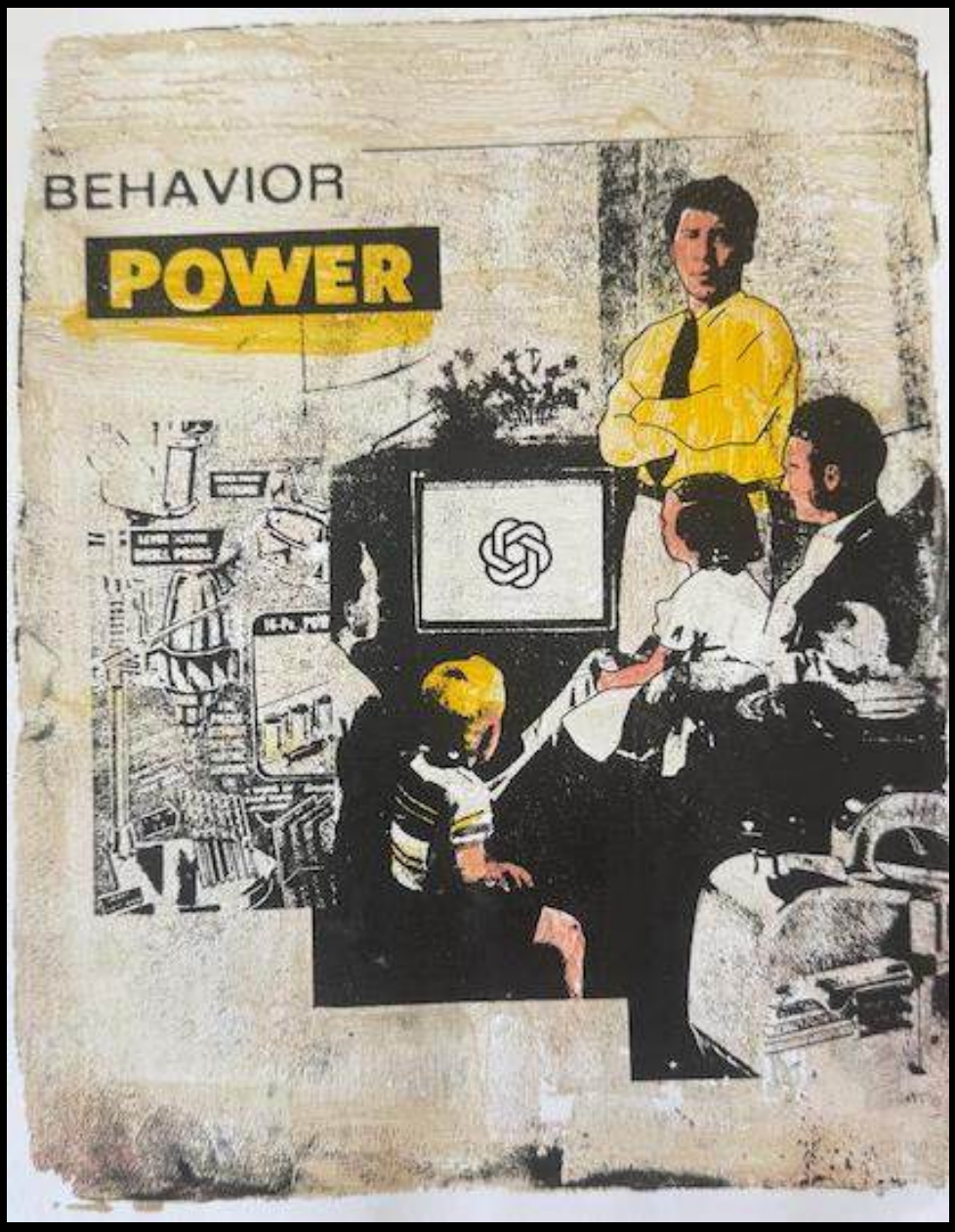
Comment expliquer les algorithmes ?

UNE GALERIE DE RESSOURCES DOCUMENTÉES PAR LE PROJET ALGO→LIT



Algo → Lit
Algorithmic literacy for effective transparency in the eu

Financé par l'Union européenne

Introduction

Ce guide présente une typologie d'activités et de projets visant à expliquer les algorithmes. Il constitue une ressource pour les médiateurs numériques, les chercheurs, les défenseurs des droits, les militants ou tout professionnel qui a le souhait de présenter le fonctionnement des algorithmes d'une manière attrayante. Il convient de noter que notre projet s'adresse principalement aux communautés touchées par les algorithmes en France, en Belgique et aux Pays-Bas. La sélection d'exemples n'est pas exhaustive, mais reflète un équilibre entre de nombreux critères :

— Nous avons veillé à présenter une diversité de formats - allant des jeux sérieux aux ateliers, des affiches imprimées aux formes collectives de prise de décision — issus de pays tels que la France, la Belgique, les Pays-Bas, l'Espagne, les États-Unis et le Royaume-Uni,
— Nous avons souhaité mettre en avant des exemples reconnus pour la qualité de leur contenu et de leur documentation, ainsi que pour leur engagement en faveur d'une approche critique des algorithmes qui échappe au « greenwashing » éthique actuel,
— Nous avons choisi de publier certains exemples longtemps négligés, tout en intégrant d'autres, bien connus dans les milieux universitaires mais rarement examinés par les acteurs du secteur de l'inclusion,
— Nous avons cherché à équilibrer la présentation de projets aisément réutilisables et d'autres, plus complexes à reproduire, mais néanmoins très inspirants.

À propos du projet Algo→Lit

Le projet Algo→Lit vise à améliorer les connaissances algorithmiques des professionnels de l'inclusion numérique pour qu'ils puissent à leur tour accompagner les citoyens dans leur compréhension et dans l'exercice de leur droit à la transparence. À l'heure où les citoyens sont de plus en plus confrontés à des algorithmes opaques, les professionnels de l'inclusion numérique doivent faire face à des demandes croissantes concernant ces systèmes. Nous visons à mettre à disposition de ces professionnels les compétences et les outils dont ils ont besoin pour aider leurs publics dans la compréhension des systèmes algorithmiques et l'exercice de leurs droits. Le projet ALGO->LIT est coordonné par Datactivist, la coopérative française spécialisée dans l'ouverture des données et des algorithmes, en partenariat avec La Mednum (Paris, FR), la coopérative française représentant le réseau des travailleurs de l'inclusion numérique, ainsi que les centres de recherche FARI - AI for the common good institute (Bruxelles, BE) et Waag FutureLab (Amsterdam, NL).

Si vous pensez qu'un exemple a été oublié, n'hésitez pas à nous contacter à l'adresse suivante : algolit@datactivi.st

Pour en savoir plus sur le projet Algo->Lit, rendez-vous sur notre site web.

Funded by
the European Union

Typologie des exemples présentés dans ce document :

1. ATELIERS COLLABORATIFS

Les ateliers collaboratifs sont des sessions pratiques de 1 à 3 heures visant à impliquer activement le public dans la compréhension des algorithmes et des impacts de l'IA. Ils consistent le plus souvent à construire, à l'aide de modèles, une représentation des enjeux abordés.

2. MÉCANISMES COLLECTIFS DE PRISE DE DÉCISION ET DE COORDINATION

Les mécanismes collectifs de prise de décision et de supervision sont des initiatives qui permettent à un groupe de citoyens de participer et de (re)façonner modérément la gouvernance, la conception ou les effets d'un système algorithmique donné. Ils prennent généralement la forme de jurys citoyens, de conventions/assemblées citoyennes, de mécanismes de supervision, d'observatoires, de projets de donation de données ou d'autres formes d'initiatives de crowdsourcing.

3. SUPPORTS VISUELS PÉDAGOGIQUES NON INTERACTIFS

Les visualisations ou diagrammes sont des représentations graphiques complexes conçues pour illustrer le fonctionnement d'un mécanisme algorithmique. Ils permettent de simplifier et de rendre accessibles les aspects des algorithmes qui sont autrement difficiles à percevoir. De plus, ils visent parfois à donner une image exhaustive des nombreux acteurs impliqués dans la création d'algorithmes ou à présenter les algorithmes différemment de leurs concepteurs, par exemple en mettant l'accent sur les dimensions liées aux impacts sociaux ou environnementaux des algorithmes et de l'IA.

4. DOCUMENTATIONS SUR LES ALGORITHMES ET RÉCITS SUR LES ENSEMBLES DE DONNÉES

Les documentations sont des documents écrits qui expliquent, commentent et justifient les choix effectués lors de la conception d'un algorithme donné. Elles peuvent potentiellement fournir des informations sur ses résultats et ses effets. Les récits de jeux de données sont des formes narratives visant à documenter la provenance des données utilisées pour former ou affiner un système d'IA, la manière dont elles ont été transformées, leurs attributs et leurs écueils.

5. SIMULATEURS ET DISPOSITIFS DE DÉMONSTRATION

Un simulateur ou dispositif de démonstration d'algorithme offre une interaction dynamique avec une interface simplifiant les calculs, mobilisant des données hétérogènes tout en donnant l'impression que l'on peut manipuler directement un algorithme. Par rapport à un algorithme écrit ou à sa documentation, l'avantage supplémentaire d'un simulateur est de permettre de tester différents paramètres et configurations, d'effectuer une interaction ludique qui invite à vivre une expérience heuristique..

