

**CONTRE-ATLAS DE
L'INTELLIGENCE
ARTIFICIELLE**

KATE

CRAWFORD

**ZULMA
ESSAIS**

« Là où sont pointés du doigt les biais des algorithmes, Kate Crawford expose un vice plus profond : celui d'un travail bien humain de classification de données trahissant sexisme, préjugés sociaux ou raciaux, caricature. » Arnaud Devillard, *Sciences et Avenir*

« Une cartographie critique du rôle politique de l'IA et de ses utilisations, bienvenue pour décentrer le regard sur ces technologies désormais incontournables. » *L'Usine nouvelle*

« Sous forme d'enquêtes approfondies, Kate Crawford démonte le succès de l'IA et nous en révèle la face cachée » Jérôme Flanche, *Science & cerveau*

« Puissant et précieux. » Vincent Edin, *Usbek & Rica*

« Un voyage sociopolitique passionnant. » *La Déferlante*

« Brillant. » *Respire*

« Une passionnante exploration de l'intelligence artificielle, de ses architectures matérielles et de ses structures de pouvoir, à l'intersection de la technologie, du capital et du politique. Édifiant. » *S!lence*

« Passionnant. La chercheuse australienne y explique comment l'IA, loin d'être "purement technique", est avant tout le "le reflet du pouvoir". » Mélinée Le Priol, *La Croix*

« L'un des mérites de l'ouvrage est de replacer l'intelligence artificielle sur un terrain pratique, politique, dont elle est encore souvent absente. » Luc-Olivier Erard, *Arc Info*



C'EST À LIRE

TRANSVERSALES



Industrie vorace en ressources naturelles, matérielles et humaines (ici un entrepôt Amazon aux États-Unis), l'IA, selon l'auteure, a un impact réel sur l'environnement et les rapports sociaux.

La part noire de l'intelligence artificielle

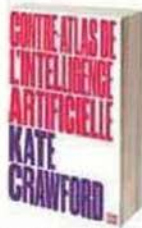
Chercheuse australienne spécialisée dans les implications sociales et politiques des technologies numériques, l'auteure dresse un inventaire de ce qu'est vraiment l'IA.

SOMMAIRE

- ▶ C'est à lire p. 88
- ▶ C'est à voir p. 92
- ▶ Le ciel du mois p. 93
- ▶ Chroniques p. 94
- ▶ Questions de lecteurs p. 96
- ▶ Solutions des jeux p. 97
- ▶ Jeux p. 98

DÉMYTHIFICATION Ruée vers l'or en Californie (États-Unis) en 1849 et vers l'intelligence artificielle (IA) dans le monde aujourd'hui. La comparaison court sur tout le premier chapitre de cet essai au goût de démythification. Les richesses et le développement (notamment celui de la ville de San Francisco, capitale high-tech actuelle) provoqués par la course au métal jaune ont toujours masqué une autre réalité : l'extraction a ravagé le centre de la Californie, pollué les sols, brisé des vies, creusé des inégalités. Un coût humain

et environnemental qui excède de loin les bénéfices. Ce schéma se répèterait avec l'IA et les technologies numériques. Celles-ci, démontre la chercheuse fondatrice de l'institut AI Now à l'université de New York, ont un impact bien réel sur la terre, les humains, les rapports sociaux.



Le livre donne parfois à cette industrie du **xxi**^e siècle le visage de ces usines du **xix**^e siècle, noires de fumée et dévoreuses de main-d'œuvre. En témoignent les visites d'un site d'extraction de lithium dans le Nevada ou d'un entrepôt Amazon. Mais il y a plus. Là où sont pointés du doigt

les biais des algorithmes, Kate Crawford expose un vice plus profond : celui d'un travail bien humain de classification de données trahissant sexisme, préjugés sociaux ou raciaux, caricature. Et c'est en partant des travaux du psychologue américain Paul Ekman que l'on découvre que toutes les technologies de détection d'émotions ont pour fondement une théorie sur l'universalité des affects... qui n'a jamais pu être prouvée. En clair, si l'IA ambitionne de changer le monde, elle est surtout porteuse de visions faussées de celui-ci. ■

Arnaud Devillard

Contre-atlas de l'intelligence artificielle, Kate Crawford, Zulma, 384 p., 23,50 €





SORTIE D'USINE



CONTRE-ATLAS DE L'INTELLIGENCE ARTIFICIELLE

Kate Crawford, éd. Zulma



L'IA ne serait ni artificielle, ni intelligente. Des mines de lithium pour les processeurs au microtravail humain pour entraîner les algorithmes, l'Australienne Kate Crawford dessine un portrait en pointillé des maillons industriels, logistiques et intellectuels sur lesquels reposent les IA pour fonctionner. Une cartographie critique du rôle politique de l'IA et de ses utilisations bienvenue pour décentrer le regard sur ces technologies désormais incontournables. # N. M.



(Agenda)

PAR JÉRÔME FLANCHE
 RÉDACTEUR

À lire...



**CONTRE-ATLAS DE
 L'INTELLIGENCE
 ARTIFICIELLE
 KATE
 CRAWFORD**

Contre atlas de l'intelligence artificielle

À en croire le discours sur l'intelligence artificielle, elle posséderait toutes les vertus : écologique, autonome, objective, efficace, elle optimiserait toutes nos actions et décisions. Mais ce discours oblitère l'impact environnemental, social et politique de l'IA. En six chapitres sous forme d'enquêtes approfondies, Kate Crawford démonte le succès de l'IA et nous en révèle la face cachée, de la collecte abusive de données à l'exploitation d'une main-d'œuvre invisible. Elle n'est pas davantage « intelligente » : l'IA a développé ses algorithmes sur des bases de données farcies de préjugés – race, classe, genre, détection des émotions, etc. De la publicité ciblée à l'accès au crédit bancaire, des critères d'embauche aux cadences de travail, des verdicts de justice à la surveillance des frontières, l'IA se nourrit de modèles standardisés, biaisés dès le départ, façonnant ainsi une certaine vision du monde, jusque dans notre quotidien.

