus récemment, les RAG (Retrieval Augmented Generation) permettent d'associer des connaissances à ces modèles généralistes (sous forme de base de données vectorielles) et permettent aux systèmes de fournir des réponses plus précises en citant la source des documents.

Ces nouveaux moteurs de recherche bibliographique offrent plusieurs possibilités pour l'utilisateur : confier sa recherche en totalité à l'IA sous forme de question, faciliter la lecture des résultats, extraire des données depuis les publications. Quelques exemples : Dimensions AI, Consensus, Research Rabbit, SCISPACE (le site <https://thereisanaiforthat.com> liste l'offre actuelle d'applications basées sur de l'IA, notamment pour la recherche documentaire). Les solutions sont souvent payantes pour accéder à toutes les fonctionnalités et un compte personnel est presque toujours nécessaire pour y accéder.

Démonstration de l'utilisation de SCISPACE qui permet de rechercher et extraire des données, « discuter » avec un article PDF, identifier les concepts clés et produire des slides ou des vidéos à partir d'un article.

Ce type d'outil présente plusieurs avantages : il peut fournir des idées sur un nouveau sujet, aider à repérer des études d'intérêt en complément de la recherche « classique », aider à rechercher du vocabulaire ou traduire, faciliter la lecture et la restitution des résultats.

Toutefois, l'interrogation sur ce type d'outil doit s'accompagner d'une certaine vigilance : manque de reproductibilité, corpus de sources souvent opaque et limité à de l'Open Access, manque d'exhaustivité dans le recueil des données. Et enfin son utilisation nécessite un contrôle sur la qualité des références mentionnées.

RESSOURCES & REFERENCES

1. Outil SCISPACE : <https://typeset.io/fr>
2. There Is An AI For That : <https://thereisanaiforthat.com>



Pitches : IA et santé-environnement dans les territoires

Paul Grignon, Ecolab (CGDD/SRI), modérateur

Les applications de l'IA en Santé Environnement sont également incarnées dans les territoires. Ecolab a mis en place une « Communauté des Acteurs de l'IA dans les territoires pour la transition écologique » en mai 2022 et plusieurs initiatives ont émergé. Nous vous présentons deux projets lauréats de l'appel à projets France 2030 « Démonstrateurs d'IA frugale au service de la transition écologique des territoires » (DIAT).

I) Predict AI'r: des données de téléphonie à une meilleure qualité de l'air grâce à l'IA

Giovanni De Nunzio, chef de projets à l'IFP Énergies nouvelles

Pour évaluer la qualité de l'air dans les territoires, les pouvoirs publics peuvent la mesurer ou modéliser en caractérisant l'impact de la mobilité sur les émissions de polluants via des simulateurs de mobilité, mais ces deux solutions présentent des limites. Les mesures ne permettent pas de discriminer la contribution de la mobilité aux émissions. Les outils d'observation et de planification de la mobilité s'appuient aujourd'hui sur des données statiques (les enquêtes de mobilité), souvent obsolètes et partielles. Ils ne permettent donc pas d'anticiper les changements de comportement et d'évaluer avec précision l'impact des politiques publiques sur les émissions de CO₂ et la qualité de l'air. Il est donc nécessaire de mettre en place des outils permettant de confirmer l'efficacité de ces politiques et mesures, mais sur quelles données peuvent-ils s'appuyer ? Les données GPS acquises par les véhicules ne renseignent pas sur les autres types de transport, les données billettiques manquent de précision et ne concernent que l'échelle locale, les données des GAFAM sont privées, non ouvertes et posent un problème de souveraineté, les données GPS des applications téléphoniques posent un problème de RGPD. Les données de bornage des opérateurs de téléphonie semblent être un bon compromis, car elles captent tous les modes de transport, ont un panel d'utilisateurs peu biaisé, mais manquent de précision. En effet, seule la position de l'antenne ayant relayé l'information (échanges de données, appels ou sms) est collectée pour la facturation. Il est donc nécessaire de reconstruire la mobilité avec une résolution plus fine.

Le projet Predict AI'r, porté par un consortium de 5 partenaires (Paris Ouest la Défense, IFP Energies nouvelles, Orange, CITEPA et Air&D) démarera en décembre 2024. Comment Predict AI'r se propose d'aller de la donnée de téléphonie à une meilleure qualité de l'air grâce à l'IA ? L'IA est au cœur du projet Predict AI'r et permettra d'exploiter les données massives de téléphonie mobile pour caractériser la mobilité et ses impacts sur le climat et l'environnement. L'idée n'est pas de proposer un modèle unique de bout en bout, mais d'utiliser un modèle d'IA spécifique à chacune des 5 étapes suivantes :

- Génération des matrices O-D (origine-destination) au pas horaire : modèles de classification ;
- Itinéraires O-D : modèles de génération ;
- Modes de transport associés aux déplacements : modèles de regroupement et classification ;
- Sources d'émissions (inventaire d'émissions dynamique alimenté par des données temps réel) : modèles d'estimation et prédiction ;
- Dispersion atmosphérique (en utilisant les sources d'émission dans des modèles de dispersion) : modèles d'estimation et prédiction.

L'objectif est de répliquer cette méthodologie sur plusieurs territoires ruraux et urbains. Ce projet s'appuie par ailleurs sur une IA frugale, car la donnée de téléphonie existe déjà,

II) AMELIA : développer une cartographie dynamique air bruit mobilités

Mathieu Vaville, WaltR

Le projet AMELIA est lauréat de l'appel à projets DIAT. Son consortium est composé de l'Université Gustave Eiffel, Bruit PARIF, IFP Energies Nouvelles et WaltR et son objectif est de réaliser une cartographie des mobilités et de l'environnement pour la planification territoriale, en se focalisant dans un premier temps sur le territoire de l'établissement public territorial de Paris Est Marne & Bois (Paris Est).

Il s'agit d'exploiter le potentiel de l'IA au service d'actions concrètes contre les pollutions en réalisant des cartographies des pollutions à fine échelle (100 m x 100 m x heure), en estimant les sources de pollution et en définissant des indicateurs et des mesures objectives et enfin en combinant les données relatives à l'air, au bruit et aux mobilités. Ces cartographies seront ensuite utilisées pour identifier l'impact des actions menées par les décideurs du territoire sur la santé.

Quelles sont les données disponibles ? Les données des stations de mesure (seulement 3 sur le territoire, dont 2 très proches de Paris) ne sont pas représentatives du territoire. Les données satellitaires informeront sur les concentrations de polluants (NO₂), mais la résolution des cartes n'est pas de bonne qualité (3 km x 5 km). La fréquence mensuelle des données ne permet pas de visualiser l'action concrète sur le territoire. Dans ce projet, des données agrégées à plus fine échelle (densité de population, de routes, météorologiques, inventaires d'émission...) seront ajoutées aux données satellites existantes en utilisant un modèle d'IA et comparées aux données des stations de mesure. Le modèle XG Boost (random forest) sera utilisé pour agréger ces données et générer des cartes de concentration.

Le projet AMELIA se veut frugal en examinant quelles données sont réellement utiles, en réalisant une analyse des gains en performance potentiels et en comparant les résultats avec d'autres modèles météorologiques sans IA (Chimère) pour identifier les données d'intérêt. Même si le modèle XGBoost est d'un seul tenant, des techniques d'explicabilité peuvent être appliquées pour comprendre quels éléments ont déterminé la décision de l'IA (densité de route, de population, météo...) et méritent de faire l'objet de plan d'action par les décideurs pour améliorer la qualité de l'air.