I) Ateliers collaboratifs

DESCRIPTION DE VOS PRATIQUES LLM

Capture d'écran du site web du projet

Traduction : *Dans quelle mesure un LLM peut-il se rapprocher de l'excellence ? Dans ce bloc, les participant·es utilisent leur modèle pour produire le travail le plus qualitatif possible. L'objectif est de tester le modèle dans toute son amplitude, en évaluant sa capacité à atteindre une véritable excellence. Où les LLMs se situent-ils dans nos pratiques professionnelles, nos normes et nos relations ? Le premier bloc porte sur l'adoption d'une posture réflexive vis-à-vis de l'usage des LLMs. Les participant·es prennent part à des exercices qui les rendent plus conscient·es de leurs pratiques, attitudes et présupposés concernant les technologies d'intelligence artificielle. Ce bloc agit comme une phase de fondation — un terrain d'entraînement — où les participant·es commencent à penser de manière critique leur relation aux LLMs.*

Le projet Ecologies of LLM développé par des designers et des sociologues du medialab de Sciences Po pose la question suivante : comment repenser le rôle des modèles linguistiques à grande échelle (LLM) tels que ChatGPT dans les pratiques professionnelles courantes ? Pour eux, il est urgent d'ouvrir la voie à d'autres approches de l'IA, afin de dépasser la seule logique prédictive et de recentrer l'attention sur le présent. L'équipe du projet invite à entamer un travail lent et silencieux pour analyser ce que l'IA produit déjà dans nos pratiques professionnelles. Ce n'est qu'après avoir prêté attention à nos pratiques que nous pourrons redéfinir cette technologie selon nos propres termes et choisir comment l'utiliser. Plus important encore, observer comment l'IA s'installe dans nos routines professionnelles pourrait éclairer les mécanismes de travail qui l'ont rendue possible. L'enjeu n'est pas seulement la manière dont nous travaillons avec l'IA, mais aussi dans la question de savoir qui détermine ce

protocole, pour documenter de près comment l'IA affecte et pourrait affecter votre travail. Leur livre agit comme une lentille artificielle pour soutenir votre recherche. Leur guide peut être vu comme un outil pour interroger votre propre utilisation des LLM et échanger vos points de vue avec d'autres professionnels confrontés à des enjeux et imprévus comparables.

Voici quelques exemples d'exercices que vous pouvez réaliser grâce à ce livre :

Ex^{2b} – All the *Things* You *Could* Do

Instructions

Think about the tasks you currently do without the help of LLMs and where you could benefit from their assistance.

This list isn't about automation (delegating to the machine) but augmentation (doing more tasks or diversifying tasks). Think about activities or projects you haven't attempted yet or tasks you might hesitate to delegate.

Even if you suspect that an LLM might perform poorly on a given task, include it in your list: this exercise is about exploring potential opportunities rather than evaluating performance.

Traduction : Réfléchissez aux tâches que vous réalisez actuellement sans l'aide des LLMs, et identifiez celles où leur assistance pourrait vous être bénéfique. Cette liste ne porte pas sur l'automatisation (déléguer à la machine), mais sur l'augmentation (faire davantage de tâches ou en diversifier la nature). Pensez à des activités ou des projets que vous n'avez pas encore entrepris, ou à des tâches que vous hésitez à déléguer. Même si vous pensez qu'un LLM pourrait mal performer sur une tâche donnée, incluez-la dans votre liste : cet exercice vise à explorer des possibilités potentielles, plutôt qu'à évaluer des performances.

Ex[3a] – Rough *Impressions*

Instructions

1. Set a timer and spend 5 minutes reflecting on your overall LLM use. Don't look at your conversation history yet!

   → Assess which tasks you delegate and assign a rough percentage to each (e.g., one-third of my tasks are translation, and the rest are drafting email answers). Consider how frequently you use the LLM (daily, weekly, or less frequently).

   → Reflect on the structure of your interactions and decide how many of your conversations are: short exchanges (a single request and answer); longer conversations (several iterations); or extended dialogue (conversations that last more than a day).

2. Finally, recall some memorable moments in your use of the LLM. Think back to two or three conversations that left a lasting impression.

   → Describe them briefly and why they were memorable.

Traduction : Réglez un minuteur et prenez 5 minutes pour réfléchir à votre utilisation générale des LLMs. Ne consultez pas encore votre historique de conversations ! → Évaluez les tâches que vous déléguez au LLM et attribuez à chacune un pourcentage approximatif (ex. : un tiers de mes tâches concernent la traduction, le reste consiste à rédiger des réponses à des e-mails). Réfléchissez à la fréquence à laquelle vous utilisez le LLM — quotidiennement, chaque semaine, ou plus rarement. → Analysez la structure de vos interactions et estimez combien de vos conversations sont : de courts échanges (une seule demande et une réponse) ; des conversations plus longues (plusieurs itérations) ; ou des dialogues prolongés (qui s'étendent sur plus d'une journée). Enfin, remémorez-vous quelques moments marquants dans votre utilisation du LLM. Pensez à deux ou trois conversations qui vous ont laissé une impression durable. → Décrivez-les brièvement et expliquez pourquoi elles vous ont marqué·e.

EXPLORER LES OUTILS D'ÉCRITURE GÉNÉRIQUE BASÉS SUR L'IA AVEC LA BOITE NOIRE DE L'IA

Une session du jeu à Nantes en avril 2025

Comment se positionner face à une technologie que même les experts ont du mal à maîtriser ? The AI Black Box est un jeu éducatif créé par la coopérative Datactivist et le réseau des travailleurs de l'inclusion de Nantes Métropole (France). Il est conçu pour aider le grand public à comprendre le fonctionnement de l'IA génératrice de texte. Les joueurs découvrent le cycle de vie d'un produit d'IA générative tel que ChatGPT, depuis l'extraction des minéraux jusqu'aux coûts énergétiques des centres de données. Ce faisant, ils débattent également des conséquences sociétales, environnementales et éthiques de l'IA. Contrairement à de nombreux formats similaires, la boite noire de l'IA se concentre sur la création de l'IA et ne se contente pas de débattre de ses impacts sur le travail, la vie sociale ou la nature. La première partie du jeu consiste à reconstituer les principales étapes menant à la création d'une IA génératrice de texte (ex : entraînement, collecte de données, etc.). Dans la deuxième partie, les joueurs doivent associer, à chaque étape, des cartes correspondant à une sous-étape de conception. Par exemple, la carte “filtrage des données” doit être placée lors de l’étape consacrée au nettoyage des données.

Dans ce jeu, de nombreux processus informatiques sont simplifiés sans jargon. Considérez-le comme une grande plongée au cœur d'un système d'IA. Il est préférable d'y jouer en petit groupe de 7 personnes. Le jeu ne nécessite aucune connaissance préalable, c'est un atelier ouvert à tous. Le format est minimaliste, ouvert (CC-BY-SA) et facilement réutilisable. Le kit comprend des cartes prêtes à être imprimées, des conseils pratiques pour animer l’atelier ainsi qu’un jeu de diapositives pour introduire la session.

Au moins, je vais pouvoir démystifier l'IA grâce à un jeu. Pour moi, c'est toujours une chance d'avoir un jeu dont on peut discuter.