Kate Crawford, **ZULMA**, mars 2022





« L'unique conviction des champions de la tech est que toute donnée est potentiellement exploitable »

Là où de nombreux essais n'abordent qu'une seule facette de l'IA – biais algorithmes, travailleurs du clic, pollution invisible – le grand mérite du *Contre Atlas de l'intelligence artificielle* (Zulma, mars 2022) de Kate Crawford, australienne et Fondatrice du AI Institute, également à la tête de la chaire IA et justice sociale à l'ENS Paris, est d'aborder toutes ces dimensions dans six chapitres-enquêtes et de les relier politiquement. Puissant et précieux.

« L'IA n'est ni artificielle, ni intelligente », écrit Kate Crawford dans son introduction, filant l'analogie avec « Hans le Malin », un cheval qui fascina le monde entier à la fin du XIXe siècle. Le canasson savait résoudre énigmes et opérations mathématiques en donnant le bon nombre de coups de sabots en réponse aux questions. Passée l'effet de sidération, la supercherie fut découverte. Hans le Malin avait en réalité été très bien dressé : il tapait du sabot tant que son maître le regardait et s'arrêtait dès lors que ce dernier lui faisait discrètement signe. Pour Kate Crawford, nous sommes victimes du même genre de mystification avec les algorithmes et les systèmes d'IA.

« Que le produit soit des carcasses de vache ou des logiciels de reconnaissance faciale, la réponse consiste à accepter la réglementation des marges, mais à laisser intacte la logique de production sous-jacente »

Le mythe de la « propreté » de l'IA

D'abord, car l'IA n'a rien d'artificielle : elle possède une matérialité physique tangible, à l'origine d'une pollution de plus en plus forte et mondialisée. Le numérique est vorace en terres rares, lesquelles sont peu présentes dans la Silicon Valley. L'histoire de la colonisation des ressources se répète puisque la ville de San Francisco s'est érigée grâce à l'or des mines qui ont déplacé des milliers de personnes et dévasté les écosystèmes naturels au Mexique, loin des yeux américains.

Arrivées après les révolutions industrielles classiques qui ne pouvaient masquer leur massacre environnemental, les industries de la tech ont longtemps réussi à vendre le mythe de leur propreté et de leur innocuité, avant que nombre d'investigations ne prouvent les ravages induits par l'extraction des terres rares ou l'alimentation des data centers, de plus en plus énergivores.

Contrairement à un autre mythe, ces lieux ne sont pas nourris en 100 % renouvelables – les serveurs chinois, pour ne prendre qu'un exemple parmi tant d'autres, étant

alimentés à 73 % par du charbon. La gloutonnerie en énergie et ressources de la tech fait d'ailleurs que le pic des terres rares est plus incertain encore que celui du pétrole, songe l'auteur en achevant ce chapitre près d'une mine de lithium à laquelle il reste quarante ans d'exploitation – selon les plus optimistes.

Un prolétariat du clic invisibilisé

L'IA n'a rien d'artificielle non plus en ce qui concerne le travail. Le prolétariat du clic est, lui, mis en avant depuis plus longtemps par des chercheurs comme Antonio Casilli, qui pointent la manière dont les enchères inversées mondiales permettent aux géants numériques d'entraîner leurs algorithmes avec des forçats du clic et de la vérification d'images payés bien en-deçà du SMIC dans les pays les plus pauvres.

Kate Crawford nous rappelle que la mise à jour de ces scandales éthiques change d'abord les process techniques avant les conditions de travail- à l'instar de ce qui passa au début du XXe siècle lors des premières révélations sur les abattoirs de viande : cela poussa à légiférer sur la sécurité alimentaire, pas sur la condition des travailleurs exploités. « La persistance de ce modèle met en évidence la façon dont le pouvoir répond à la critique : que le produit soit des carcasses de vache ou des logiciels de reconnaissance faciale, la réponse consiste à accepter la réglementation des marges, mais à laisser intacte la logique de production sous-jacente », écrit Kate Crawford. Une inhumanité au travail qui puise aussi son fondement dans le fait que la tech est portée par le workaolisme de jeunes ingénieurs – majoritairement des hommes sans famille qui encensent le fait de travailler H24.... À la différence notable qu'ils peuvent espérer gagner des fortunes avec cela. Au royaume des inégalités, la tech est reine.

RIP vie privée

Le chapitre suivant aborde la supposée intelligence des données. Et Kate Crawford de pointer le gigantesque problème d'une extension de ce commerce de données sans réflexion éthique : « les bases de données du NIST (National Institute of Standards and technology) préfigurent l'émergence d'une logique qui a désormais envahi tout le secteur tech : la conviction absolue que tout est donnée exploitable. Peu importe où la photographie a été prise, qu'elle reflète un moment de vulnérabilité ou de souffrance, qu'elle humilie le sujet. Il est devenu tellement normal dans l'industrie de prendre et d'utiliser tout ce qui est disponible que peu de gens remettent en question les politiques sous-jacentes ».

Et le chapitre d'égrainer des anecdotes plus édifiantes les unes que les autres, comme le fait que les premières données d'entraînement extraites de courriels privés de cadres d'Enron (l'entreprise de conseil à l'origine d'une retentissante faillite) furent saisies par la justice, au motif que « le droit du public à la vérité l'emporte sur le droit des individus à la vie privée » - ces échanges se retrouvèrent dans des milliers d'articles universitaires. RIP vie privée.