Session 2 : L'IA au service de l'analyse de données

Yann Le Strat, Santé publique France, modérateur

I) Mobiliser le *machine learning* pour mesurer l'hétérogénéité de l'effet de la pollution de l'air sur le jeune enfant

Milena Suarez Castillo, data scientist et économètre et directrice adjointe du Lab Santé de la DREES

Le Lab Santé est chargé de diffuser les innovations sur les données de santé. La littérature fait émerger l'hypothèse selon laquelle la vulnérabilité à la pollution de l'air est variable d'un enfant à l'autre, mais évoque essentiellement des effets moyens. Cette étude montre qu'il est possible d'utiliser des techniques de *machine learning* pour analyser certaines données et s'intéresser à l'hétérogénéité de ces effets.

Si l'on suppose deux groupes d'enfants aux caractéristiques individuelles comparables, mais avec un niveau d'exposition à la pollution de l'air différent, l'accroissement de la probabilité d'un asthme hospitalisé en moyenne peut être mesuré via une approche classique par groupe (est-ce que le groupe 1 est plus affecté que le groupe 2 selon les niveaux de vie l'état de santé à la naissance...) via une approche de *machine learning* (qui connaît un événement aigu lorsqu'il y a un surcroît de pollution de l'air et ne l'aurait pas connu sinon ? un groupe peut-il émerger à partir des caractéristiques individuelles ?). Cette deuxième approche suppose de disposer de nombreuses caractéristiques individuelles. La base de données EDP Santé permet de repérer des événements aigus en lien avec les maladies respiratoires via les données de consommation de soins et les bases de données INSEE (fiscales) permettent d'obtenir les caractéristiques socioéconomiques de l'enfant.

L'idée est d'appliquer des méthodes de *machine learning* sur la moitié de l'échantillon afin de faire émerger des prédictions pour estimer la vulnérabilité sans poser un trop grand nombre d'hypothèses initiales (approche non supervisée). La deuxième moitié de l'échantillon est alors utilisée pour confirmer les prédictions, en classant les effets selon leur niveau d'affection.

Quels sont les résultats ? Sur la base des données d'admissions aux urgences pour bronchiolites, l'effet de la pollution de l'air masquerait en réalité de fortes disparités : 10% des enfants concentrent l'essentiel des effets détectables statistiquement lors d'une augmentation de l'exposition avant leur premier anniversaire. Sur la base des données de consommation de médicaments contre l'asthme, l'effet est plus diffus, mais fait également ressortir un groupe d'enfants. Les enfants de ce groupe sont caractérisés par la concentration d'un certain nombre de facteurs de risques (prématurité, niveau de vie plus faible).

II) Geospatial AI for One Health

Jessica Abbate, épidémiologiste spécialisée en maladies infectieuses et analyse géospatiale à Geomatys

Le projet EpiWise, monté par la société française Geomatys en relation avec la communauté académique, vise à rendre les données géospatiales plus accessibles et opérationnelles dans le champ des maladies infectieuses émergentes, liées aux changements climatiques ou au contexte *One Health*. 60 à 75 % des maladies infectieuses émergentes sont zoonotiques. L'approche *One Health* consiste à considérer l'interaction entre l'environnement, l'animal et l'humain. L'organisation PREZODE vise à prédire les émergences de maladies infectieuses zoonotiques et la réunion des données issues de sources et de champs très variés est un réel défi technique, auquel Geomatys et son équipe EpiWise contribuent.

Il est probable que les maladies infectieuses sensibles au climat soient de plus en plus présentes en Europe, comme dans d'autres régions du monde, en lien avec le changement climatique. L'intégration de ces changements et leurs effets sur les dynamiques et des interactions des populations animales et l'homme sont nécessaires pour prévoir l'évolution des risques épidémiologiques. L'IA géospatiale fournit beaucoup d'informations à partir des données d'observation de la terre, qui permettent d'extraire des caractéristiques physiques, de réaliser des analyses spatiotemporelles multispectrales et également à partir d'informations contextuelles géolocalisées (LLM pour la surveillance d'événements et pour les facteurs de surveillance syndromique, culturels et socio-économiques).

Les modèles de risque épidémique sont basés sur les données de surveillance d'apparition de maladies, mais également sur l'adéquation de l'habitat de l'hôte et du vecteur, sur les interactions affectant la propagation de la maladie de l'animal à l'humain et sur la transmission intrahumaine. L'ensemble de ces données permet de cartographier des risques de maladie et de modéliser les scénarios afin de stratifier les risques de maladies et de lancer des alertes destinées aux décideurs. Les données spatiales (utilisation du terrain, occupation du sol, changement d'habitat, empiétement sur l'habitat, densités humaines, voies de communication) sont particulièrement utiles pour actualiser les modèles et extraire des informations.

L'exemple d'Ebola peut être généralisé à d'autres maladies (dengue, grippe aviaire, etc.). Les données relatives aux maladies émergentes sont limitées et il est nécessaire d'utiliser des modèles mécanistiques permettant notamment de savoir quels éléments sont réellement impactants. Les habitats des animaux hôtes naturels d'Ebola sont distincts et prédictibles. La compilation des différentes données avec un modèle de *machine learning* permet de prédire les risques de présence de ces hôtes. L'émergence d'Ebola est également étroitement corrélée aux perturbations de l'habitat forestier. Les modèles d'IA permettent d'identifier les perturbations forestières (incursions humaines, fragmentation de zones forestières) qui stressent les animaux, les obligent à se déplacer et ainsi contribuent à augmenter les contacts entre les hommes et les animaux. La possibilité d'introduction du virus à l'homme est plus grande. De nombreuses populations humaines confrontées au risque Ebola sont peu documentées (zones de conflits, nomadisme...), et les cartographies de densité obtenues par IA, à partir des images satellites, apportent de précieuses informations. Enfin, prédire la contagion et avoir une bonne connaissance des déplacements humains est essentiel pour la mise en place de plans d'action et de réponses.

Le [projet EpiWise](#) vise à mettre en place une plateforme d'intégration pour l'ensemble de ce type de données et permettre une télé-épidémiologie opérationnelle.

III) Amélioration de scores de risque environnemental par machine learning informé et IA explicable

Jean-Baptiste Guimbaud, docteur du Laboratoire d'informatique en images et systèmes d'information (LIRIS)

En épidémiologie, les deux principales méthodes actuellement utilisées pour calculer les scores de risque sur la santé (basées sur la littérature et sur les données) sont linéaires et limitées, car en incapacité de capturer des relations complexes entre les facteurs environnementaux et leurs impacts sur la santé. Le premier objectif de ma thèse a été d'utiliser des méthodes *machine learning* non paramétriques combinées à des méthodes d'IA explicable pour calculer ces scores de risque. Les méthodes non paramétriques ont l'avantage de ne faire aucune hypothèse sur la nature des relations capturées par l'IA, de capturer des relations complexes (relations non linéaires, effet cocktail entre groupes de variables) et d'avoir de meilleures performances prédictives. Toutefois, elles ne sont pas utilisables en l'état pour garantir la significativité statistique des relations capturées ni leur causalité.

Dans une première étude, cette approche basée sur l'IA a été utilisée pour calculer des scores de risques environnementaux pour trois axes de symptômes de maladies chroniques

(maladies respiratoires, cardiovasculaires et mentales). Les performances prédictives et le type de relations obtenues ont été comparés à ceux des approches traditionnelles. Des données d'une cohorte européenne mère-enfant ont été utilisées. Des indicateurs populationnels et individuels ont été obtenus pour déterminer quels facteurs environnementaux avaient le plus d'impact sur les scores de risque. Il est également possible de grouper les scores de risque par individus ou variables similaires. Cette méthode permet également de déterminer la nature de la relation entre un facteur environnemental et le risque estimé pour la santé. Les limites de cette approche sont tout d'abord liées à la taille de l'échantillon de données (le choix a été d'inclure des connaissances de la littérature) et ensuite au fait que certaines relations capturées ne sont pas en adéquation avec l'état actuel des connaissances (là encore, l'utilisation des connaissances de la littérature a guidé l'apprentissage du *machine learning*).