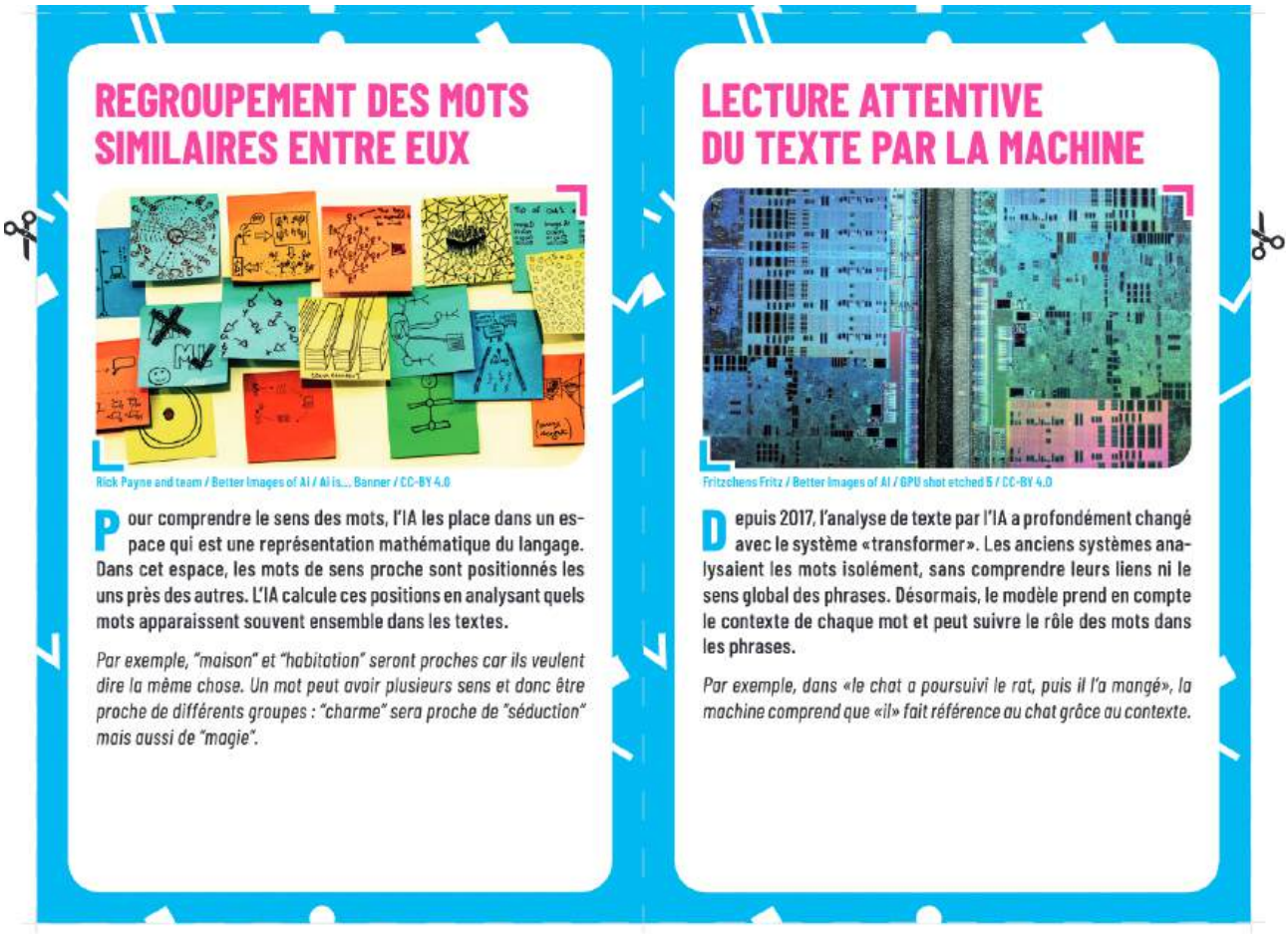
– *Participant à la Boîte Noire de l'IA*

J'ai apprécié d'avoir une vue d'ensemble de toutes les étapes nécessaires à la création d'une IA. Cela m'a aidé à comprendre les utilisations et les enjeux environnementaux liés à l'IA.

– *Participante à la Boîte Noire de l'IA*

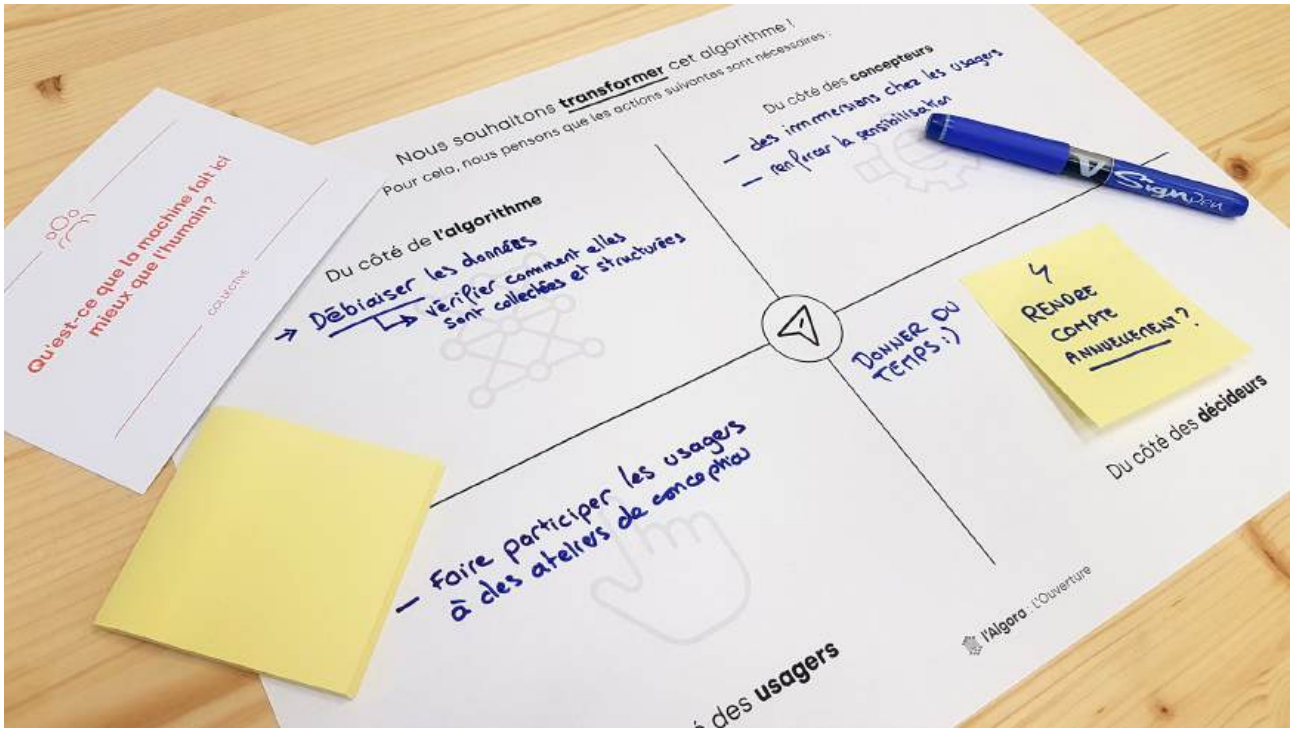
J'ai aimé réorganiser les étapes. J'ai encore plus aimé le jeu social, car l'IA redevient un objet social et non plus seulement un objet technique.

– *Participant à la Boîte Noire de l'IA*



Exemple de deux cartes « action »

DISCUTER DES IMPACTS SOUHAITÉS DE L'IA AVEC THE ALGORA



Une image du modèle d'affiche utilisé lors d'un atelier

The Algora est un outil de discussion permettant de comprendre les impacts souhaités, inattendus et indésirables d'un algorithme existant, en cours d'exécution, et d'agir en conséquence. Il a été développé par Design Friction et commandé par Etalab, l'équipe chargée des données ouvertes de l'État français. Conçu pour être utilisé dans le cadre d'un atelier collectif, l'Algora facilite la communication entre différents profils (décideurs, concepteurs, développeurs, fonctionnaires, utilisateurs). À l'aide d'un jeu de cartes de conversation, chaque participant est invité à exprimer son point de vue et à le confronter à la réalité des autres. L'objectif du jeu est d'identifier collectivement ce que produit réellement l'algorithme discuté. Le kit Algora comprend : un jeu de 40 cartes de discussion, des tableaux de synthèse et d'action, des ressources pour l'animation, une version pour les ateliers en ligne.

DÉCOUVRIR L'IA AVEC ST[IA]MMTISCH

Le plateau de jeu

Le jeu StIAmmtisch, conçu par la Délégation académique régionale du Grand Est (France), s'inspire de la tradition du Stammtisch (une table traditionnelle en Alsace et dans d'autres régions germanophones) : il s'agit avant tout de se retrouver pour passer un bon moment, autour d'un sujet d'actualité : l'intelligence artificielle. C'est un jeu de société facile à utiliser, pour 2 à 6 joueurs ou équipes, accompagné d'un animateur. Le jeu dure environ 50 minutes et est conçu pour être utilisé en classe. Cependant, il peut être joué aussi bien par des adultes que par des élèves. Les niveaux junior et intermédiaire sont actuellement disponibles. Un niveau expert est prévu depuis la création du jeu, mais celui-ci n'est pas encore disponible.

II) Mécanismes collectifs de prise de décision et de coordination

ORGANISER UN PARLEMENT HUMAIN DU NUMÉRIQUE

Un parlement humain en 2023 s'est tenu sur la place Saint-Jean, près du Parlement francophone de Bruxelles.