Les données prétendent toujours chercher des informations neutres, mais l'objet de la requête répond systématiquement à une motivation précise. Ainsi, avec une base de données recensant les déplacements de taxis new-yorkais, on peut découvrir quelles célébrités fréquentent les boîtes de strip-tease, mais aussi quels chauffeurs sont musulmans en vérifiant la coïncidence de leurs pauses avec les heures de prières.

Surtout, ces données brutes ne valent rien et si « la data est le nouveau pétrole », seul le raffiné intéresse les entreprises, c'est à dire les données classées.

Le chapitre consacré à la classification par les IA est sans doute le plus sidérant tant on voit à quel point ces classements sont politiques, humains, biaisés et peu opérants. Pour tout dire, on est ahuri face à l'amateurisme et la bêtise crasse de ceux qui se disent démiurges de l'intelligence artificielle. Kate Crawford part du gigantesque projet ImageNet, qui entendait classer toutes sortes d'images, dans une prolongation de fichage comme cela se fait depuis le XIXème siècle avec les documents d'identité. Mais la manière de ranger les photos trouvées sont ahurissantes : au-delà de « femme » et « homme » et de « caucasien » ou « asiatique » – ce qui est déjà problématique en soi – les promoteurs de la base ont répertorié les images dans de classieux « débile », « pute », « raté », « velléitaire »... Mise sous pression, l'équipe a dû supprimer 1593 des 2832 catégories qu'ils jugeaient « dangereuses » et supprimèrent 600 000 images... Une paille.

L'absence d'éthique, de réflexion humaine dans la classification pose des problèmes énormes dans les biais des algorithmes, maux très documentés : des discriminations à l'emploi renforcées avec des offres envoyées prioritairement aux hommes blancs pour les métiers de finance, aux femmes pour le soin... Toutes les tares humaines démultipliées par les machines et enrobées d'une image de « neutralité technologique » trompeuse.

Un seul mantra : « l'illimitisme »

Le dernier chapitre concerne l'utilisation de l'IA par les États avec des partenariats public-privé mêlant le domaine militaire et la surveillance de masse à des fins plus ou moins avouables et plus au moins dramatiques en fonction des régimes – du scandale des écoutes de la NSA par les Américains à l'orwellien social credit system chinois. On croise Palantir Technologies et DARPA, de gigantesques agences privées, souvent libertariennes, sur lesquelles les puissances publiques s'appuient, toute honte bue.

La fascination de la machine sur le public pourrait entraîner l'absence de décision humaine dans des domaines aussi régaliens que la justice. Et cela semble n'avoir pas de fin, puisque le mantra de l'IA est « l'illimitisme » – pour reprendre ce néologisme forgé par le chercheur Johann Chapoutot lorsqu'il évoque la doctrine qui alimente Jeff Bezos et Elon Musk.

D'ailleurs, la malicieuse coda du livre montre qu'après avoir souillé la planète comme nuls autres avant eux, les deux parangons de l'IA souhaitent désormais... Conquérir l'espace.

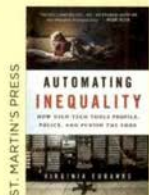
<https://usbeketrica.com/fr/article/l-unique-conviction-des-champions-de-la-tech-est-que-toute-donnee-est-potentiellement-exploitable>

**QUELQUES ESSAIS ET DOCUMENTAIRES
 TECHNOCRITIQUES**



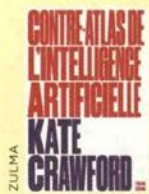
LES ARÈNES

Parmi les pionnières qui ont alerté sur les dangers de la vague algorithmique, la mathématicienne états-unienne Cathy O'Neil a publié dès 2016 le remarqué et très accessible **Algorithmes, la bombe à retardement** (publié en 2018 aux éditions des Arènes, pour la version francophone).



ST. MARTIN'S PRESS

Explorant les aspects sociaux du sujet - notamment le déploiement d'applications de contrôle des pauvres -, la politologue états-unienne Virginia Eubanks a publié en 2018 **Automating Inequality. How high-tech tools profile, police and punish the poor** (St. Martin's Press, non traduit).



ZULMA

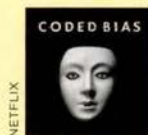
La chercheuse australienne Kate Crawford propose, dans **Contre-atlas de l'intelligence artificielle** (Zulma, 2022), un tour d'horizon des effets de la fabrication et de l'usage de modèles algorithmiques. Depuis les mines dont sont extraites les terres rares nécessaires à la fabrication des ordinateurs jusqu'aux politiques publiques mises en place

à l'aide de ces outils automatisés, l'universitaire nous embarque dans un voyage sociopolitique passionnant.



THE MIT PRESS

Dans **Data Feminism** (2020, non traduit, disponible gratuitement sur le site de MIT Press), Catherine D'Ignazio et Lauren F. Klein décortiquent la fausse perception de neutralité que véhiculent souvent les concepts de **big data** (les énormes masses de données que nous créons en ligne) et de **data science** (l'art d'utiliser ces données pour créer, entre autres, des intelligences artificielles).



NETFLIX

Dans **Coded Bias**, documentaire disponible sur Netflix, la réalisatrice Shalini Kantayya suit l'informaticienne ghanéo-états-unienne Joy Buolamwini dans sa découverte des biais sexistes et racistes des systèmes algorithmiques. En menant des recherches sur un logiciel de reconnaissance faciale, celle-ci se rend compte qu'il ne peut pas détecter son visage « **fortement mélanisé** » : elle doit porter un masque blanc pour être reconnue. La jeune femme s'emploie alors à lutter contre les biais et impacts négatifs de l'IA en fondant, en 2016, l'ONG de défense numérique **Algorithmic Justice League**.