Dans une deuxième étude, une nouvelle approche a été développée pour intégrer les connaissances de la littérature et les ajouter à une base de données observationnelles du *machine learning* (données générées). Le score de risque ajusté doit présenter une meilleure performance prédictive et être en mesure de capturer des relations plus alignées avec le consensus scientifique et donc plus fiables.

Dans une troisième étude, cette approche a été validée sur de véritables données médicales issues du projet espagnol GCAT (données d'hypertension chez l'adulte). L'objectif était de comparer les performances prédictives et la causalité des relations capturées entre l'environnement et la santé entre notre approche de *machine learning* informée (données d'entrée + données de la littérature) et une approche de *machine learning* agnostique (données d'entrée). L'impact d'un facteur environnemental sur le score de risque associé pour la santé prédit par l'approche *machine learning* informée est beaucoup plus aligné avec le consensus scientifique que celui obtenu par *machine learning* agnostique. Notre approche permet également de corriger les variables pour lesquelles il n'existe pas d'information dans la littérature.

Cette approche offre une meilleure stabilité des résultats, permet de mieux démêler les effets et associations et enfin améliore les performances et la capacité à généraliser les prédictions notamment pour les sous-populations mal représentées dans les jeux de données lorsqu'ils sont limités. Néanmoins, ce modèle ne fournit pas de garantie sur la significativité des relations capturées ni sur la causalité. Enfin, il faut veiller à la robustesse des connaissances incorporées dans le processus d'apprentissage.

IV) L'IA en santé environnementale : cas d'étude avec la chaleur extrême

Jérémie Boudreault, INRS & INSPQ (Québec, Canada)

Les vagues de chaleur en 2003 ont provoqué 70 000 décès en Europe dont 15 000 en France, le Québec a connu d'importantes vagues de chaleur en 2010 et 2018. L'Institut national de recherche scientifique et l'Institut national de santé publique du Québec ont lancé un projet visant à utiliser des modèles d'IA afin de mieux comprendre les effets de la chaleur extrême sur la santé et développer des modèles prédictifs. Les objectifs étaient notamment de modéliser la mortalité lors de vagues de chaleur, définir les variables importantes pour développer des systèmes d'alerte précoce et enfin développer un modèle de surveillance régional. Les outils d'aide à la décision ont été construits sur la base de trois grandes catégories de modèles : les modèles basés sur les arbres (apprentissage automatique), les réseaux de neurones (apprentissage automatique et profond) et les modèles statistiques. Les approches non supervisées ou par renforcement n'ont pas été utilisées, mais seulement des approches supervisées.

Concernant la *prédiction de la mortalité lors des vagues de chaleur*, la modélisation des décès quotidiens toutes causes a concerné la région métropolitaine de Montréal de mai à septembre. Les prédicteurs retenus sont des variables de température, d'humidité, de

vitesse des vents (jusqu'à 7 jours avant la survenue des décès) et de pollution de l'air. Les trois catégories de modèles ont été utilisées. Pendant la période d'entraînement des modèles (1998 à 2013) les modèles complexes reproduisent très bien le pic de mortalité de 2010. Sur la période de validation 2014-2019, la vague de chaleur de 2018 ayant entraîné une surmortalité est bien visible, mais les modèles qui l'ont bien reproduite ne sont pas les mêmes (arbre décisionnel, modèles statistiques non linéaires). Si les modèles sont entraînés sur une plus longue période de 1981 à 2009 (les vagues de chaleur de 2010 et 2018 sont désormais dans la vague de validation), les performances des deux catégories de modèles (arbres et modèles statistiques non linéaires) sont relativement élevées. Les réseaux de neurones continuent à ne pas performer (volume insuffisant de données). Référence : J Boudreault, C Campagna, F Chebana (2023). Machine and deep learning for modelling heat-health relationships. *Science of the Total Environment*, 682(164660): 1-12.

Quelles variables sont importantes pour les systèmes d'alerte ? La modélisation de la mortalité toutes causes a été réalisée sur les régions métropolitaines de Montréal et Québec. Plutôt que d'inclure toutes les variables dans les modèles d'apprentissage automatique et profond, 17 groupes de prédicteurs ont été définis : les 9 premiers sur des variables de température (température moyenne, min et max, point de rosée et humidex) et les suivantes ont combiné la température moyenne et d'autres indicateurs (changements de température, météo, pollution de l'air...). Un modèle statistique non linéaire, deux modèles basés sur les arbres décisionnels et deux modèles de réseau de neurones ont été utilisés. Les groupes de prédicteurs incluant seulement la température avaient des performances en validation équivalente à celles obtenues avec des groupes de prédicteurs plus complexes. Pour Montréal, les indices de température présentant les meilleures performances sur la phase de validation sont l'humidex maximum, la température moyenne, le point de rosée maximum et l'humidex moyen. Pour Québec il s'agit de la température minimale, de l'humidex minimal, de l'humidex moyen et de la température moyenne. Référence : J Boudreault, C Campagna, F Chebana (2024). Revisiting the importance of temperature, weather and pollution variables in heat-mortality relationships with machine learning. *Environmental Science and Pollution Research*. 31 (9): 14059-14070.

Quels modèles de surveillance des impacts régionaux retenir pour englober 15 régions sanitaires du Québec ? 9 indicateurs de santé ont été examinés (mortalité, hospitalisations, visites aux urgences ambulances chez la population générale et 65+ ans...). Différents prédicteurs météorologiques, régionaux et temporels ont été utilisés, et 5 modèles statistiques et d'apprentissage automatique/profond ont été considérés, l'objectif étant de réaliser des prévisions à court terme des indicateurs de santé. Les modèles d'apprentissage profond (réseaux de neurones) ont le mieux performé pour une majorité d'indicateurs de santé. En majorité, les performances en phase de validation étaient supérieures à 60%. Les modèles d'apprentissage automatiques sont un outil prometteur pour modéliser les impacts de la chaleur sur la santé, mais le choix du meilleur modèle dépend de différents facteurs (mode de validation, critères de performance retenus, volume de données...). Il s'agit des premiers travaux de ce genre dans la littérature et ils sont sujets à amélioration. Ces approches peuvent aider à avoir une surveillance améliorée de la chaleur. Ces modèles peuvent être utilisés pour étudier les facteurs de vulnérabilité à la chaleur (prise de certains médicaments, îlots de chaleur...) et c'est ce sur quoi le groupe de recherche travaille actuellement (publications à venir)..



V) Échanges

Grégoire REY (Inserm, ancien directeur du CépiDc)

Quelle est la méthode utilisée pour évaluer la pertinence des données issues de la littérature et injectées dans le modèle de Jean-Baptiste Guimbaud ?

Jean-Baptiste GUIMBAUD

Nous avons utilisé des méta-analyses qui agrègent les résultats de plusieurs études, prennent en compte le niveau d'incertitude de chaque étude et les agrègent pour obtenir des jeux de données robustes. Dans son cas d'usage, l'utilisateur peut également avoir une idée du niveau de confiance associé : une méthode permet d'identifier la meilleure solution en fonction des besoins et des pondérations effectuées.

Modérateur

Milena Suarez Castillo, face aux nombreuses variables en jeu, avez-vous créé des classes d'individus ?

Milena SUAREZ CASTILLO

Nous avons construit des classes d'impact sur l'effet du traitement (ex. : les 10% les plus affectés par le traitement, les 50 % les moins affectés) et nous les avons comparées. Les covariables ne sont pas regroupées en classe et restent individuelles.

Modérateur

Jessica Abbate, est-il envisageable de faire une modélisation comme la vôtre sur des exemples français ? Est possible d'inclure des données sociales ?