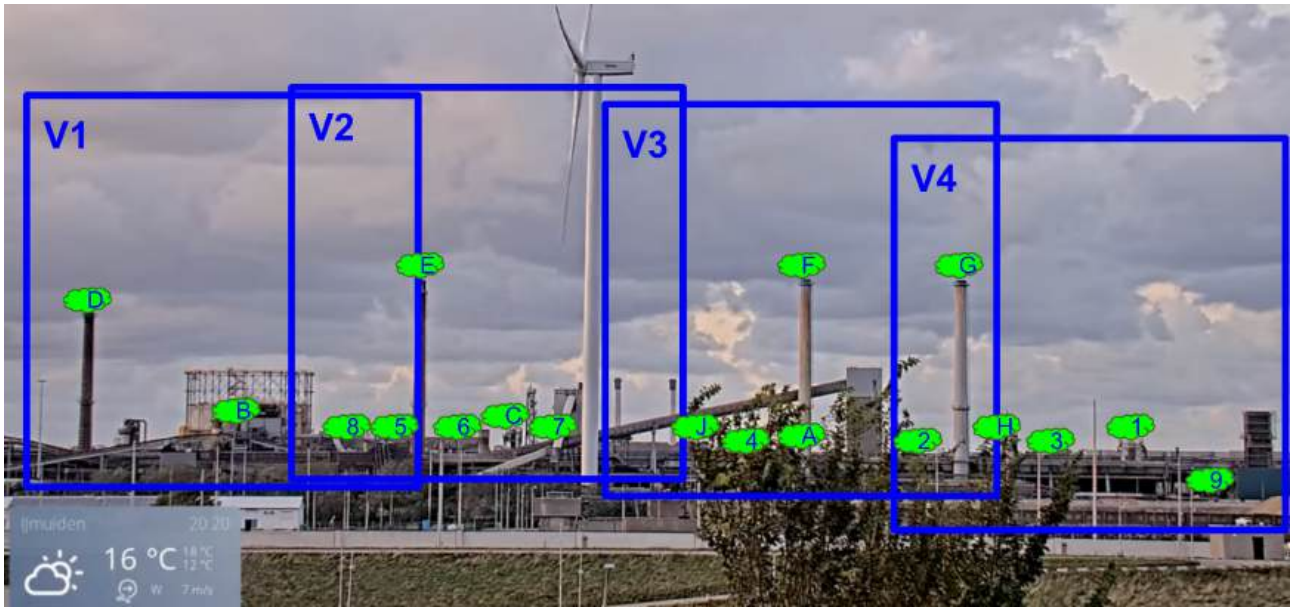
Un « Parlement humain », organisé par le Comité humain du numérique, est un moment pour discuter et écouter, partager nos expériences, nos souffrances et nous entraider au sujet du numérique dans notre vie quotidienne. Il est conçu comme un lieu où nous parlons à partir de notre propre expérience, et non à partir de théories ou d'idées. Un espace où chacun est un expert et où un monde plus juste peut exister ; un outil d'animation pour élaborer des lois, des règles collectives ou des revendications. Cela peut prendre du temps et nécessiter plusieurs réunions : l'efficacité n'est pas une priorité. L'objectif est de défendre le visage humain de la technologie numérique dans nos vies lors d'une réunion et en rédigeant des lois pour garantir que la technologie numérique s'adapte à l'humain et non l'inverse. À l'issue du Parlement humain, un Code humain du numérique a été publié : il s'agit d'une loi citoyenne autoproclamée en gestation depuis 2021. Ce livre rassemble plus de 3 ans d'actions et de rencontres à Bruxelles et en Wallonie : réunions avec des citoyens, chercheurs en droit, politiciens, discussions intimes, ateliers, créations de décors en carton, repas partagés, manifestations, etc.

ORGANISATION D'UNE CONVENTION CITOYENNE SUR L'UTILISATION DE CHATGPT

Le groupe d'étudiants et d'enseignants de Sciences Po Aix participant à la convention

Avec l'accélération de l'utilisation de ChatGPT par les étudiants dans l'enseignement supérieur et la recherche, il est urgent de repenser sa place dans la relation pédagogique et d'accompagner son utilisation dans la formation des étudiants. Pour relever ce défi, l'école Sciences Po Aix (France) a lancé, du 6 mars au 12 mai 2025, une Convention citoyenne universitaire sur l'IA générative. Cette initiative sans précédent vise à ouvrir le débat sur le sujet en créant un espace d'échange et de discussion impliquant tous les membres de la communauté éducative de l'école : enseignants, étudiants et personnel administratif. À l'issue de ses travaux, les membres de la Convention ont soumis à la direction de l'école une série de propositions pour la formation et l'enseignement à Sciences Po Aix à l'ère de l'IA.

ÉTIQUETAGE DES DONNÉES POUR L'IA AVEC LE PROJET IJMONDCAM



Images de nuages toxiques cartographiées par le réseau neuronal

Waag FutureLab et l'équipe de Yen-Chia Hsu de l'université d'Amsterdam ont organisé deux réunions au cours desquelles les habitants de la région d'Ijmond (Pays-Bas) ont pu donner leur avis sur l'étiquetage des données d'un système d'IA destiné à détecter la fumée. En d'autres termes, il s'agissait de déterminer avec précision ce qu'est un nuage de fumée suspect. Le système d'IA affiche une image avec un marqueur carré autour du nuage de fumée suspect. Les habitants peuvent ajuster ce marquage afin d'améliorer la précision du système. En étiquetant les données pour constituer un jeu d'apprentissage, la communauté influence directement les choix opérés par le modèle. Si les réseaux neuronaux profonds se sont avérés utiles dans diverses applications (par exemple, la reconnaissance d'objets), l'apprentissage d'un tel modèle nécessite une grande quantité de données étiquetées. L'étiquetage manuel de toutes ces données peut facilement nécessiter plusieurs centaines d'heures pour une seule étude. C'est pourquoi ils ont besoin de bénévoles. Le système détermine une étiquette en fonction de toutes les étiquettes fournies par les chercheurs et les citoyens. Chaque vidéo ou image est visionnée par au moins deux volontaires et un chercheur. Si les volontaires ont attribué la même étiquette à la vidéo ou à l'image, le système l'enregistre. Sinon, le processus se répète. Le résultat est donc déterminé par un vote à la majorité. Les travaux antérieurs sur la surveillance de la qualité de l'air montrent que la visualisation des preuves d'émissions de fumée contribue à attirer l'attention des régulateurs. Elle contribue également à renforcer la confiance de la communauté dans la lutte contre la pollution atmosphérique.

ÉVALUER LES RISQUES LIÉS À L'IA GRÂCE AU « LABOR AI AUTO-TEST »

Illustration du projet LaborIA

Avant d'intégrer un système d'IA, il est essentiel d'évaluer ses avantages et ses risques potentiels. À titre de diagnostic initial, l'outil d'auto-évaluation développé par le projet de recherche Labor AI (INRIA, France) permet d'évaluer les aspects critiques liés à l'intégration d'un système d'IA dans une organisation et d'identifier les domaines nécessitant une vigilance particulière. Cet outil fournit une première évaluation des risques liés à l'intégration d'un système d'IA, sans se substituer aux obligations légales applicables à l'introduction de nouvelles technologies en entreprise.

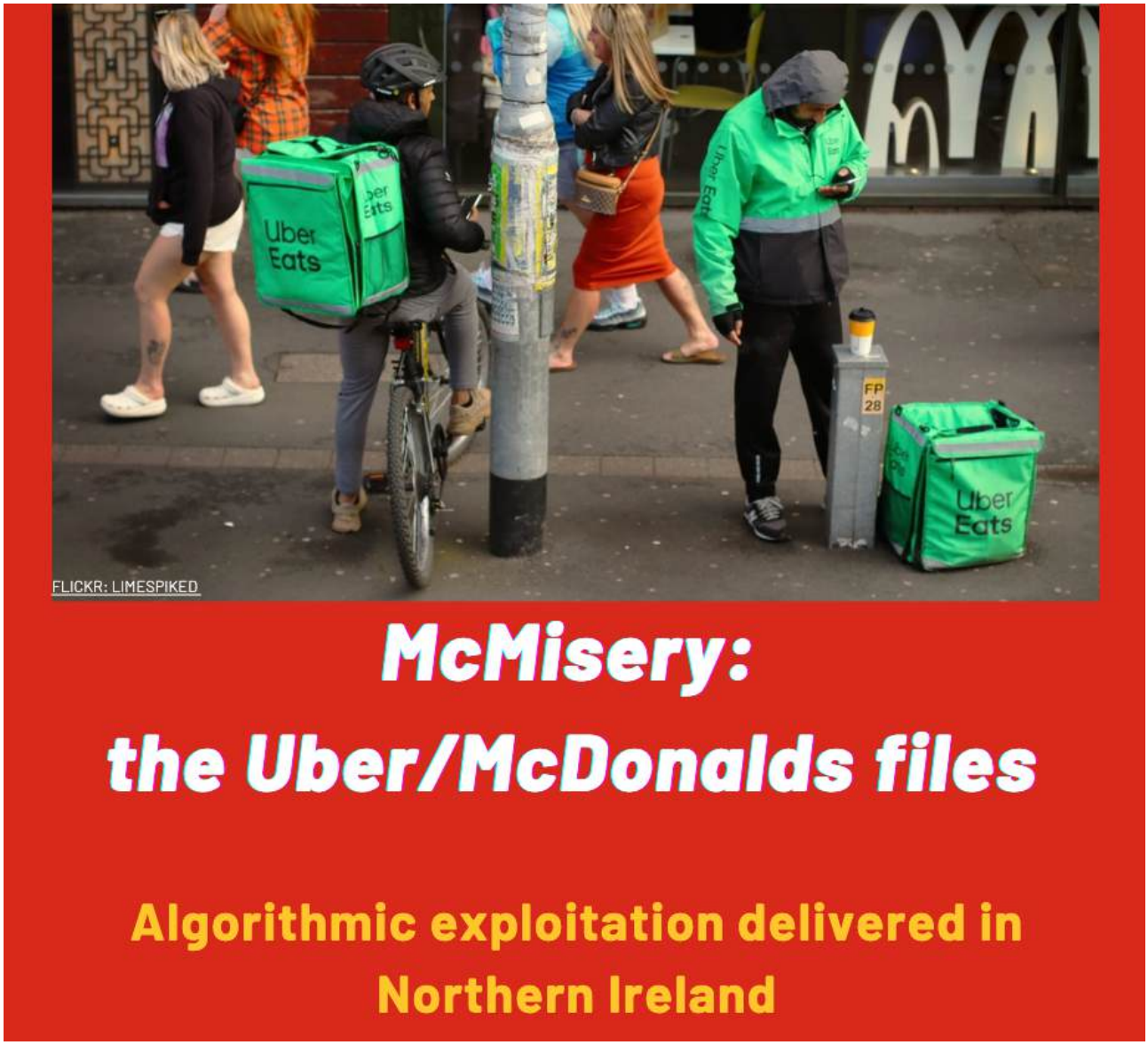
DONNÉES ISSUES DU CROWDSOURCING AVEC L'OBSERVATOIRE DES ALGORITHMES DES TRAVAILLEURS