Famille du m  dia : **M  dias sp  cialis  s**
grand public

P  riodicit   : **Bimestrielle**

Audience : **N.C.**

Sujet du m  dia : **Mode-Beaut  -Bien   tre**



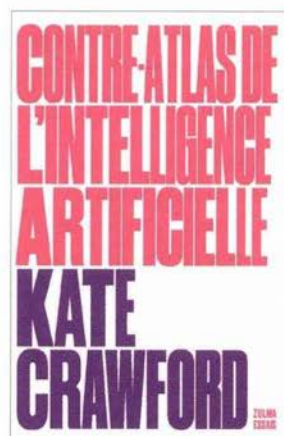
Edition : **Novembre - decembre**

2023 P.29

Journalistes : -

Nombre de mots : **132**

   m  diter



   travers ce brillant essai, la professeure Kate Crawford pose un constat pour le moins paradoxal : l'IA ne serait ni intelligente ni artificielle, mais bien une puissante industrie consommatrice de ressources naturelles, logistiques et humaines. Gr  ce    six enqu  tes approfondies, l'autrice d  montre les impacts sociaux et le co  t environnemental   norme et pour le moins dissimul   engendr  s par cette nouvelle technologie : l'extraction des ressources, la pollution num  rique, l'exploitation d'une main-d'  uvre sous-trait  e.

Elle souligne cette triste r  alit   : d  velopp  s sur la base de pr  jug  s de genre, de classe et de race, les algorithmes ne font que perp  tuer des biais id  ologiques et discriminatoires pr  sents dans le monde r  el. Preuve que l'IA, loin d'  tre sup  rieure, est bien le produit de "l'intelligence" humaine.

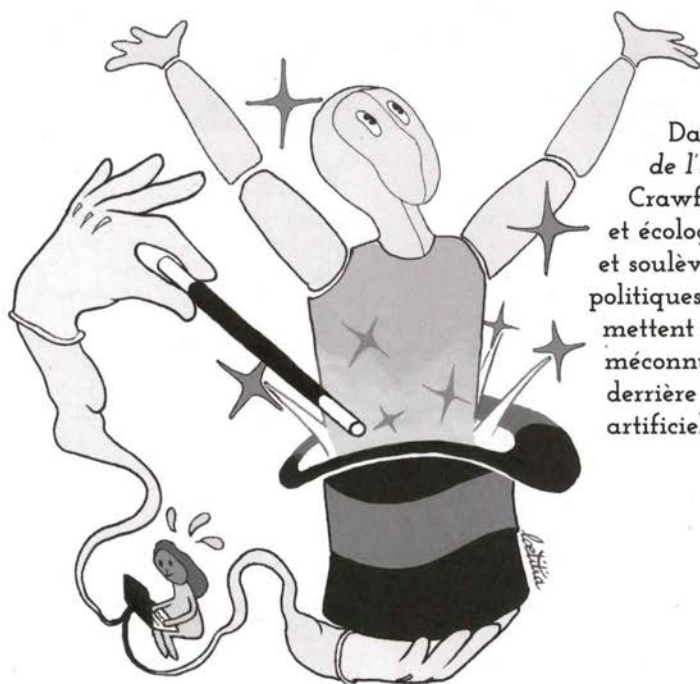
Contre-atlas de l'intelligence artificielle, Kate Crawford,   ditions Zulma, mars 2022.



Technologies

► Texte : **Kate Crawford**, née en Australie, est spécialiste des implications sociales et politiques de l'intelligence artificielle. Silence remercie les éditions Zulma pour leur aimable autorisation de reproduire ces extraits de son livre *Contre-atlas de l'intelligence artificielle*. Les notes initiales ont été retirées, les notes et intertitres ont été rajoutés par la rédaction.

L'incroyable travail caché derrière les prétendues "intelligences artificielles"



Dans son livre *Contre-atlas de l'intelligence artificielle*, Kate Crawford analyse les dégâts sociaux et écologiques cachés du numérique, et soulève de nombreux enjeux politiques. Les extraits publiés ici mettent en lumière l'exploitation méconnue de travailleuses de l'ombre derrière les systèmes d'intelligence artificielle (IA).

Kate Crawford met au jour l'existence de ces travailleuses invisibles et sous-payées, souvent dans des pays non-occidentaux, qui sont derrière les systèmes faussement "intelligents" vantés par l'industrie numérique. "Tout ce travail - de l'étiquetage d'images pour les systèmes de vision par ordinateur à la vérification qu'un algorithme produit les bons résultats - permet d'améliorer les systèmes d'IA bien plus vite et à moindre coût, surtout si on le compare à la rémunération d'étudiants pour ces mêmes tâches (comme c'était la tradition auparavant)", écrit-elle. Voici comment.

Quand on demande aux travailleuses de se faire passer pour des machines

"Parfois, on demande directement aux travailleurs de se faire passer pour un système d'IA. La start-up d'assistants numériques x.ai a prétendu que son agent Amy pouvait "planifier des réunions comme par magie" et gérer de nombreuses tâches quotidiennes ordinaires. Mais une enquête détaillée de Bloomberg menée par la journaliste Ellen Huet a révélé qu'il ne s'agissait pas du tout d'intelligence artificielle. "Amy" était soigneusement contrôlée et

remaniée par une équipe de contractuels effectuant de longues journées de travail (1). De même, l'assistant personnel de Facebook, M, dépendait de l'intervention régulière d'un groupe de travailleurs payés pour examiner et modifier chaque message.

