

Aux algorithmes, citoyens !

par Marie Le Clainche-Piel

Comment peut-on créer des algorithmes qui soient aussi justes qu'efficaces ? David Robinson propose des pistes de réflexion et tire les conséquences d'une expérience collective.

À propos de : David Robinson, *Voices in the code. A story about people, their values, and the algorithm they made*, Russell Sage Foundation, 2022, 2024, 379 p.

En 2024, non seulement les algorithmes sont entrés dans le fonctionnement ordinaire des entreprises autant que des administrations publiques françaises, mais ces outils sont aussi mis en avant par ces organisations comme des marqueurs de leur modernité et de l'efficacité de leurs services. Que ce soient les algorithmes de reconnaissance faciale utilisés par Google et par les forces de l'ordre, les algorithmes de Parcoursup qui affectent les bacheliers à des places dans l'enseignement supérieur ou ceux de l'administration fiscale appelés à traquer des fraudes, peu de personnes et d'espaces de la vie sociale sont susceptibles d'y échapper.

Promus, critiqués ou mis en concurrence, les algorithmes ont perdu de leur aura mystérieuse. Ces suites d'instructions transmises à un ordinateur pour qu'il utilise des données et fournisse un résultat passent de l'imaginaire des insondables boîtes noires à celui des sacs de nœuds : sans savoir tout-à-fait comment ils sont agencés, les usagers et usagères en devinent le fonctionnement à tâtons et les chercheurs et chercheuses en sciences sociales en analysent les logiques sous-jacentes. C'est dans ce contexte de présence grandissante des algorithmes et d'exposition de leurs logiques que David Robinson publie *Voices in the Code. A Story About People, Their Values, And The Algorithm They Made*. Au lieu de décoder le fonctionnement de ces outils dans le cours de leur

implémentation, l'auteur propose d'enquêter sur ce qui se passe en amont de leurs usages : dans le processus de leur conception. Il cherche plus précisément à identifier les voies par lesquelles permettre au plus grand nombre d'influer sur leur fonctionnement. Comment intégrer « des voix dans le code », c'est-à-dire multiplier les voix au chapitre de la création des algorithmes, au nom de la maîtrise collective des logiques sociales et politiques sur lesquelles ils reposent ?

Le choix de cette démarche s'éclaire par le positionnement de David Robinson, chercheur et entrepreneur à la croisée des espaces disciplinaires. Initialement formé aux sciences politiques et au journalisme, l'auteur a co-fondé en 2011 un collectif – *Upturn* – qui promeut « l'équité et la justice dans le design, les usages et la régulation des technologies » avant de devenir directeur de l'organisme de formation *Apple University* de l'entreprise du même nom. Ce livre est à l'image de sa trajectoire puisqu'il cherche à faire tenir ensemble une posture d'alerte vis-à-vis du pouvoir des algorithmes sur tous les aspects de nos vies, et une ambition réformatrice quant à la capacité de créer des outils qui soient aussi justes qu'efficaces.

Algorithme = données + valeurs

Toute la logique du livre repose sur la considération selon laquelle les algorithmes sont constitués de deux catégories d'ingrédients clefs rarement repérés ensemble : des données, c'est-à-dire des chiffres extraits de la vie et des réalités humaines, et des valeurs morales. Le plus souvent, seuls les premiers ingrédients, les données utiles au fonctionnement des programmes, sont considérés comme constitutifs des algorithmes et les professionnels de la programmation comme les principaux acteurs de leur construction : ce sont ces derniers qui vont sélectionner des données et les organiser. Or toutes ces étapes techniques engagent des considérations morales. Traduire des informations de la vie humaine en catégories, sélectionner les données pertinentes, organiser les réponses au sein du programme : toutes ces étapes constituent en réalité des choix éthiques. Mises bout à bout elles peuvent servir des modèles de sociétés extrêmement variés, si ce n'est opposés. Pour le montrer, David Robinson se concentre tout au long du livre sur des cas sensibles dans lesquels le résultat produit par un algorithme affecte directement des trajectoires humaines. Ce sont typiquement des algorithmes qui répartissent des ressources en tension : un accès aux soins en médecine, un logement, une place dans une université ou dans un service d'aide à l'enfance. Dans la mesure où il n'y a pas assez de ressources pour toutes les

personnes en attente, comment construit-on un outil de décision à leur allocation ? Quel enfant extraire d'un cadre familial à risque ? À quelle personne en attente de greffe allouer un organe qui vient d'être rendu disponible ? À quel étudiant accorder une place dans une formation sélective ? La fabrique des outils algorithmiques d'aide à la réalisation de ces « compromis éthiques » est au cœur de sa réflexion.