Capture d'écran du site web du WAO

Traduction : Audit mené par les travailleurs dans l'économie des plateformes

Le WAO est un projet collaboratif d'audit participatif basé aux États-Unis et hébergé par l'université de Princeton. Le WAO aide les travailleurs et leurs alliés à enquêter sur les systèmes algorithmiques opaques. Il collecte de manière participative les données nécessaires à des audits scientifiques pertinents, et développe des outils et des ressources pour les travailleurs et leurs alliés. Il a été lancé en 2022. Il s'agit actuellement d'une initiative à but non lucratif financée par le Mozilla Tech Fund 2023 « Auditing AI ». Les projets de la WAO comprennent par exemple FairFare, une initiative collective visant à aider les chauffeurs, les organisateurs et les décideurs politiques à comprendre le secteur des voitures de transport avec chauffeur (VTC) en collectant de manière participative les données tarifaires des chauffeurs. Un projet précédent était The Shipt Calculator, une autre étude menée par des travailleurs sur un algorithme opaque utilisé par la société de livraison Shipt. Willy Solis, membre du Gig Workers Collective qui a dirigé l'étude, a collaboré avec Coworker.org et l'équipe de WAO pour collecter et analyser les données partagées par plus d'une centaine de travailleurs à l'aide d'un chatbot de messagerie personnalisé. Leur analyse montre que les salaires ont baissé pour la majorité mais augmenté pour une minorité, révélant que l'algorithme ne détermine pas seulement la rémunération : il impacte également l'autonomie et le bien-être des employés.

AIDER LES TRAVAILLEURS À LUTTER CONTRE LES ALGORITHMES GRÂCE À WORKERINFOEXCHANGE



Couverture de l'enquête « McMisery » concernant les chauffeurs Uber et McDonald's en Irlande du Nord

Traduction : les dossiers Uber/McDonald's / Exploitation algorithmique mise en œuvre en Irlande du Nord

WorkerInfoExchange est une organisation (coopérative de données) basée au Royaume-Uni qui recueille des données personnelles et des cas de discrimination algorithmique afin de reconstituer les problèmes systémiques liés à la gestion algorithmique. Elle a par exemple démontré que l'introduction par Uber d'une rémunération dynamique a réduit le salaire des chauffeurs, augmenté la commission d'Uber et rendu la rémunération plus imprévisible et inégale pour sa main-d'œuvre. Le système de rémunération dit « dynamique » d'Uber utilise l'intelligence artificielle et l'apprentissage automatique pour fixer les salaires et attribuer le travail en temps réel, sans aucune transparence pour les chauffeurs et sans possibilité de contester les décisions automatisées en matière de rémunération et d'attribution du travail.

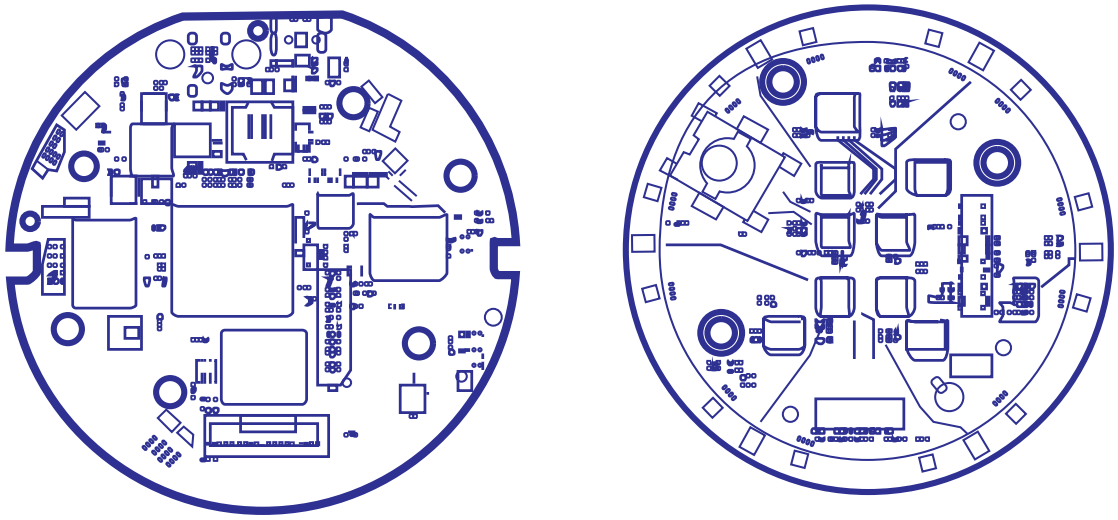
AU-DELÀ DES ROBOTS : POURQUOI NOUS AVONS BESOIN D'UNE MEILLEURE IMAGE DE L'IA

Jamillah Knowles & Digit / Better Images of AI / CC-BY-4.0

La prédominance d'images inspirées de la science-fiction et anthropomorphisées, ainsi que le manque d'images ou d'idées alternatives facilement accessibles, rendent difficile une communication précise sur l'IA. En effet, ces images générées par l'IA renforcent également la méfiance du public à l'égard de l'IA, un problème croissant pour l'innovation dans un domaine parfois considéré comme biaisé, opaque et extractif. Les images dominantes actuelles véhiculent des idées fausses potentiellement dangereuses et, au mieux, freinent la compréhension du public sur l'usage réel, le fonctionnement, le potentiel et les implications des systèmes d'IA.

Le projet Better Images of AI vise à créer une nouvelle banque d'images de meilleure qualité sur l'IA, accessible à tous, en commençant par une collection d'images inspirantes. La première étape de ce projet consiste à explorer à quoi pourraient ressembler ces nouvelles images et à inviter des personnes issues de différents horizons créatifs, techniques et autres à travailler ensemble pour développer de meilleures images. Pour produire de nouvelles images, il nous faut d'abord définir ce qui constitue une bonne image d'archive. Pour quelles raisons les utilise-t-on et de quelle manière ? L'image représente-t-elle un aspect particulier de la technologie ou tente-t-elle de raconter une histoire plus large ? Quelle réaction émotionnelle cette image devrait-elle susciter chez le public ? Aide-t-elle les gens à comprendre la technologie et en donne-t-elle une représentation fidèle ?