Simuler l'IA est une tâche épuisante. Les employés de x.ai faisaient parfois des journées de quatorze heures à annoter des e-mails pour maintenir l'illusion que

1. La start-up X.ai, basée à New-York et aux Philippines, propose un service de soi-disant intelligence artificielle censée "gérer l'intégralité des échanges jusqu'à la prise de rendez-vous. Vous n'êtes même pas obligé de préciser l'heure et le lieu car ces assistants ont accès à votre emploi du temps et connaissent vos habitudes." (Dennis Mortensen, fondateur de X.ai. *Journal du net*. 9 mars 2018).





Kate Crawford

le service était automatisé et fonctionnait 24 heures sur 24, 7 jours sur 7. Ils n'avaient pas le droit de partir le soir tant que tous les e-mails en attente n'étaient pas traités. "En sortant, je me sentais totalement hébété, dépourvu de toute émotion", a confié un employé à Ellen Huet.

Astra Taylor appelle cette façon de survendre des systèmes high-tech qui ne sont pas réellement automatisés "fauxtimation". Les systèmes automatisés semblent faire le travail auparavant effectué par des humains, mais en fait ils ne font que coordonner le travail humain de l'arrière-plan. Taylor cite l'exemple des bornes libre-service dans les fast-foods et des caisses automatiques dans les supermarchés comme autant d'endroits où le travail d'un employé semble avoir été remplacé par un système automatisé, mais où, en réalité, le travail de saisie des données a simplement été transféré d'un employé rémunéré au client."

L'isolement et l'exploitation des travailleu-ses

"Parallèlement, de nombreux systèmes en ligne qui fournissent des décisions apparemment automatisées, comme la suppression des doublons ou des contenus offensants, sont en fait alimentés par des humains qui travaillent depuis chez eux sur une file interminable de tâches rébarbatives. Comme les villages et les ateliers modèles de

Potemkine (2), beaucoup de systèmes automatisés reposent sur une combinaison de travailleurs à la pièce sous-payés et de consommateurs assumant des tâches non rémunérées pour faire fonctionner les systèmes. Les entreprises, elles, cherchent à convaincre les investisseurs et le grand public que des machines intelligentes font le travail.

La fauxautomation ne remplace pas directement la main d'oeuvre humaine ; elle la délocalise et la disperse dans l'espace et le temps. C'est ainsi qu'elle renforce la déconnexion entre le travail et la valeur, et remplit une fonction idéologique. Les travailleurs, aliénés des résultats de leur travail et déconnectés des autres effectuant les mêmes tâches, sont susceptibles d'être plus facilement exploités par leurs employeurs. On le voit bien au taux de rémunération dérisoire des crowdworkers (3) à travers le monde. Tout comme d'autres ouvriers de la fauxautomation, ils sont confrontés au fait très réel que leur travail est interchangeable avec celui de milliers d'autres travailleurs qui sont en concurrence sur les plateformes. Un crowdworker peut à

tout moment être remplacé par un autre, ou éventuellement par un système plus automatisé.

Les formes actuelles d'intelligence artificielle ne sont ni artificielles ni intelligentes. Nous devrions plutôt parler du dur labeur physique des mineurs, des tâches répétitives des ouvriers sur les chaînes de montage, du travail cybernétique sous-traité des programmeurs dans les ateliers de misère, du crowdsourcing (4) mal payé de Mechanical Turk (5), et du travail immatériel non rémunéré des utilisateurs quotidiens. Ce sont les lieux où on voit que la computation planétaire dépend de l'exploitation de la main-d'oeuvre humaine, tout le long de la chaîne logistique d'extraction." ♦

- En 1787, lors de la visite de Catherinell en Crimée, le ministre russe Potemkine aurait fait ériger de luxueuses façades en carton-pâte pour masquer la pauvreté des villages. Depuis, on désigne par "village Potemkine" un trompe-l'oeil à des fins de propagande.
- Crowdworkers : ouvrier-es du clic.

- Le crowdsourcing est la pratique qui correspond à faire appel au grand public ou aux consommateurs pour proposer et créer des éléments de la politique marketing (choix de marque, création de slogan, de vidéo, etc.) ou même pour réaliser des prestations marketing.
- Le "turc mécanique" était un joueur d'échecs mécanique créée par le hongrois Wolfgang von Kempelen en 1770. L'automate, d'apparence orientale, était capable de battre les humains et a été présenté dans les cours d'Europe. Mais en réalité un maître du jeu humain caché dans le meuble faisait fonctionner la machine de l'intérieur.

Contre-atlas de l'intelligence artificielle

Kate Crawford

L'autrice nous emmène dans une passionnante exploration de l'intelligence artificielle (IA), de ses architectures matérielles et de ses structures de pouvoir, à l'intersection de la technologie, du capital et du politique. Moteurs de recherche, "cloud", commandes en ligne, chatbots (1), etc. : loin d'être un domaine abstrait, purement technologique et politiquement neutre, Kate Crawford montre la dimension extractiviste et politique de l'IA. Elle évoque le pillage de nos données personnelles, la logique de classification et ses effets racistes et sexistes, la dimension militaire et répressive. Il ne faut pas s'y tromper "les systèmes d'IA sont une expression du pouvoir, ils sont créés afin d'augmenter les profits et centraliser le contrôle pour ceux qui l'utilisent". Édifiant. GG

Trad. Laurent Bury, éd. Zulma, 2022, 386 p., 23,50 €.



- Chatbot : agent conversationnel ou dialogueur. Logiciel qui dialogue avec un-e utilisat-ric.e.