Jessica ABBATE

Il existe effectivement des exemples de zoonose en France transmissibles à l'homme, comme la grippe aviaire, la toxoplasmose... J'ai travaillé sur les facteurs environnementaux et écologiques de la prévalence des pathogènes zoonotiques chez les petits mammifères en France. Il est possible d'obtenir des cartes de risque à partir de ces informations, intéressantes non seulement pour la santé humaine, mais aussi pour celle de la faune et de la flore. Pour les travaux sur Ebola et la dengue, nous essayons d'inclure les données sociales, le niveau de ressources, le statut vaccinal pour identifier les populations « résistantes » aux interventions locales et qui doivent être sensibilisées.

RESSOURCES & REFERENCES

1. Milena Suarez Castillo (DREES), avec la collaboration de Vianney Costemalle (DREES), David Benatia (CREST) et Christine le Thi (Insee), Plus exposés à la pollution de l'air, les jeunes enfants des ménages modestes, plus fragiles, sont les plus affectés : <https://drees.solidarites-sante.gouv.fr/sites/default/files/2024-01/ER1292MAJ2.pdf>
2. Guimbaud, JB., Siskos, A.P., Sakhi, A.K. et al. Machine learning-based health environmental-clinical risk scores in European children. *Commun Med* 4, 98 (2024). <https://doi.org/10.1038/s43856-024-00513-y>
3. Projet EPIWISE : <https://www.epiwise.com/>
4. J Boudreault*, A Ruf, C Campagna, F Chebana (2024). Multi-region models built with machine and deep learning for predicting several heat-related health outcomes. *Sustainable Cities and Society*. 115(105785): 1-12.



Session 3 : L'IA pour répondre aux enjeux de restitution des résultats scientifiques et perceptions

Theo Alvez Da Costa ONG Data for Good, Modérateur

I) L'IA générative au service de la démocratisation de l'accès aux données

David Thoumas, cofondateur et directeur technique d'Opendatasoft

Opendatasoft aide ses clients des secteurs publics et privés à démocratiser l'accès à leurs données et construire des portails de données pour les diffuser au plus grand nombre. Opendatasoft a par exemple travaillé sur l'accessibilité des données synthétiques et intelligibles de vaccination contre le Covid-19 du portail Ameli. Chaque jeu de métadonnées est traité par des fonctions de recherche et de filtrage, visualisé, exporté, intégré à une application. L'utilisateur peut alors faire du *data story telling* pour les rendre accessibles.

Opendatasoft peut intégrer tous types de données à un portail (développement de connecteurs et de fonctions de normalisation, diverses sources de données, enrichissements de données issues de diverses sources), sans toutefois de transformation significative de la donnée, qui était jusqu'à présent réalisée en amont ou en aval par les utilisateurs. Avec la démocratisation de l'IA, Opendatasoft s'est demandé comment l'utiliser pour démocratiser l'accès à la donnée.

Les capacités d'interactivité et d'interprétation de l'IA sont des atouts et l'IA est un élément clé pour simplifier l'utilisation des portails qui accueillent une grande variété de données. Opendatasoft a réutilisé des fonctions proposées par les opérateurs de LLM pour construire des fonctionnalités permettant d'enrichir la plateforme. Ainsi, les plateformes proposent une recherche intelligente (versus lexicale) et des recommandations de données similaires via des modèles *d'embedding* et une exploration interactive de données via des modèles d'IA générative.

Les bénéfices sont à la fois pour les utilisateurs finaux (sans connaître l'intégralité du glossaire métier, ils peuvent trouver des jeux de métiers et interagir avec la donnée), mais aussi pour les producteurs data et les administrateurs de portail (augmenter l'adoption du portail à l'échelle et valoriser le remplissage et la contribution des métadonnées).

Opendatasoft a commencé à travailler avec Open AI en 2023 et plus récemment avec Mistral AI. Les deux approches sont proposées dans les plateformes, avec plusieurs modèles. Un objectif est de permettre aux organisations utilisatrices de la plateforme de brancher leur propre modèle IA.

Démonstration réalisée sur le portail du ministère de l'Économie et des Finances.

II) Constituer un écosystème et mutualiser les corpus d'expertise pour élaborer une IA publique de confiance sur la transition écologique

Philippe Guillouzic, Aderme

La découverte et l'expérience de l'utilisation de Climate Q&A (<https://www.climateqa.com/>) sont à l'origine du projet SofIA de l'ADEME. J'ai découvert Climate Q&A qui est une solution qui permet de saisir simplement sa question liée au climat. ClimateQ&A vous fournira une réponse avec des références aux rapports pertinents du GIEC. ClimateQ&A récupère des passages spécifiques des rapports du GIEC pour vous

aider à répondre avec précision à votre question. Les informations sur la source, y compris les numéros de page et les passages, sont affichées sur le côté droit de l'écran.

L'ADEME produit une quantité importante d'expertises et de connaissances et il est parfois difficile de trouver simplement et rapidement la bonne information. L'idée a simplement été d'expérimenter à l'ADEME une solution équivalente à Climate Q&A. Pour mettre en place le projet SOFIA à l'ADEME, j'ai contacté Ekimetrics qui a développé le projet Climate Q&A.

La volonté a été rapidement de créer ce projet en écosystème en associant des acteurs de la transition écologique. Ainsi, l'ADEME, avec le Cerema et le CGDD, via EcoLab, et accompagnés d'Ekimetrics, nous développons, SofIA, une intelligence artificielle (IA) publique de confiance pour faciliter la diffusion et l'usage de leurs ressources documentaires mutualisées. Producteurs de nombreuses connaissances sur les thématiques de la transition écologique qui visent à accompagner la mise en œuvre des politiques publiques, nous avons souhaité mutualiser nos ressources textuelles de référence pour développer un agent conversationnel. L'utilisateur peut formuler des requêtes en langage naturel et interroger le ou les corpus documentaires de son choix pour identifier, trouver, synthétiser des connaissances et obtenir des informations générées grâce à des moteurs de recherche intelligents.

La mise en commun et la mutualisation des corpus de données conjuguées à la puissance des modèles de langage (LLM) permettent de dialoguer avec une diversité de gisements de données textes, fiables, souveraines et régulièrement actualisées pour accéder à des informations ciblées et qualifiées.

Le projet SofIA ambitionne à court terme d'élargir cette mutualisation à d'autres acteurs publics de référence qui produisent des connaissances techniques et visent à accompagner la mise en œuvre des politiques publiques dans le domaine de la transition écologique.

Parallèlement en interne ADEME, nous avons ouvert un appel à cas d'usage, qui embarquerait des solutions d'intelligence artificielle, via notre plateforme interne d'open Innovation. Une vingtaine de projets est remontée, avec notamment un besoin de pouvoir interroger le stock d'expertise. Cet appel à cas d'usage nous a permis de structurer une feuille de route IA, organisée en 4 piliers :

- Système d'IA basé sur des corpus documentaires mutualisés sur la transition écologique (projet SOFIA) ;
- Data : exploiter l'IA sur des données structurées sur l'ensemble du spectre de la transition écologique (diagnostics de performance énergétique, qualité de l'air...) pour faire émerger des cas d'usage, faire de la prédiction, du monitoring ;
- Bureautique et pratiques professionnelles : IA pour aider à l'amélioration de tâches quotidiennes, automatiser un certain nombre de tâches (assistance utilisateur...) : un projet est en cours pour savoir si l'IA peut intervenir sur des études réalisées par l'ADEME (ex. : assistance à la rédaction de cahier des charges, analyse des projets reçus...) et réinternaliser des tâches aujourd'hui majoritairement externalisées ;
- Aides et services : utilisation de l'IA pour participer à la génération, à l'étude et au financement d'appels à projets liés à la transition écologique ; pour fluidifier le parcours utilisateur, faciliter l'accès aux aides et le traitement des données

III) Perception et acceptabilité de l'Intelligence Artificielle : l'apport des sciences comportementales

Paul Grignon, Ecolab (CGDD/SRI)

La perception, l'adoption et l'utilisation de l'IA sous-tendent une dimension intrinsèquement humaine. Ainsi, quel est l'apport des sciences comportementales pour bien se saisir des enjeux de perception et d'acceptabilité de l'IA ?