DÉCOUVRIR L'ANATOMIE D'UN SYSTÈME D'IA



amazon-echo-dot

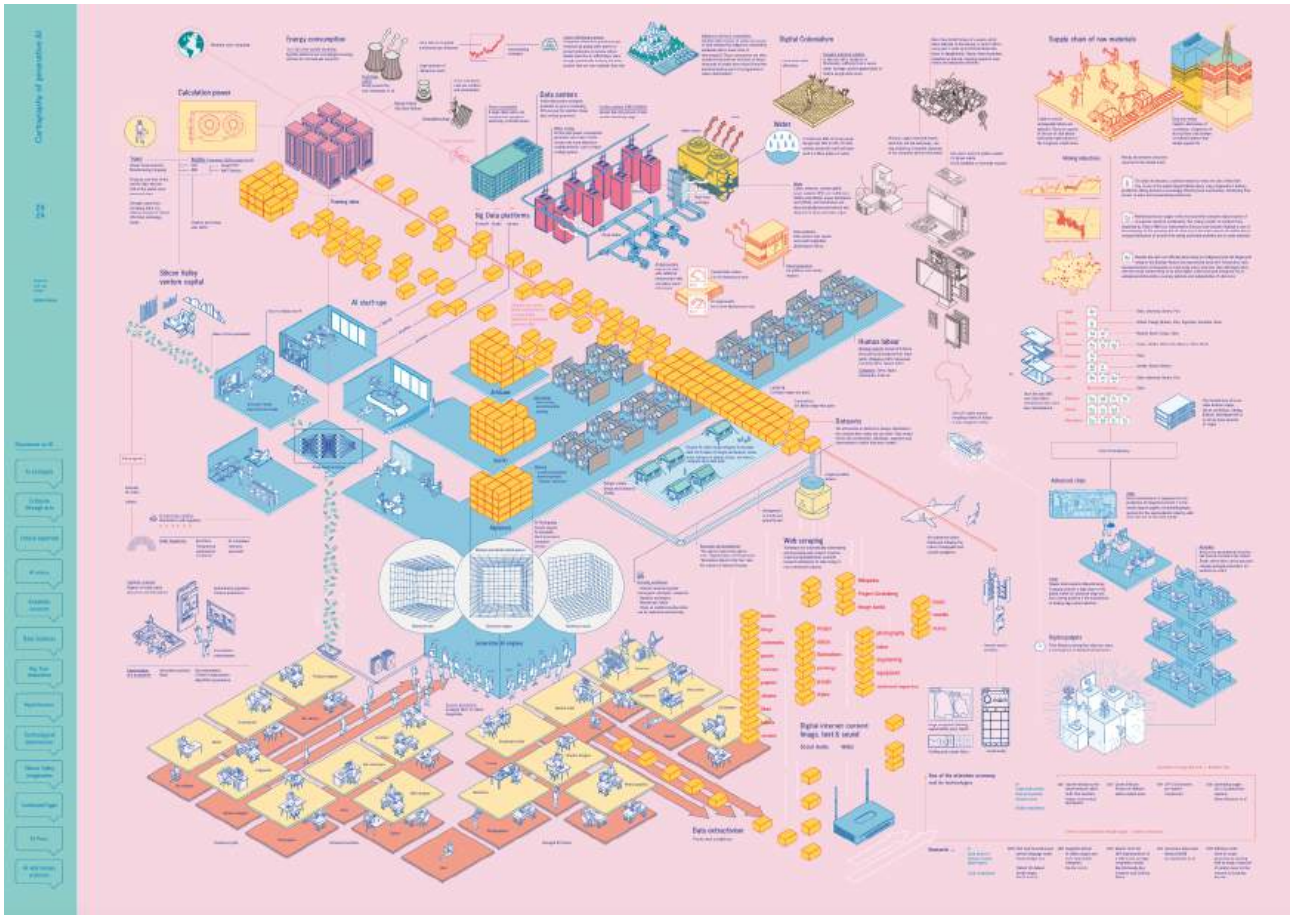
Anatomy of an AI System est la carte anatomique du travail humain, des données et des ressources planétaires nécessaires au fonctionnement de l'appareil Amazon Echo. En prendant l'exemple de l'Amazon Echo, les auteurs illustrent la multitude de composants et de facteurs impliqués dans la production des systèmes d'intelligence artificielle. Mais ce processus est si complexe que son ampleur totale est presque impossible à saisir.

> Un cylindre se trouve dans une pièce. Il est impassible, lisse, simple et petit. Il mesure 14,8 cm de haut et est doté d'une seule lumière circulaire bleu-vert qui trace le contour de son bord supérieur. Il est silencieux. Une femme entre dans la pièce, portant un enfant endormi dans ses bras, et s'adresse au cylindre.
>
> Alexa, allume les lumières du couloir.
>
> Le cylindre s'anime. « OK. » La pièce s'illumine. La femme fait un léger signe de tête et emmène l'enfant à l'étage.
>
> *– Une interaction avec un appareil Amazon Echo*

Le projet commence par une esquisse : une vue éclatée d'un système planétaire à travers trois étapes (naissance, vie et mort), accompagnée d'un essai en 21 parties. Ensemble, cela devient une carte anatomique d'un seul système d'IA. Comment pouvons-nous commencer à le percevoir, à saisir son immensité et sa complexité en tant que forme connectée ? Les ressources nécessaires à la production de systèmes tels qu'Amazon Echo, un assistant personnel connecté et contrôlé par la voix, dépassent largement les aspects purement techniques de la modélisation de données, du matériel, des serveurs et des réseaux: les véritables coûts – sociaux, écologiques, économiques et politiques – demeurent pour la plupart invisibles.

Elle a été créée par l'artiste Vladan Joler et la chercheuse Kate Crawford.

VISUALISATION DE L'INFRASTRUCTURE DERRIÈRE L'IA GÉNÉRATIVE



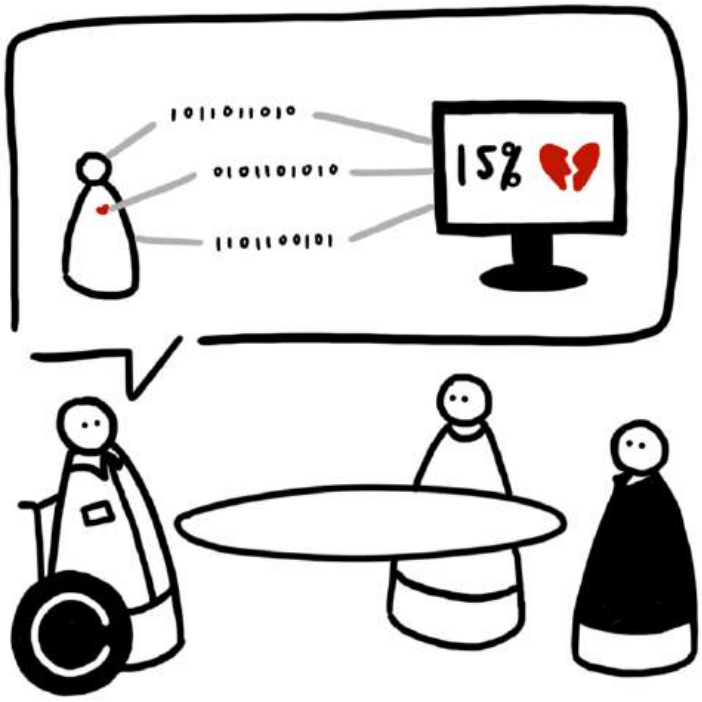
L'ensemble de la cartographie de l'IA générative

Quelles formes d'extraction, quels systèmes autonomes et quelles ressources rendent possible le fait de converser en ligne avec un générateur de texte ou de créer une image en quelques secondes ?

Le projet Cartographie de l'IA générative La cartographie de l'IA générative, conçu par le studio de design espagnol Estampa, propose une représentation visuelle de l'ensemble de la chaîne d'approvisionnement d'une IA générative : depuis la constitution des ensembles de données jusqu'à la consommation énergétique. La carte met également en lumière les dimensions matérielles, économiques et humaines de ces systèmes : puissance de calcul, travail humain et structures financières. Cette mosaïque relie des objets, des savoirs et des échelles très différentes. Les discours sur l'IA sont d'ailleurs souvent teintés de mythes et d'imaginaires technologiques : l'idée d'une autonomie totale des algorithmes, d'une technologie inévitable qui façonnerait notre avenir, ou encore d'une supposée universalité des données. Ces récits, qu'ils soient experts ou populaires, participent eux-mêmes à construire la manière dont nous percevons et développons ces technologies. Face à cela, le projet Cartographie de l'IA générative cherche à offrir une carte conceptuelle critique, complexe, capable de rendre visibles les acteurs, les ressources et les tensions qui traversent cet écosystème. Inscrite dans la tradition des cartographies critiques, cette visualisation ne se limite pas à décrire le phénomène : elle invite à penser les liens, les dépendances et les controverses qui rendent l'IA générative possible.