Ces enjeux de traduction de choix moraux en décisions techniques ne sont bien sûr pas propres aux algorithmes ou à l'informatique – même s'ils se singularisent par leur automatisation. D'ailleurs, pour Robinson, s'interroger sur la meilleure manière de créer un algorithme à fort enjeu de société présente les mêmes défis que la création de lois. Prenons le cas le plus travaillé dans l'ouvrage, la répartition des reins prélevés sur des défunts en vue de réaliser une greffe d'organe. Lorsque le Congrès américain déclare que les organes doivent être attribués « selon des critères médicaux », il s'appuie sur un a priori similaire d'objectivité des catégories médicales. La dimension morale d'une telle déclaration est largement invisibilisée. Or, non seulement les critères médicaux les plus fins ne parviennent pas toujours à départager plusieurs patients en attente de greffe, mais en plus, les critères médicaux sont eux-mêmes chargés moralement. Choisir le patient pour lequel l'organe a le moins de risque d'être rejeté ou bien celui qui gagnera le plus d'années de vie, ou encore celui qui a le plus attendu sur liste d'attente : ce sont chaque fois des choix moraux que la technique médicale ne saurait trancher seule. Bureaucraties et lignes de code, même combat pour Robinson qui les définit comme des « dispositifs de conversion de problèmes moraux difficiles en problèmes apparemment techniques et ennuyeux » (p. 30).

L'auteur reconnaît d'ailleurs qu'en pensant à travers les algorithmes et en se focalisant sur les enjeux éthiques qu'ils contiennent, on se détourne pour partie des racines des problèmes qu'ils soulèvent. Les algorithmes sensibles sur lesquels se focalise l'ouvrage s'inscrivent en effet dans un contexte social plus large : ce sont des outils d'optimisation de la rareté et du manque de moyens. Que ce soit pour l'attribution de places dans des organismes d'accueil d'enfants en danger ou des organes à des patients sur liste d'attente d'une greffe, l'essor des algorithmes s'inscrit ici dans une logique de gestion de pénuries et de triage.

Faire entrer les algorithmes dans la vie démocratique

La délégation de la création des algorithmes à des équipes techniques repose ainsi sur l'illusion d'objectivité des données chiffrées, déjà bien étudiée en histoire et en sociologie par Theodore Porter, Olivier Martin ou Lorraine Dalston et Peter Galison. Mais Robinson s'inquiète surtout de son corollaire : à travers les chiffres et la transformation de données de la vie sociale en catégories programmables, des choix sont faits sans en avoir l'air, et cette délégation technique fait peser sur les épaules des programmeurs et programmeuses un « fardeau moral ». Autrement dit, l'illusion d'objectivité va de pair avec un inégal partage des responsabilités. En regardant de près comment les algorithmes sont pensés et construits, l'auteur interroge la liberté laissée aux programmeurs et mathématiciens de résoudre des questions morales au gré de choix techniques. Il défend que pour fonctionner les algorithmes conjuguent toujours une expertise technique et une autorité morale. Par exemple, lorsque les créateurs et créatrices d'un logiciel d'aide à la décision judiciaire décident d'utiliser un code couleur vert-jaune-rouge pour catégoriser des degrés de probabilité de risque de récidive, ils traduisent des situations complexes en panneaux « STOP – ne pas libérer » faciles à repérer pour les juges mais terriblement réducteurs, pour qui n'a pas connaissance ni du système carcéral ni des logiques de la délinquance. Pour Robinson, une bien trop grande responsabilité est ainsi accordée à des professionnels qui sont pourtant structurellement dans l'incapacité de faire des choix éclairés au plus près des enjeux humains et du fonctionnement des mondes sociaux dans lesquels ils se matérialisent.

Dans une dynamique volontairement optimiste, Robinson invite dès lors à penser les algorithmes comme des outils de régulation nécessairement discutables et critiquables. Il applique au domaine de l'essor technologique le slogan issu des luttes des personnes handicapées « Rien pour nous sans nous » (*Nothing about us without us*) : c'est-à-dire pas d'algorithme à fort enjeu de société sans participation citoyenne. Il fouille l'histoire de la régulation des algorithmes aux États-Unis, et tout particulièrement les débats autour du Kidney Allocation System (KAS), compile 1500 pages de documents fédéraux, de littérature médicale et informatique, et mène une dizaine d'entretiens avec des acteurs de la transplantation et de sa régulation. À partir de cela, il identifie une série de stratégies à travers lesquelles des acteurs extérieurs au monde de la programmation ont pu contribuer, à des degrés divers, au cadrage des valeurs implémentées.