Les sciences comportementales intègrent un ensemble de champs disciplinaires permettant d'apprécier l'étude des comportements individuels et collectifs (sciences cognitives, économie comportementale, psychologie sociale, sociologie, neurosciences, anthropologie). De nombreux concepts, notions et méthodes sont mobilisés.

Les sciences comportementales peuvent notamment favoriser la prise en compte des enjeux éthiques de l'IA et le rapprochement de l'IA et de l'humain : compréhension des enjeux de perception et d'acceptabilité, analyse des conséquences d'utilisation de l'IA, conception centrée sur les individus, ou encore prise en compte des besoins et des valeurs des individus et relatives à l'éthique dans le développement et l'utilisation des outils d'IA.

Comment étudier les enjeux de perception et d'acceptabilité concernant l'usage de l'IA ? Plusieurs concepts s'enchaînent : la perception (compréhension et interprétation), l'acceptabilité (évaluation anticipée), l'acceptation (évaluation *ex-post*) puis l'intention d'usage (disposition à utiliser). Plusieurs modèles ont été développés pour comprendre les mécanismes sous-jacents à l'acceptabilité de la technologie (IA). Les deux principaux mécanismes sont son utilité perçue et la perception de facilité d'utilisation (effort attendu), auxquelles peuvent s'ajouter la prise en compte de l'influence sociale et des conditions facilitantes. Il peut également être pertinent de prendre en compte l'hétérogénéité individuelle en introduisant des variables comme l'âge, le genre, le niveau éducatif. Certains modèles prennent également en compte d'autres facteurs tels que la motivation hédonique, l'habitude ou encore la valeur prix (modèle UTAUT 2) pour étudier l'intention comportementale d'utiliser une technologie. Au-delà des facteurs psychosociaux et individuels, il est nécessaire de tenir compte des facteurs culturels (une étude montre qu'en France, la principale émotion ressentie à l'évocation de l'IA est l'inquiétude alors qu'en Inde, son évocation est associée à une impatience et à un certain espoir), mais aussi des facteurs contextuels selon le domaine concerné.

Il est important de vérifier la déclinaison opérationnelle des questionnaires psychométriques basés sur ces modèles au regard du contexte de chaque étude. Des enquêtes qualitatives peuvent être éclairantes pour mettre en œuvre la déclinaison opérationnelle des outils transverses développés, et les adapter à chaque étude spécifique..

Que dit la recherche sur les enjeux de perception et d'acceptabilité de l'IA ? La majorité des articles ne définit pas explicitement l'IA, ce qui rend difficile la comparaison et la généralisation des résultats obtenus. Un article propose 4 grands clusters de type d'IA selon deux dimensions : l'autonomie de l'IA et le degré d'implication humaine dans l'IA. Le niveau d'acceptabilité diffère selon l'IA considérée. Il existe également une corrélation positive entre l'acceptation de l'IA et la compréhension de son fonctionnement et de la manière de l'utiliser (littératie). Enfin, une consultation citoyenne menée en 2024 par Ekitia a souligné l'importance de dimensions éthiques de l'IA (finalité et impartialité) comme facteurs clés d'acceptation.

Le manque de connaissances concernant l'IA peut conduire à une distorsion et à une défiance ou à une confiance pouvant être jugée comme excessive. Comment corriger les fausses croyances ? Il faut tenir compte des biais de confirmation et de l'heuristique de disponibilité. Cette dernière implique une réaction de conservatisme et de biais de récence. Pour lever ces fausses croyances, plusieurs pistes existent : engager les individus dans des tâches qui les obligent à évaluer les preuves, identifier des erreurs logiques et amener les individus à réfléchir sur leur processus de pensée. En conclusion, les sciences comportementales peuvent être pertinentes sur les enjeux de transparence et d'explicabilité de l'IA également.



IV) Échanges

Théo ALVES DA COSTA

La dernière présentation fait-elle écho à l'acceptabilité des fonctions d'IA générative chez Opendatasoft ou l'ADEME ?

Philippe GUILLOUZIC

Nous n'avons pas encore adressé la question de l'acceptabilité à l'ADEME, mais nous avons associé le terme « confiance » à l'IA publique et ajouté des notions d'éthique dans la gouvernance. La démarche est progressive et opérationnelle, pour répondre aux besoins des utilisateurs. SOFIA ne traite pas de sujets moraux (discrimination dans l'attribution des aides...). Le contexte culturel influe fortement sur la manière d'appréhender ces technologies (Cf. l'expérience de the Moral Machine).

Table Ronde : Futur de l'IA en santé environnement - quels enjeux juridiques, éthiques et environnementaux

Régis Chatelier, CNIL

Florent Martin, Ecolab (CGDD/SRI)

Philippe Guillouzic, Ademe

Théo Alves Da Costa, Data for Good

Sébastien Denys modérateur



Messages clés de la table ronde :

- Pour rester dans la course à l'IA, la France peut choisir de travailler sur d'autres enjeux que la performance que sont l'éthique, l'environnement, le respect de la vie privée.
- L'usage des outils d'Intelligence Artificielle doit être questionné tout au long du projet, dans les réflexions en amont mais aussi sur ses conséquences en aval : l'IA est-elle la seule solution pour répondre au besoin, les bonnes pratiques ont-elles été mises en œuvre (maîtrise du volume de données, estimation de la consommation du modèle choisi, optimiser l'usage de l'équipement...), l'usage s'inscrit-il dans les limites planétaires ?
- Le RGPD s'applique aussi aux données personnelles traitées pour ou par les systèmes d'IA. Le Règlement européen sur l'intelligence artificielle (RIA) le complète notamment lorsqu'un système d'IA à haut risque nécessite des données personnelles pour son développement ou dans son déploiement. Les risques ne sont pas nouveaux mais peuvent être amplifiés avec l'usage des technologies d'IA.
- L'acculturation à l'IA est nécessaire ainsi qu'aux outils disponibles au sein des organisations (SOFIA Ademe).
- Il est primordial de :
 - o penser la mutualisation des données, des compétences, des connaissances, des architectures techniques, des puissances de calcul
 - o mettre en place les bonnes pratiques pour mesurer, contrôler voire réduire les biais

Sébastien DENYS

L'objectif de cette table ronde est d'aborder les enjeux éthiques, réglementaires, de frugalité, qui sont structurants afin de donner des points de repère aux institutions telles que Santé publique France et ses partenaires. J'accueille :

- Théo Alvez Da Costa : data scientist et ingénieur IA chez Ekimetricx, coprésident de l'ONG Data for Good, qui rassemble 6 000 bénévoles dans les métiers de la data et de l'IA qui mettent leurs compétences au service d'ONG et de services publics
- Régis Chatelier : chargé d'études d'innovation et prospectives au sein du laboratoire d'innovation numérique de la CNIL, et travaille en proximité avec le récent service de l'IA de la CNIL
- Philippe Guillouzic : responsable de l'IA à l'ADEME

- Florent Martin : chef du bureau « InnovationDATA/IA » d'Ecolab (laboratoire d'innovation du Commissariat général au développement durable)

Ces enjeux sont particulièrement importants, s'agissant de créer des liens entre les données environnementales et les données sensibles (données de santé).

Théo Alves Da Costa, vous avez initié « un livre blanc sur l'IA générative » et contribué à l'élaboration d'un serment d'Hippocrate pour les *data scientists*. Comment mettre en œuvre une IA de confiance dans le sens de l'intérêt général tout en respectant ces différents enjeux ?