IV) Documentations sur les algorithmes et récits sur les ensembles de données

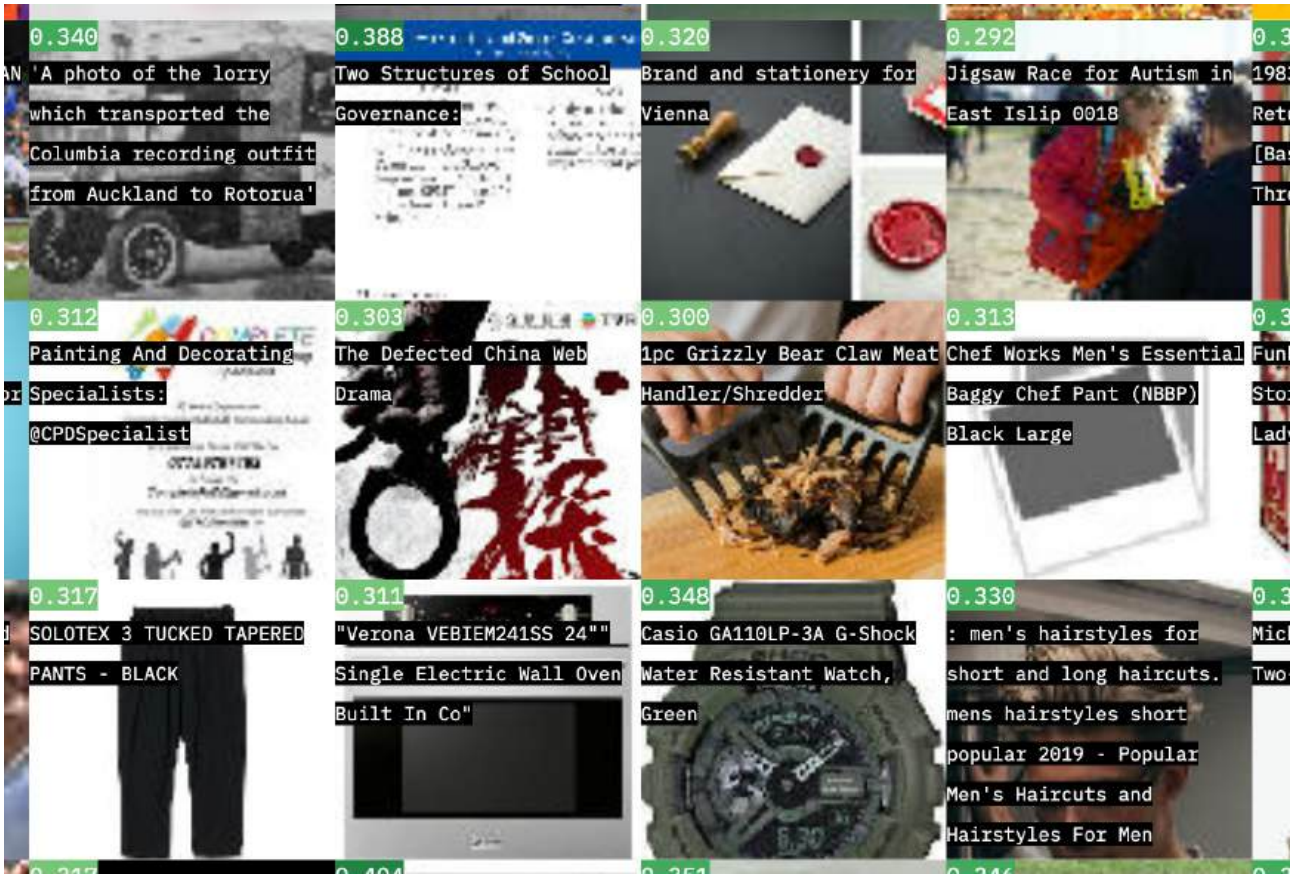
COMMENT FONCTIONNE LE JUMELAGE ? LE CAS DES GREFFES DE REIN



Prediction by Yaning Wu / Better Images of AI / CCBY-4.0

L'Agence française de biomédecine a publié un guide pour comprendre le score qui détermine l'attribution des greffes de rein. Cela s'inscrit dans un contexte où la pénurie de greffes continue de soulever des questions médicales et éthiques difficiles concernant le processus de sélection des patients qui se verront attribuer un greffon provenant d'un donneur décédé. L'attribution des greffons est un processus complexe et sensible, tant dans sa conception que dans son exécution. Il n'existe pas de solution globale ou définitive. Elle varie d'un pays à l'autre, en fonction des critères d'attribution pris en compte, et dans le temps, en fonction de l'évaluation périodique des résultats. Un système d'attribution doit être adapté, efficace et aussi équitable que possible par rapport aux besoins de santé des patients concernés. Pour certains patients, une solution consiste à proposer des greffes sur une base « prioritaire ». C'est le cas des patients en situation d'urgence ou présentant une hyperimmunité, qui sont actuellement prioritaires, généralement au niveau national. Pour la grande majorité, il est impossible de donner la priorité à un groupe plutôt qu'à un autre. Le système d'attribution doit donc trouver un compromis entre équité, efficacité et faisabilité. Une solution efficace consiste à utiliser un score. Cette documentation officielle de l'Agence biomédicale française explique comment ce score est calculé.

QUELLES SONT LES DONNÉES QUI SOUS-TENDENT L'IA ?



Capture d'écran de l'article montrant les métadonnées des images

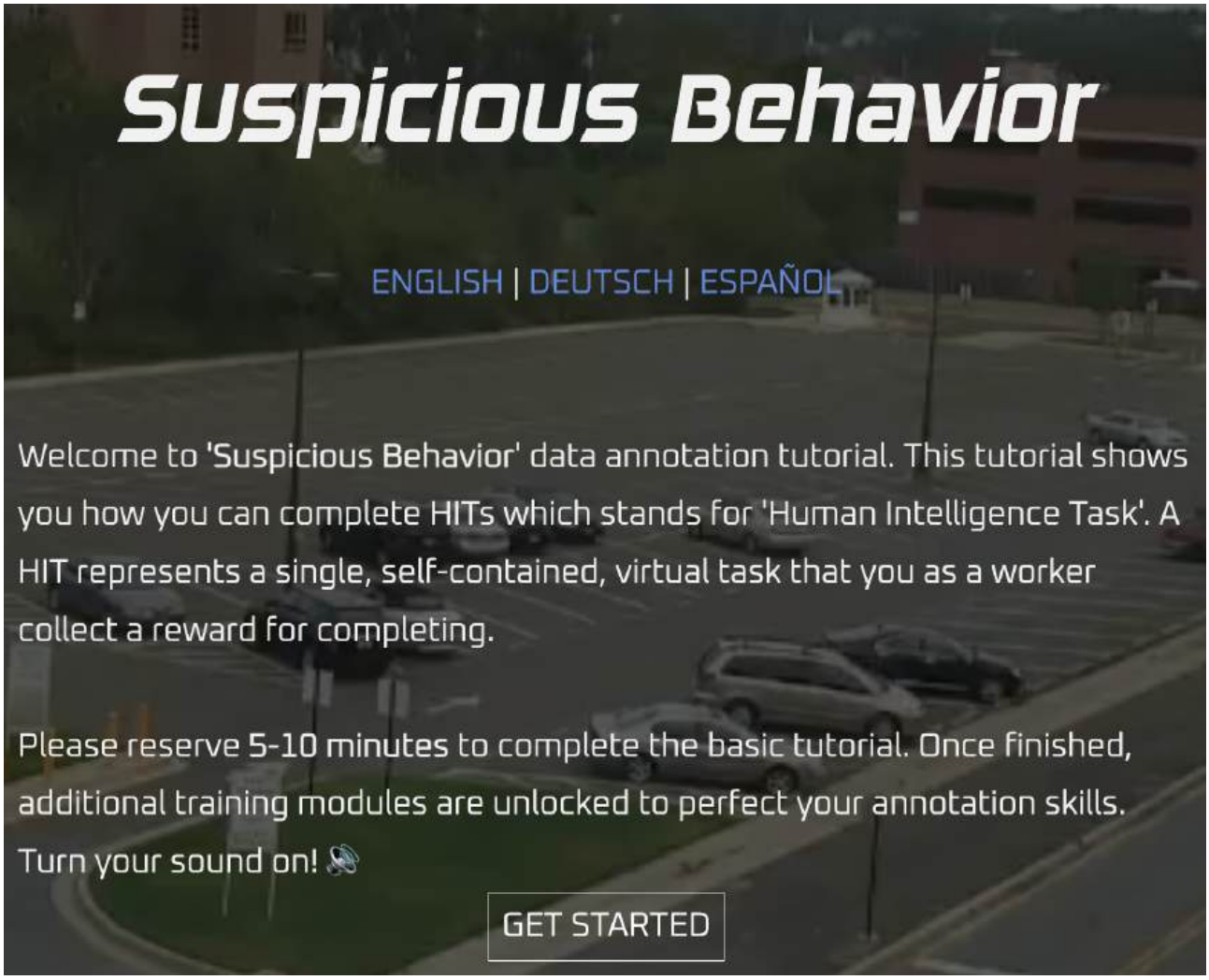
Models all the way down est un récit visuel qui peut être compris comme une forme narrative de documentation. Il raconte l'histoire de LAION-5B, un ensemble de données open source utilisé pour entraîner des modèles d'IA tels que Stable Diffusion. Il contient 5,8 milliards de paires d'images et de textes, une taille trop importante pour être comprise. Dans cette enquête visuelle créée par Christo Buschek et Jer Thorpthe dans le cadre du projet de recherche Knowing Machines, nous suivons la construction de l'ensemble de données afin de mieux comprendre son contenu, ses implications et ses imbrications.

DÉMONSTRATION DE LA FIABILITÉ DES RÉSEAUX NEURONAUX AVEC PYRAT

Le démonstrateur PyRAT

Une équipe de chercheurs du CEA-List, un laboratoire de recherche appartenant à l'organisme public français de recherche technologique et énergétique, a développé un nouveau logiciel appelé PyRAT. Conçu pour prouver la fiabilité de la réponse des réseaux neuronaux dans des configurations données, PyRAT peut aider à évaluer la fiabilité des systèmes basés sur l'IA dans des applications critiques allant de la prévention des collisions aériennes aux diagnostics médicaux. Le laboratoire avait à cœur de présenter PyRAT à un large public, notamment des chercheurs, des scientifiques, des partenaires industriels ou des prospects, de manière à la fois technique et pédagogique, à travers un démonstrateur exposé dans le showroom du CEA à Paris-Saclay. Le studio de design français Units a collaboré avec le laboratoire afin de trouver la meilleure façon de représenter et d'expliquer les processus hautement complexes et opaques à l'œuvre dans les grands réseaux neuronaux. Le résultat est un démonstrateur qui dévoile la logique derrière PyRAT et les différentes opérations mathématiques en jeu, grâce auquel les utilisateurs peuvent expérimenter différents paramètres d'analyse. L'interface est visible dans les showrooms du CEA et se contrôle à l'aide d'un dispositif matériel dédié, conçu et développé par Units.

COMMENT LES ALGORITHMES INTERPRÈTENT-ILS LES IMAGES DE COMPORTEMENTS SUSPECTS ?