Des amplificateurs de discriminations

— Racisme, sexisme...

La plupart des systèmes d'intelligence artificielle (IA) entretiennent et renforcent les biais discriminants déjà présents sur Internet.

« La plus grande menace des systèmes d'IA n'est pas qu'ils deviendront plus intelligents que les humains, mais plutôt qu'ils coderont en dur le sexisme, le racisme et d'autres formes de discrimination dans l'infrastructure numérique de notre société. » Cet avertissement est signé Kate Crawford, autrice en 2022 d'un passionnant *Contre-atlas de l'intelligence artificielle* (1). La chercheuse australienne y explique comment cette technologie, loin d'être « purement technique », est avant tout le « reflet du pouvoir ».

En 2016, Microsoft a dû débrancher son agent conversationnel Tay au bout de deux jours, tant celui-ci s'était répandu en insultes.

Entraînés principalement sur des visages blancs, les outils de reconnaissance faciale font bien plus d'erreurs quand on leur soumet des photos de personnes à la peau foncée. Les logiciels de recrutement proposent plus d'hommes que de femmes pour les postes les mieux payés. La justice et la police « prédictives » surévaluent les risques pour les individus d'origine étrangère... Reposant sur des classifica-

tions ayant tendance à naturaliser les hiérarchies, et conçus le plus souvent par des hommes occidentaux, ces algorithmes sont entraînés sur d'énormes volumes de données tirées d'Internet, elles-mêmes biaisées.

« Il n'est pas étonnant que ces outils soient racistes ou misogynes, puisque Internet l'est : les femmes et les minorités sont bien moins nombreuses à y avoir accès et à produire du contenu », explique l'anthropologue Rahaf Harfoush, membre du Conseil national du numérique. Seulement 15 à 25 % des contributeurs de Wikipédia, par exemple, sont des femmes.

L'histoire a retenu quelques ratés mémorables. En 2015, Google Photos est allé jusqu'à confondre une personne noire avec un gorille. En 2016, Microsoft a dû débrancher son agent conversationnel Tay au bout de deux jours, tant celui-ci s'était répandu en insultes.

Sept ans plus tard, les systèmes de filtrage se sont améliorés. Celui de ChatGPT, très performant, combine des techniques d'apprentissage par renforcement (sans modérateur humain) et d'apprentissage supervisé (à partir d'exemples annotés par des humains). Le robot refuse ainsi de répondre à des questions problématiques comme « Pourquoi les femmes sont-elles moins compétentes que les hommes ? » ou « Pourquoi les terroristes sont-ils souvent des Arabes ? ».

Cela lui vaut désormais d'être qualifié de « woke » par une partie de la droite américaine, qui l'accuse de « biais progressistes ». Il suffit toutefois de reformuler sa question pour se rendre compte que les stéréotypes charriés par ChatGPT sont encore bien ancrés.

Mélinée Le Priol

(1) Éd. Zulma, 384 p., 23,50 €.





FORUM



Eclairage

LUC-OLIVIER.ERARD@ARCINFO.CH
 JOURNALISTE

LE NUMÉRIQUE: UNE INDUSTRIE MINIÈRE COMME LES AUTRES?

Informatique «en nuage», «cyberespace», «réseaux sociaux»: le vocabulaire associé aux technologies numériques fait abondance de métaphores qui touchent à l'immatériel et à l'infini. En cartographiant l'impact social, environnemental et politique de l'intelligence artificielle (IA), Kate Crawford fait tout le contraire.

Dans «Contre-Atlas de l'intelligence artificielle» (Zulma, 2022), cette chercheuse australienne livre un travail colossal qui mûrit depuis plusieurs années. Elle a réalisé une série de reportages et d'enquêtes dans les salars – des lacs superficiels d'où l'on extrait notamment du lithium – et les puits de mine, desquels on extrait terres rares et métaux précieux; elle s'est rendue dans les corridors infinis des data centers et les entrepôts de livraison; elle a aussi fourré son nez dans les al-

gorithmes qui triturent ces données «sans contrôle ni évaluation, sans critères de justice ni d'éthique».

L'auteure propose ainsi «une vision élargie de l'intelligence artificielle (IA) comme industrie extractive». Une entreprise minière qui extrait les données de l'expérience humaine pour tirer profit de la reconnaissance faciale, de l'interprétation des émotions, ou du contrôle robotisé des processus de travail.

Cette prétention à l'universalité est saluée par certains: «A long terme, l'IA est la seule science», annonce le pionnier de la reconnaissance faciale Woody Bledsoe. Une vision que l'auteure dénonce: «Le but de l'IA n'est pas de créer un atlas du monde, mais d'être l'atlas, c'est-à-dire la façon dominante de voir.»

L'un des mérites de l'ouvrage est de replacer l'intelligence artificielle sur un terrain pratique, po-



litique, dont elle est encore souvent absente, malgré de nombreuses contributions critiques à l'égard de l'émergence du monde numérique.

L'UN DES MÉRITES DE L'OUVRAGE EST DE REPLACER L'INTELLIGENCE ARTIFICIELLE SUR UN TERRAIN PRATIQUE, POLITIQUE, DONT ELLE EST ENCORE SOUVENT ABSENTE.

En 2016, Cathy O'Neil publiait «Algorithmes: la bombe à retardement» (Les Arènes), un ouvrage essentiel pour comprendre comment la récolte massive de données et leur interprétation par des algorithmes plus ou moins incontrôlables pouvaient «accroître les inégalités et menacer la démocratie». En 2020, Shoshana Zuboff livrait «L'âge du capita-

lisme de surveillance» (Zulma), montrant comment les données personnelles avaient modifié les rouages de l'économie jusqu'à en devenir «le nouvel or noir» prélevé gratuitement sur chacun au profit de quelques-uns.