Robinson montre ainsi que les politiques de *transparence* sont les plus courantes, mais aussi les plus insatisfaisantes, car elles se traduisent par des pratiques de divulgation de lignes de codes peu compréhensibles du grand public. Il s'intéresse alors à deux autres ensembles de politiques qui lui semblent complémentaires et prometteuses : audit et prévision. Les politiques d'*audit* des résultats fournis par les algorithmes consistent pour les institutions concernées à produire à échéance régulière des analyses des effets de l'automatisation des décisions selon une série de critères déterminés, relatifs à la performance ou à l'équité. Quant aux politiques de prévision (*forecast*), elles consistent à projeter et à mettre en série les bénéfices attendus du recours à tel algorithme plutôt qu'à tel autre, et ses risques par exemple en matière de discrimination. Dans le cas de l'algorithme de répartition automatisé des reins (KAS), des pratiques d'audit et de prévision ont été mobilisées tour à tour au cours de sa préparation. Elles ont assisté les vifs débats entre représentants des autorités sanitaires, de la médecine et des associations de patients sur la recherche d'un équilibre entre performance médicale, équité du partage des organes et minimisation des discriminations dans l'accès aux organes. L'écriture de l'algorithme a impliqué de trancher des questions morales avec des implications techniques, telles que : est-il acceptable de corrélérer la qualité des organes alloués à la durée de vie estimée des patients et donc à leur âge ? À l'appui d'exercices de prévision et de discussions entre professionnels et membres du public, un ratio 20/80, qui consiste à diriger le top 20 % des reins au top 20% des patients qui ont l'espérance de vie la plus haute, a finalement été jugé équilibré et implémenté dans l'algorithme. Autre exemple, est-il recevable de maximiser la compatibilité entre organe donneur et patient receveur pour limiter les rejets, si cela défavorise l'accès des minorités raciales à la greffe ? Les points accordés à la compatibilité ont fait l'objet d'une série d'ajustements pour équilibrer efficacité médicale et justice sociale. Des rapports d'audits ont ensuite montré une progression de l'accès des catégories de patients « noirs » et « hispanics » par rapport aux « blancs ».

Pour Robinson, chacune de ces stratégies – transparence, audit, prévision et participation du public – offre des prises à la critique et à la transformation des algorithmes. En les cumulant, il défend la possibilité de créer des processus « participatifs, collaboratifs, compréhensibles et mutuellement tolérables » de négociation de la « substance morale » des algorithmes à fort enjeu social.

Une morale sans pratiques ?

Entre enquête journalistique et historique, mais aussi promotion de bonnes manières de gouverner les algorithmes, Robinson évite l'écueil d'une partie de la littérature sur l'éthique des algorithmes qui se perd parfois dans des réflexions détachées de tout lien avec les pratiques. Il offre de nombreux points de contacts avec des archives relatives à la régulation des algorithmes, et avec une littérature protéiforme qui traite tant des processus démocratiques que de la gouvernance technologique. Cadrer les manières de créer des algorithmes suffisamment justes l'intéresse certes pour la finalité des algorithmes qui seront nourris de perspectives critiques issues des personnes visées par les outils, mais aussi pour faire adhérer un public élargi à l'essor de ces technologies. En cela le slogan serait peut-être ici plutôt « Rien pour eux sans eux » : la participation citoyenne est limitée à quelques représentants reconnus par les institutions et apparaît autant comme un outil de justice sociale que de développement technologique qui sert l'intérêt des professionnels du secteur de l'intelligence artificielle dont l'auteur fait dorénavant partie.

La focale du livre sur l'éthique de la conception des algorithmes contient cependant deux angles morts qui méritent d'être soulevés. Le premier est le plus connu puisqu'en mettant l'accent avant tout sur les enjeux de partage du poids moral des algorithmes, l'auteur évite d'étudier les intérêts des professionnels et des entreprises les plus influentes à garder le pouvoir sur ces outils. Le second est plus surprenant. En effet, la réflexion de Robinson sur l'éthique des décisions et les critères de répartition des organes s'arrête aux portes des programmeurs, elle est déconnectée de la manière dont ils traduisent effectivement dans le logiciel les choix moraux longuement établis. Détachée aussi de la manière dont les résultats produits par l'algorithme sont appliqués dans le quotidien de la pratique, si ce n'est contournés. Le livre ne propose à aucun moment des observations des pratiques, ni auprès des équipes techniques quand elles reprennent la main ni auprès des médecins et des agences sanitaires qui se servent du logiciel. Robinson reconnaît que les usages des algorithmes pourraient différer des intentions émises lors de sa conception, comme le montre notamment Danielle Citron dans *The Fight for Privacy* (W. W. Norton & Company, 2022). Mais il décide pourtant de suivre l'hypothèse selon laquelle les règles de répartition des reins démocratiquement tranchées se retrouvent sans altération dans l'outil qu'il étudie et dans ses usages. En se situant en amont, on se retrouve parfois en dehors, ici à la porte d'entrée du bureau des programmeurs et programmeuses.

On peut dès lors se saisir du livre de Robinson comme point de départ pour poursuivre l'enquête. La démonstration de l'épaisseur des compromis éthiques avec lesquels les professionnels des algorithmes doivent composer offre en effet des prises aux chercheurs & chercheuses, mais aussi aux activistes, pour analyser les dimensions morales et techniques des activités algorithmiques. Le recours croissant aux outils d'automatisation de décisions à fort enjeu social met les sciences sociales au défi de tenir de bout en bout la chaîne de construction des algorithmes et de leurs usages, pour mieux en comprendre les effets et les leviers de leur transformation.

Publié dans lavedesidees.fr, le 15 juillet 2024.