Théo ALVES DA COSTA

Le serment d'Hippocrate du Data Scientist a été créé en 2018, alors que les sujets d'IA éthique et de confiance commençaient à émerger de la société civile (présence de biais, manque de vigilance sur les données personnelles, explicabilité...). Cette charte regroupe ces différents éléments et a été signée par de nombreux professionnels du secteur. Ce serment a été représentatif des enjeux jusqu'au début des années 2020 sur l'IA prédictive, le *machine learning*... qui posait des questions éthiques en fonction des cas d'application. L'arrivée de Chat GPT dans le grand public et la popularisation de l'IA générative en 2021 a fait émerger le besoin de faire évoluer ce serment. 150 contributeurs ont participé et ont identifié de nouveaux enjeux éthiques :

- Les biais sur l'IA générative, notamment dans la génération d'images ;
- Les sujets d'authenticité et de droit : clonage de voix, cybersécurité... ;
- Les enjeux environnementaux : une étude récente de Deloitte a chiffré les impacts environnementaux globaux de l'IA et montré que les data center consomment 2% de l'électricité mondiale, l'IA représente 10% des data centers aujourd'hui et devrait monter à 50 % dans les prochaines années : globalement, les émissions des data centers devraient passer de 10 à 100 millions de tonnes de CO_{2,eq} ;
- L'« accélérationisme » débridé et la concentration des pouvoirs, qui peuvent entraîner une perte de possibilité d'y naviguer.

Sébastien DENYS

Le Comité consultatif national d'Éthique soulignait déjà en 2018 la nécessité de réévaluer les enjeux éthiques dans le cadre de l'utilisation des données massives par l'IA.

En matière de données de santé, le RGPD impose déjà beaucoup de précautions pour protéger la confidentialité des données de santé. Régis Chatelier, pouvez-vous expliquer ce que va amener de plus le règlement européen sur l'IA publié cet été ? Va-t-il alourdir les procédures déjà conséquentes qui ont été mises en place par les instituts de santé publique qui utiliseraient des outils d'IA, voire les freiner ?

Régis CHATELIER

Le niveau de départ des agences de santé est plus élevé que celui d'autres secteurs, dans la mesure où elles traitent déjà de données de santé. La CNIL examine des dossiers à base d'IA depuis de nombreuses années. Son rapport sur l'éthique des algorithmes et de l'IA publié en 2017 traite notamment de la difficulté à qualifier un biais, souvent produit par l'usage de données personnelles. Le Défenseur des Droits a récemment publié un rapport sur les algorithmes publics. L'IA est un champ technologique qui fait l'objet de différents usages, qui ne sont par nature pas nécessairement contradictoires avec la protection des données dès lors que leur mise en place respecte le principe de la protection des données. La nécessité de contrôler l'usage de nos données, de les connaître et de les gérer sera d'autant plus importante que nous aurons recours à des algorithmes dont le fonctionnement peut être obscur. La gestion des données en entrée et en sortie s'appliquera. La CNIL, qui traitait essentiellement le sujet de l'IA au cas par cas, produit désormais des

recommandations plus générales : 8 fiches pratiques IA ont été publiées et 5 autres le seront prochainement (<https://www.cnil.fr/fr/les-fiches-pratiques-ia>). Les sujets traités sont notamment la maîtrise des grandes bases d'apprentissage, les données personnelles dans les modèles, la maîtrise des usages, accompagner et sensibiliser les personnes qui alimentent un modèle avec leurs données personnelles, une proposition de création d'un registre des acteurs ayant eu recours à de grandes bases de données pour l'apprentissage de modèles pour produire de la transparence et corriger l'existant. La logique du RIA est différente de la nôtre et consiste à traiter certains types d'IA selon leur niveau de risque (logique d'AMM). La CNIL s'intéressera aux modèles d'IA traitant des données personnelles et d'autres acteurs pourront également intervenir. Dès lors qu'il ne s'agit pas de prise de décision pour les personnes, les niveaux d'autorisation appliqués aux systèmes d'IA plus traditionnels qui seront mobilisés seront moins restrictifs. Les IA impliquant des risques de biais ou de prise de décision sur la personne feront probablement l'objet d'AMM, avec des analyses de risque sur les droits fondamentaux des personnes. Être en mesure d'appliquer correctement le RGPD est déjà un bon point de départ pour l'application du RIA.

Sébastien DENYS

L'ADEME a un temps d'avance sur la mise en œuvre opérationnelle des outils d'IA. Comment a-t-elle été confrontée concrètement à ces enjeux et a-t-elle été confrontée à une réticence interne au regard des enjeux éthiques relatifs à la consommation de ressources ?

Philippe GUILLOUZIC

Le système SOFIA n'est pas concerné aujourd'hui par des problématiques éthiques. En interne, l'impact environnemental de la solution a été rapidement évoqué et le dispositif permet de visualiser l'équivalent CO₂ généré par une requête. Ce genre de dispositif peut être coûteux (nombre de requêtes suscitées par l'écosystème, les coûts au sein de l'écosystème doivent-ils être mutualisés, individualisés... ?). Ces questions doivent être anticipées. La gestion des aspects juridiques est également questionnée. Il est nécessaire de savoir quelles conditions juridiques permettent de mettre les documents dans des bases d'IA générative, quel est le statut juridique des productions générées à partir de plusieurs documents issus de plusieurs acteurs ? Ces éléments doivent être clairement précisés dans les contrats de partenariat. Une charte de coopération entre les membres d'un écosystème ne suffit pas, car les membres sont demandeurs de clauses spécifiques sur le statut juridique des documents et des données. Enfin, il est nécessaire de trouver le juste compromis entre environnement, performance, fiabilité, souveraineté et sécurité. SOFIA utilise le dispositif Mistral IA, mais ne peut élaborer un système d'IA 100 % souverain. Lorsque l'on fait appel à des LLM non souverains, il est nécessaire de savoir où partent les données, ce qu'elles deviennent, si elles sont utilisées comme données d'entraînement... Il faut anticiper le système IA dans sa globalité.

Sébastien DENYS

Cela nécessite donc de prévoir une formation de l'ensemble des collaborateurs.

Philippe GUILLOUZIC

Dans les usages personnels, on a souvent tendance à faire d'embrée confiance à des systèmes d'IA (ex. : Waze). Dans l'environnement professionnel de l'ADEME, les personnes ont un esprit critique qui permet de challenger le dispositif et donc d'évaluer la performance de notre solution pour améliorer la pertinence des réponses. Nous diffusons des campagnes régulières d'analyse des réponses proposées par le système, avec une traçabilité de cette évaluation, ce qui permet d'évaluer la pertinence, la redondance, les hors-sujet...) et par la même de renforcer l'acceptabilité et la crédibilité du système.

Il existe de nombreux modules de formation clés en main, mais ils doivent être personnalisés. L'individu doit comprendre ce que sont les différentes facettes de l'IA, aller

jusqu'au questionnement de l'éthique en IA, et mobiliser les mesures en capacité de corriger les erreurs. L'éthique en IA doit s'appliquer sur l'ensemble du périmètre, amont, mais aussi aval...

Sébastien DENYS

D'où la nécessité de toujours avoir un regard humain sur les travaux, ce qui permet de parler d'IA de confiance.

Ecolab s'est beaucoup questionné sur la frugalité de l'IA en lien avec son écosystème et a élaboré avec l'AFNOR un référentiel sur la frugalité de l'IA, qui a vocation à être diffusé à l'international. Comment la France a-t-elle été pionnière sur ces questions ?

Florent MARTIN

Ecolab porte la feuille de route publique « IA et Transition écologique »¹, qui estime que l'on ne peut pas laisser le développement de l'IA « s'emballer » sans anticiper et évaluer son impact environnemental. Dans le cadre de sa convention avec l'AFNOR, le CGDD a rassemblé près de 150 acteurs de l'écosystème pour réaliser un AFNOR SPEC (document volontaire, préalable à une norme). Ainsi, le premier référentiel général sur l'IA frugale a été publié en juin 2024² pour outiller les acteurs à l'évaluation des impacts du recours à l'IA.

