

PAR EMILE MARZOLF

25 octobre 2023, 15:19, mis à jour le 25 octobre 2023, 15:34

Camille Girard-Chanudet : "L'IA ne fait pas disparaître le travail, mais se contente de le déplacer"

Doctorante en sociologie à l'EHESS, Camille Girard-Chanudet s'est plongée dans le travail de conception d'un outil d'intelligence artificielle à la Cour de cassation. Son enquête de terrain rappelle combien l'IA n'a rien d'automatique, ni de magique. Elle repose au contraire sur le travail fastidieux des petites mains, les annotatrices, chargées d'aider la machine à saisir le monde réel dans toute sa complexité. du moins en partie. Entretien.



Vous avez récemment publié plusieurs articles sur le fruit de vos recherches sur le travail de l'intelligence artificielle*. Comment avez-vous été amenée à vous intéresser au développement de l'IA à la Cour de Cassation ?

J'ai commencé à m'intéresser plus généralement à l'intelligence artificielle dans la justice

afin de regarder, au-delà des discours, comment ces technologies étaient déployées ou non, notamment parmi les start-ups qui se sont multipliées à ce sujet dans la foulée de la loi pour une République numérique et de l'ouverture des décisions de justice. L'idée était de creuser, par un travail sociologique, ce qui se cache derrière la façade : étudier la manière dont ces IA sont conçues et déconstruire cette idée qu'il s'agit d'outils autonomes, objectifs et immatériels. C'est ainsi que je me suis penchée sur la Cour de cassation, l'institution judiciaire la plus avancée en la matière, mais j'ai aussi pu suivre le développement de l'algorithme d'évaluation des préjudices corporels Datajust par le ministère de la Justice par exemple, ou bien étudier les projets du conseil d'Etat.

Vous mettez en lumière, notamment, le rôle des annotateurs dans le développement de ces outils...

Tout à fait, et c'est justement ce qui explique que la Cour de cassation soit l'institution la plus avancée. Dans le cadre de son projet d'outil de pseudonymisation des décisions de justice en vue de leur publication en open data, elle s'est dotée d'une véritable équipe d'annotateurs, ou plutôt d'annotatrices, dont ne disposent pas les autres institutions judiciaires. Entre janvier et juin 2021, j'ai donc mené une enquête de terrain ethnographique au sein du service chargé du développement de cet outil pour rendre compte de ce processus de conception. A l'époque, certaines briques étaient déjà posées, mais toute la discussion sur les paramètres de l'algorithme étaient encore vives entre les développeurs, les datascientists, les magistrats et les annotatrices. J'ai ainsi pu me plonger dans les négociations et les arbitrages, entre d'un côté les possibilités techniques et la puissance de calcul forcément contraintes, et de l'autre les ambitions et le cadre du projet, davantage maîtrisé par les magistrats. Mais j'ai surtout orienté mon travail sur celui des annotatrices, car il s'agit du nerf de la guerre dès lors qu'on parle d'IA, et pourtant, il reste invisibilisé.

A quoi correspond le travail de ces petites mains ? En quoi est-il indispensable ?

Dans l'IA, il y a d'un côté les données, de l'autre des modèles, mais entre les deux, on retrouve des humains chargés d'aider la machine à lire ces données. A la Cour de

cassation, une vingtaine d'annotatrices menaient et continuent de mener ce travail minutieux, qui consiste à lire des décisions de justice les unes après les autres, et à signaler dans le texte toute donnée potentiellement identifiante : un nom, une adresse, un numéro de cadastre, etc. C'est grâce à l'annotation fastidieuse de dizaines de milliers de décisions de justice que le modèle d'apprentissage automatique est ensuite capable de s'entraîner à repérer puis masquer des éléments qui pourraient permettre d'identifier une personne, avant leur publication en accès libre. Ces mêmes annotatrices sont également chargées de vérifier que la machine a bien fait son travail. L'équipe n'a donc pas vocation à disparaître ni même à être réduite. Leur rôle reste primordial pour continuer d'aider l'IA à apprendre.

Que voulez-vous dire ?

Le monde réel n'est pas celui, figé et sans inventivité, de l'IA, il change constamment. De nouveaux mots, comme la visioconférence, apparaissent ou s'imposent, de nouveaux cas émergent. En continuant d'annoter de nouvelles décisions, la Cour de cassation permet à la machine d'appréhender la nouveauté. C'est d'ailleurs, entre autres choses, ce qui a fait défaut au projet Datajust : le ministère n'était pas prêt à recruter des annotateurs sur le long terme [**le développement de Datajust a été finalement abandonné**]. Les annotatrices permettent aussi de faire évoluer le système de catégorisation du monde réel retenu par les magistrats - en ajoutant de nouvelles catégories pour mieux le refléter -, mais aussi d'en révéler les limites, et les partis pris culturels et politiques. Lors de mon enquête, j'ai ainsi pu constater qu'elles étaient régulièrement confrontées à des casse-têtes, ne sachant pas dans quelle catégorie placer un élément. Par exemple, en Outre-mer, une décision mentionnait un nom de parcelle, sous forme de nom propre. L'annotatrice s'est finalement résolue à la rentrer de force dans la catégorie "numéro de cadastre". Ces choix, et parfois ces erreurs, ont néanmoins une incidence directe sur le développement de l'algorithme, qui ne fera que les mimer. Elles en étaient d'ailleurs parfaitement conscientes.

Pourquoi le travail d'annotation est-il selon vous invisibilisé ?

Comme je le disais, il s'agit d'un travail fastidieux, difficile à décrire, et peu valorisant. On parle de personnes, des agents de catégories C, simplement assises devant un ordinateur sur lequel défilent des décisions de justice. Cela ne correspond pas tout à fait à l'image que l'on se fait autour de la magie de l'IA. Ce que j'ai voulu montrer, et ce qu'a montré notamment Antonio Casilli avec l'externalisation de ces tâches à Madagascar, c'est que l'IA ne fait pas disparaître le travail, mais se contente de le déplacer, avec des conséquences très concrètes sur l'entraînement et donc le fonctionnement des algorithmes.

*Camille Girard-Chanudet, [« Le travail de l'Intelligence Artificielle : concevoir et entraîner un outil de pseudonymisation automatique à la Cour de Cassation »](#), RESET, mars 2023.
[« Mais l'algo, là, il va mimer nos erreurs ! »](#). *Contraintes et effets de l'annotation des données d'entraînement d'une IA* », Réseaux, vol. 240, no. 4, 2023, pp. 111-144.