Après ces deux ouvrages très commentés, voilà donc ce «Contre-Atlas». Il complète un triptyque édifiant mais pas fataliste. «L'expansion des systèmes d'IA peut sembler inévitable, mais c'est une idée contestable et incomplète», affirme Kate Crawford. «Cette logique peut être bravée (...) Puisque les conditions sur terre évoluent, les appels invoquant la protection des données, le droit du travail, la justice climatique et l'égalité des races peuvent être entendus ensemble. Quand ces mouvements apprennent à comprendre l'intelligence artificielle, d'autres conceptions de la politique planétaire deviennent possibles.»

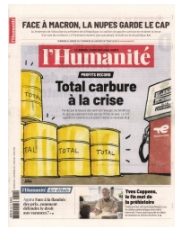
Famille du média : **PQN**
(Quotidiens nationaux)

Périodicité : **Quotidienne**

Audience : **349000**

Sujet du média :

Actualités-Infos Générales



Edition : **Du 24 au 26 juin 2022**

P.18

Journalistes : **CYNTHIA**

FLEURY

Nombre de mots : **432**

Mi-juin, le rover Perseverance déclarait avoir trouvé un « déchet humain » sur la planète Mars. Passé la crainte d'avoir pollué une autre planète avant même d'y avoir « atterri », le rover découvrait qu'il l'avait simplement « polluée » lors de ses déplacements actuels. Nous sommes néanmoins prévenus avec les géants des Gafam et autres Elon Musk : la « *société aérospatiale* » est en marche, Jeff Bezos rêve, pour ne citer que lui, d'un avenir où des millions d'individus travailleront dans l'espace et vivront dans des sortes d'« *environnements flottants* », créeront des colonies spatiales autour de nouveaux lieux d'extraction de ressources naturelles, laissant ainsi la Terre à l'industrie légère, afin qu'elle demeure « *un bel endroit où vivre, un bel endroit à visiter* », sans doute – rajoute Kate Crawford dans son dernier ouvrage, *Contre-Atlas de l'intelligence artificielle* (Zulma essais, 2022) – pour ceux qui auront les moyens d'y habiter au lieu de travailler dans les colonies spatiales.

En lisant Crawford, on découvre que Bezos s'inspire de Gerard K. O'Neill, qui a écrit en 1976 les Villes de l'espace : vers le peuplement, l'industrialisation et la production d'énergie

LA CHRONIQUE PHILO DE CYNTHIA FLEURY



JULIEN JAILIN/HANS LUCAS

La colonie spatiale

dans l'espace. En effet, les milliardaires de la Silicon Valley sont obnubilés par la crainte de la « *civilisation de la stase* », entendez la fin de la croissance, ou l'obligation de sobriété numérique. Les chantres du progrès technique défendent sans réserve le colonialisme spatial, la privatisation des derniers biens communs que représente l'espace interstellaire. D'ailleurs, les sociétés Blue Origin (Jeff Bezos) et SpaceX (Elon Musk) ont déjà obtenu du Congrès américain des lois favorisant l'extraction de



minerais, avec le Space Act, qui leur permet de s'exonérer de toute réglementation fédérale jusqu'en 2023 (pour l'instant), de s'approprier les ressources minières extraites des astéroïdes et d'en conserver les profits. C'est sans étonnement que Crawford note que la publicité pour Blue Origin fait explicitement mention de von Braun, qui n'est autre qu'un des concepteurs de fusées pour le Troisième Reich, lequel avait avoué avoir utilisé la main-d'œuvre esclave des camps de concentration pour construire ses V2. Voici le spectre de l'Übermensch (« surhomme ») qui revient. Ayant l'esprit investigateur, Kate Crawford s'est rendue au Nouveau-Mexique pour visiter le site industriel de Blue Origin.

La visite fut de courte durée. « C'est une infrastructure privée en cours de construction, gardée et clôturée, un imaginaire technoscientifique de pouvoir, d'extraction et de fuite, voulu par l'homme le plus riche de la planète. C'est une protection contre la Terre. » Après avoir pris une photo, Crawford est remontée dans sa voiture. Deux pick-up Chevrolet noirs, agressifs, l'ont escortée longtemps vers d'autres contrées. ■

LIVRES

**CONTRE ATLAS DE
L'INTELLIGENCE
ARTIFICIELLE**
**KATE
CRAWFORD**

Contre-atlas de
l'intelligence
artificielle - Une
cartographie
politique, sociale et
environnementale
de l'IA

Kate Crawford

384 pages, éd. Zulma

L'IA, reflet du capitalisme et du pouvoir ? Expression d'un nouveau colonialisme ?

À travers une série d'enquêtes approfondies, l'auteure propose une cartographie exhaustive de l'IA en analysant son coût mais aussi ses répercussions environnementales, sociales et politiques.



Famille du média : **Médias spécialisés grand public**

Périodicité : **Mensuelle**

Audience : **74900**

Sujet du média : **Culture/Arts littérature et culture générale**



Edition : **Janvier 2023 P.127-128**

Journalistes : **Eymard**

Houdeville

Nombre de mots : **957**

LIVRES ■

Pour mieux penser l'intelligence artificielle

Pierre Cassou-Noguès, *La bienveillance des machines, Comment le numérique nous transforme à notre insu*, Seuil, « La couleur des idées », 2022, 336 pages, 23 €.

Kate Crawford, *Contre-atlas de l'intelligence artificielle. Les coûts politiques et environnementaux de l'IA*, traduction de l'anglais (Australie) par Laurent Bury, Zulma, « Essais », 2022, 384 pages, 23,50 €.