L'IA frugale doit respecter 3 grands principes :

- S'assurer qu'il n'existe pas une meilleure solution connue que l'IA pour tenter de répondre à un besoin ;
- Mettre en place des bonnes pratiques sur la chaîne de vie du projet ;
- Incrire la démarche (usage et moyens déployés) dans les limites planétaires.

La question de la frugalité peut être intégrée à chaque étape d'un projet, ce qui nécessite une formation et l'acquisition de compétences pour permettre aux gens de comprendre les enjeux. Il faut qualifier le besoin en amont, optimiser les données dans le projet quand cela est possible, bien choisir un modèle en fonction du besoin (certains sont performants et moins énergivores que d'autres), évaluer l'impact et tenir compte des enjeux de gouvernance.

Sébastien DENYS

Cela implique d'avoir conscience de ces enjeux en amont et souligne l'importance de sensibiliser, former, s'entourer de personnes compétentes pour appréhender ces dimensions.

Quels éléments clés de recommandations et de points de vigilance pourriez-vous adresser à la Directrice générale de Santé publique France ?

Théo ALVES DA COSTA

Il faut veiller à bien définir ce que l'on entend par IA : il existe des différences fondamentales entre les nombreuses familles d'IA (celles qui aident à répondre aux questions climatiques ne sont pas celles comme *MidJourney* qui permettent de générer des images). Il ne faut également pas oublier la dimension systémique de l'IA, qui est non seulement un système technologique, mais également un système économique. On peut s'attrister de voir que la consommation électrique de grandes entreprises privées soit plus importante que celle de certains pays, que 30 % des achats d'énergies renouvelables ont été réalisés par les Big Tech en 2024 et que certaines centrales nucléaires sont dédiées à l'alimentation électrique de data center en trichant sur les normes de comptabilité

d'émissions carbone... L'infrastructure globale de l'IA est de plus en plus aux mains d'acteurs privés, ce qui impacte les avancées réalisées.

Régis CHATELIER

Les systèmes d'IA peuvent être utilisés *by design* (se poser la question des objectifs et des moyens mobilisés dès le début du projet). La CNIL a publié en 2023 un cahier Innovations et Prospectives sur les intersections entre protection des données et environnement, traitant des sujets de l'IA et de l'écoconception. La première question posée par le référentiel d'écoconception des services numériques est de déterminer si ce service est nécessaire.... Le RGPD a été un véritable outil d'apprentissage et d'éducation au numérique et à la donnée au sein des organisations. Après une phase d'apprentissage sur les données personnelles et les droits des personnes, nous passons à une phase d'apprentissage sur ce que sont les systèmes d'IA et comment les mettre en place.

Philippe GUILLOUZIC

Il faut développer la capacité de discernement des gens, revenir aux besoins métier et examiner si l'IA peut y contribuer ou non (vigilance sur les effets rebonds et les impacts environnementaux). L'IA a une capacité à développer de nouveaux imaginaires et adopter une approche différente. Enfin, la mutualisation des données, des compétences, des connaissances, des architectures techniques, des puissances de calcul est une force et représente des gisements de performances.

Sébastien DENYS

D'autant plus que l'Agence interagit avec de nombreux acteurs de l'écosystème.

Florent MARTIN

Pour rester dans la course à l'IA, la France peut choisir de travailler sur d'autres sujets que la course à la performance (éthique, environnement, respect de la vie privée). Les acteurs qui sont présents aujourd'hui sont intéressés : l'enjeu est d'informer ceux qui s'y intéressent peu pour creuser le sujet et leur donner un premier niveau d'informations, à l'occasion d'évènements pédagogique sur l'IA, comme les Cafés IA par exemple.



Échanges

Un participant

Comment proposez-vous d'entraîner des IA en s'affranchissant des biais éthiques liés aux données historiques et à la génération de données d'entraînement artificielles ?

Théo ALVES DA COSTA

Cela dépend du type de données (textes, images, données structurées...) : tous les types d'algorithmes permettent de réduire les biais, de même que certaines techniques organisationnelles (choix des genres, âges, origines ethniques) ou de collecte de données peuvent réduire des biais. Il faut en avoir conscience et mettre en place les bonnes pratiques pour mesurer, contrôler voire réduire les biais. Stanford réalise chaque année un classement des LLM et algorithmes sur leur respect d'un certain nombre de principes éthiques.

Jessica ABBATE

Les compétences mathématiques reconnues de la France offrent des opportunités de développement de différents modèles d'IA et peuvent changer le paysage international. Existe-t-il une organisation internationale qui essaie de regrouper les initiatives d'IA frugale ?

Florent MARTIN

La France souhaite être cheffe de file du travail international de normalisation sur l'impact environnemental de l'IA. Le sommet mondial de l'IA est organisé en France en février 2025 et l'objectif est de travailler sur une feuille de route commune de normalisation de l'impact environnemental de l'IA.

Nous espérons que ces règles permettront aux entreprises de considérer les critères environnementaux, éthiques... et pas uniquement la performance. Ce faisant, le classement des entreprises pourra évoluer.

Un participant

Que fait l'IA pour les personnes précaires, dont la principale préoccupation est de manger, et qui font passer leur santé et les préoccupations environnementales au second plan ?

Régis CHATELIER

L'IA ne fait rien : c'est la manière dont on l'utilise qui a des conséquences.

Théo ALVES DA COSTA

On peut classer les IA en 4 grandes catégories : les *gods*, les assistants, les engrenages et les jouets. La première et la dernière catégorie ne sont pas utiles aux populations défavorisées. Les deux catégories intermédiaires d'IA sont intéressantes. Data for Good a par exemple travaillé avec des ONG pour identifier dans les PDFs d'entreprises des montants d'évasion fiscale sur la comptabilité fiscale multi-pays, et permettre à ces ONG de mettre en place des projets pour lutter contre les inégalités.

Milena SUAREZ CASTILLO

Quels conseils donner à la Direction générale pour sa stratégie d'IA en santé ?

Régis CHATELIER

Les problématiques sur le traitement des données en santé sont bien identifiées et traitées, mais il ne faut pas s'en remettre entièrement à des systèmes sans les questionner sur toute la chaîne. Les risques ne sont donc pas nouveaux, mais peuvent être amplifiés si ces technologies ne sont pas abordées avec les précautions nécessaires. Dans certains projets de recherche en santé, le manque d'accès à des données européennes oblige à aller rechercher des bases de données constituées ailleurs, ce qui peut amener des biais. Vérifier la qualité des données d'entrées est essentiel.