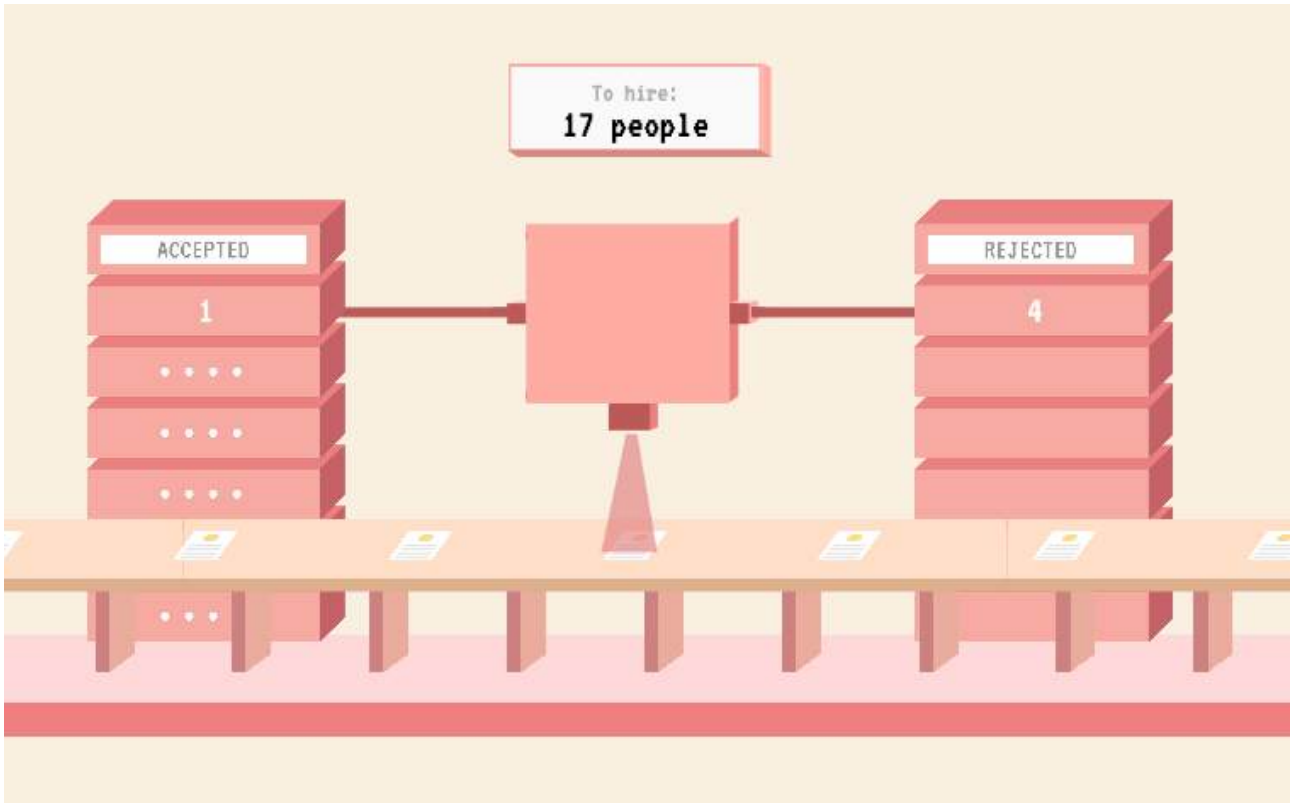


Une capture d'écran du jeu

Traduction : Bienvenue dans le tutoriel d'annotation de données « Suspicious Behavior ». Ce tutoriel vous montre comment accomplir des HITs, ce qui signifie « Human Intelligence Task » (tâche d'intelligence humaine). Un HIT représente une tâche virtuelle unique et autonome pour laquelle vous, en tant que travailleur, recevez une récompense après l'avoir complétée. Veuillez prévoir 5 à 10 minutes pour terminer le tutoriel de base. Une fois celui-ci terminé, des modules de formation supplémentaires seront débloqués afin de perfectionner vos compétences en annotation. Activez le son !

Le jeu en ligne Suspicious Behavior développé par les artistes Kairus révèle le travail humain invisible, qui alimente la manière dont les systèmes de vision par ordinateur « intelligents » interprètent nos actions. Grâce à une mise en situation, avec un bureau physique et un tutoriel d'étiquetage d'images, l'utilisateur expérimente concrètement le travail fastidieux des annotateurs externalisés. Dans le cadre d'un tutoriel interactif simulant une entreprise fictive, l'utilisateur est sollicité pour effectuer l'étiquetage des comportements jugés suspects.

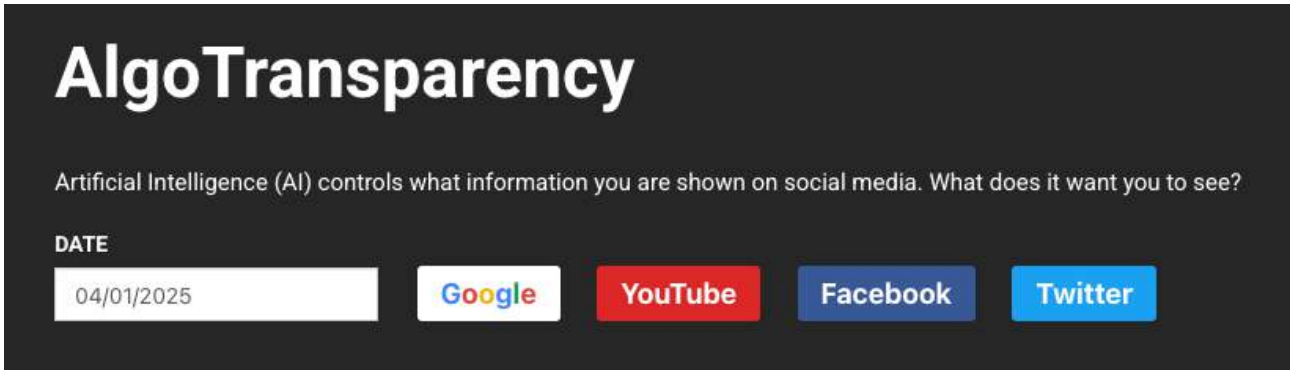
EN SAVOIR PLUS SUR LES BIAIS DANS LES ALGORITHMES DE RECRUTEMENT



Capture d'écran du jeu

Survival of the Best Fit est un jeu éducatif sur les préjugés dans le recrutement liés à l'IA. Il montre comment une utilisation inappropriée de l'IA peut conduire les machines à reproduire les préjugés humains et à accentuer les inégalités. Une grande partie du débat public sur l'IA la présente comme une menace qui nous est imposée, plutôt que comme un outil sur lequel nous conservons une part de maîtrise. Les concepteurs du jeu ont voulu changer cela en aidant le public à comprendre la technologie et en exigeant davantage de responsabilité de la part de ceux qui développent des systèmes logiciels de plus en plus omniprésents. Le projet a été développé par Gabor Csapo, Jihyun Kim, Miha Klasinc et Alia ElKattan et soutenu par le Creative Media Award de la Fondation Mozilla. Ce jeu vise à toucher des personnes qui ne sont pas les concepteurs de la technologie, mais qui en ressentent directement les effets dans leur vie quotidienne. Il est important de faire comprendre aux publics comment fonctionne l'IA et comment elle peut les affecter, afin qu'ils puissent mieux exiger la transparence et la responsabilité des systèmes qui prennent un nombre croissant de décisions à leur place.

COMMENT FONCTIONNE L'ALGORITHME DE RECOMMANDATION DE YOUTUBE ?



Capture d'écran du site web

Deux milliards de personnes utilisent YouTube chaque mois. L'algorithme de recommandation de YouTube détermine ce que les gens regardent dans plus de 70 % des cas. Cela représente 700 millions d'heures, soit l'équivalent de 1 000 vies humaines, chaque jour. Le projet AlgoTransparency observe plus de 800 chaînes d'information de premier plan, notamment des chaînes d'actualités, des vlogueurs et des émissions de tous horizons politiques.

Financé par l'Union européenne. Les vues et opinions exprimées n'engagent que leur (s) auteur(s) et ne reflètent pas nécessairement celles de l'Union européenne ou de l' Agence Erasmus+ France / Education Formation. Ni l'Union européenne ni l'autorité chargée de l'octroi ne peuvent en être tenues pour responsables

Guide créé par le projet Algo->Lit, financé par l'agence Erasmus +, coordonné par Datactivist et développé en collaboration avec les organisations Waag Futurelab, FARI - AI for the common good institute et la Mednum.
Crédit de l'image de couverture : Bart Fish & Power Tools of AI / https://betterimagesofai.org / https://creativecommons.org/licenses/by/4.0/
Première publication en septembre 2025.
Publié sous licence [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/deed.fr).

Pour toute question, n'hésitez pas à nous contacter à l'adresse suivante : algolit@datactivi.st.
Pour en savoir plus sur le projet Algo->Lit, rendez-vous sur notre [site web](https://www.algolit.eu/).

Les vues et opinions exprimées n'engagent que leur(s) auteur(s) et ne reflètent pas nécessairement celles de l'Union européenne ou de l'Agence Erasmus+ France / Education Formation. Ni l'Union européenne ni l'autorité chargée de l'octroi ne peuvent en être tenues pour responsables.

Financé par l'Union européenne