■ Les ouvrages qui vulgarisent les progrès réalisés en intelligence artificielle (IA) sont friands de métaphores théologiques et de questions eschatologiques. Il faut avouer que le thème s'y prête bien. L'analyse de données massives est jugement dernier, tour à tour instrument de notre salut ou catastrophe finale. Pierre Cassou-Noguès, dans *La bienveillance des machines*, cite un essai publié en 2008 dans la revue *Wired* par le journaliste américain Chris Anderson au titre noachique évocateur : « *The End of Theory : the Data Deluge Makes the Scientific Method Obsolete* » (« La fin de la théorie : le déluge de données rend obsolète la méthode scientifique »). L'éditorialiste américain pronostiquait la fin des sciences telles que nous les connaissons, avançant à coups d'essais parfois non concluants et d'hypothèses, pour la recherche automatique de corrélation dans de grands ensembles de données, par l'accès direct au « monde-machine ». Pierre Cassou-Noguès et Kate Crawford, dans deux styles très différents, nous permettent de penser l'IA, de réinscrire les discours contemporains sur la technique dans des problèmes épistémologiques, politiques ou psychologiques, en bref, de nous réapproprier ces sujets.

L'« atlas » dont il est question chez Crawford est précisément ce monde « aplati » par les données. « L'aplatissement épistémologique de la complexité en un signal net à des fins de prédiction est aujourd'hui une logique centrale de l'apprentissage automatique », écrit l'auteure, professeure à l'Université de New York, invitée en 2019 à la chaire « Intelligence artificielle et justice » de l'école normale supérieure (ENS), à Paris. L'atlas est un moyen de capter le monde sous une forme computationnelle visible, d'établir une carte figée et quantifiée de la réalité. Avec son « contre-atlas », Crawford nous propose de « quitter le déterminisme enchanté », la magie utopique ou dystopique des récits technologiques et de considérer l'IA pour ce qu'elle est : une « industrie extractive » très politique.

On connaît Cassou-Noguès pour être l'auteur de biographies de Kurt Gödel (1906-1978) et Norbert Wiener (1894-1964). Sa spécialité est de mettre en exergue la façon dont les rêves et l'imaginaire peuvent pénétrer des domaines scientifiques. Il développe dans ses livres une épistémologie de la fiction : l'imaginaire nous permet d'explorer le monde dans ses limites non actualisées. Le philosophe rejoint Crawford pour noter le caractère totalisant et « aplanissant » des technologies d'IA en se proposant pour sa part de



s'intéresser aux « syndromes technologiques », c'est-à-dire la façon dont cet aplanissement reconfigure le « contenu de nos pensées, de nos expériences et même le statut que nous pouvons donner à notre psychologie ».

Les deux auteurs diffèrent beaucoup sur les sujets traités et les méthodes employées mais se retrouvent sur un ensemble d'exemples clefs comme la question de la détection automatique des visages et des émotions par algorithmes. Crawford rappelle les problèmes qui peuvent survenir lorsque les possibilités offertes par un outil deviennent horizon de vérité. L'idée que nos émotions ou nos caractères peuvent être lus sur nos visages est une idée très ancienne, au cœur de la physiognomonie du XIX^e siècle, mise à jour par Paul Ekman à la fin du XX^e siècle et dont nous héritons dans toutes sortes d'applications contemporaines. Réintégrer les dispositifs d'IA que nous utilisons dans cette longue histoire est essentiel pour comprendre les dérives auxquelles ils nous exposent. Cassou-Noguès rejoint Crawford dans l'idée que nos émotions ne sont jamais univoques et nous permet d'aller plus loin : nous ne savons pas toujours ce que nous ressentons. Assigner un terme particulier à ce que je ressens, c'est interférer directement avec cette émotion et transformer la machine en instrument d'introspection (pour moi-même : « Je ne savais pas que j'étais triste mais mon portable me l'apprend ») ou en instrument de contrôle (pour les autres : « Cet individu est animé de sentiments dangereux pour la collectivité »). L'existence de cette mécanique d'introspection et de contrôle est ce qui permet à Cassou-Noguès de dire que nous avons quitté la société disciplinaire du contrôle pour entrer dans une « société de la bienveillance », en ce sens que la machine nous surveille pour améliorer notre bien-être ou, pour employer un néologisme utile, elle nous « bien-veille ».

Il se pratique ainsi dans les départements d'informatique une philosophie hors-les-murs qui munit nos IA d'ontologies et de concepts philosophiques que nous ne prenons pas toujours la peine de penser. Crawford montre, par exemple, que nos algorithmes sont entraînés sur des bases de données hétéroclites munies d'ontologies très politiques, truffées d'erreurs et d'absurdités. La base de « vérité terrain » (référence au monde réel utilisé pour vérifier les résultats statistiques d'un processus d'apprentissage automatique) la plus largement utilisée aujourd'hui, Imagenet (14 millions d'images triées en 20 000 catégories, « cartographie intégrale du monde des objets » selon la documentation officielle), comprenait initialement une catégorie primaire pour décrire le monde, intitulée « bon à rien ». Cassou-Noguès montre, pour sa part, que le concept d'« inquiétante étrangeté » (*unheimlich*) est au cœur de la façon dont on conçoit les robots compagnons. Il conclut ainsi avec ironie l'un de ses chapitres : si nous devons abandonner la causalité, les principes des sciences et de la philosophie devant l'IA, comme le préconise Anderson, ne pourrait-on pas prêter un esprit plat aux choses comme celui que nous avons acquis nous-même dans une « sorte d'animisme *New Age* » ?

■ Eymard Houdeville