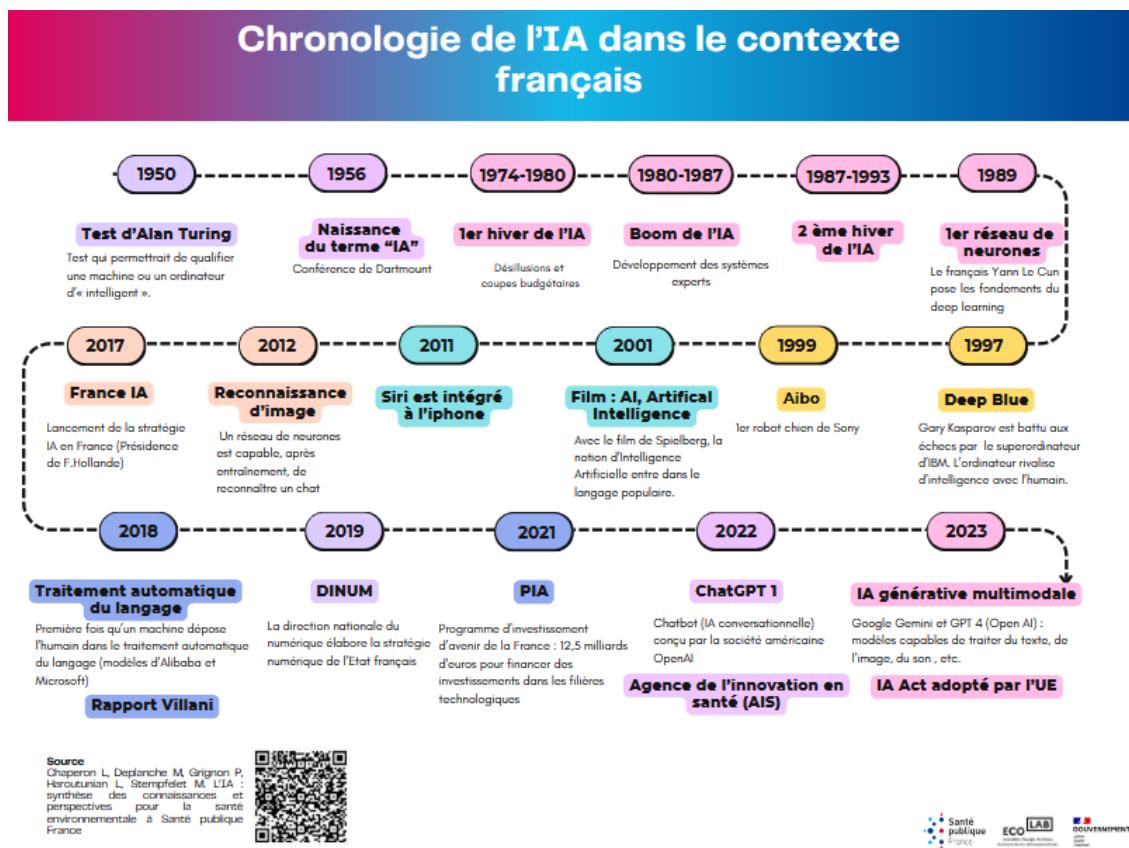
Théo ALVES DA COSTA

Il existe assez peu d'études sur l'utilité de l'IA générative sur les questions de santé et de santé publique. Les risques des nouveaux enjeux sont peu chiffrés : un nombre croissant de jeunes utilisent chat GPT comme un psychologue, les hallucinations de chat GPT utilisé comme coach de santé posent question... Des opportunités intéressantes existent sur les parcours de santé, l'accès aux données de santé publique...

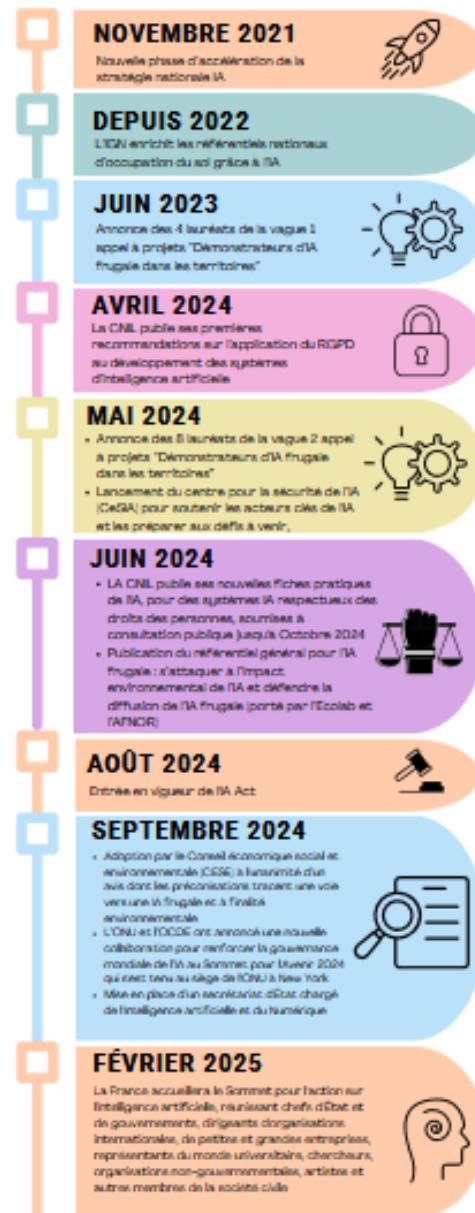
RESSOURCES & REFERENCES

1. Etude Deloitte : “As generative AI asks for more power, data centers seek more reliable, cleaner energy solutions” :
<https://www2.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions/2025/genai-power-consumption-creates-need-for-more-sustainable-data-centers.html>
2. Les fiches pratiques IA de la CNIL : <https://www.cnil.fr/fr/les-fiches-pratiques-ia>
3. Entrée en vigueur du règlement sur l'IA :
https://commission.europa.eu/news/ai-act-enters-force-2024-08-01_fr
4. Publication du référentiel général pour l'IA frugale : s'attaquer à l'impact environnemental de l'IA et défendre la diffusion de l'IA frugale :
<https://www.ecologie.gouv.fr/presse/publication-du-referentiel-general-ia-frugale-sattaquer-limpact-environnemental-ia>
5. Sommet pour l'action sur l'IA : <https://www.elysee.fr/sommet-pour-l-action-sur-l-ia>
6. Le livre blanc « Les grands défis de l'IA générative » :
<https://dataforgood.fr/iagenerative>

Annexes



QUELQUES ACTU DANS LE MONDE DE L'IA



Santé publique France

ECO LAB

ÉCOLE NATIONALE
DES SCIENCES

EXEMPLES D'OUTILS UTILISANT L'IA

S.A.R.A.H.

Sarah est le prototype d'une agente numérique chargée de promouvoir la santé, disponible 24 heures sur 24 et 7 jours sur 7, en huit langues, au format vidéo ou texte. Elle peut vous donner des conseils et vous donner des informations sur plusieurs autres sujets de santé.



CLIMATE Q&A



ClimateQ&A exploite les techniques modernes de reconnaissance optique de caractères pour analyser et prétraiter les rapports du GIEC. ClimateQ&A est en mesure de passer au crible la vaste collection de rapports scientifiques sur le climat et d'identifier les passages pertinents en réponse aux demandes des utilisateurs.

PIXTRAL 12B

La licorne française MISTRAL a lancé Pixtral 12B. Qu'il s'agisse d'URL ou d'images encodées en base64 (représentation textuelle d'une image), Pixtral 12B est capable d'analyser à la fois du texte et des images.



Pixtral 12B



L'IA générative s'immisce de plus en plus dans nos vies à travers des outils capables de générer des contenus de qualité quasi-humaine et à une vitesse inégalée. Il faut questionner l'impact de cette révolution technologique sur notre société. Il n'est plus seulement question de savoir comment exploiter ces outils, mais aussi et surtout pourquoi.

LIVRE BLANC IA GÉNÉRATIVE

ECOLOGITS

EcoLogits suit la consommation d'énergie et les impacts environnementaux de l'utilisation de modèles d'IA générative par le biais d'API. Il prend en charge les principaux fournisseurs de LLM tels que OpenAI, Anthropic, Mistral AI et d'autres encore.



EcoLogits

Consensus



CONSENSUS

Consensus est un moteur de recherche académique alimenté par l'IA et nous avons pour mission de rendre les meilleures connaissances du monde plus accessibles.

BATAILLE DE L'IA

En s'appuyant sur le rapport de Data for Good sur l'IA Générative, la Bataille de l'IA vous permet de découvrir de manière ludique les enjeux sociaux et environnementaux de l'IA. L'atelier est pensé pour générer la discussion par petits groupes de 5 personnes, une bonne manière de créer du lien entre participants et participantes.



La
Bataille
de l'IA



