

[...]

Contrairement à l'idée mise en avant par Google à ses débuts, le caractère mathématique d'un algorithme n'est en rien un gage d'objective neutralité. Même si les apports de la science des données à la connaissance de la société sont indéniables, ses outils et productions peuvent prêter à la critique.

Dans un livre paru en 1962, *La structure des révolutions scientifiques*, le philosophe Thomas Kuhn présente les faits scientifiques non pas comme une représentation objective de la nature, mais comme le produit du travail des scientifiques, travail socialement conditionné tout comme le reste des activités humaines. On peut illustrer l'idée de conditionnement social des scientifiques par l'histoire de la statistique. À la fin du XIXe siècle, les britanniques Francis Galton et Karl Pearson ont inventé les techniques mathématiques qui forment encore maintenant l'outillage de base de la statistique et de l'économétrie (régression linéaire, écart-type, coefficient de corrélation...). Dans le même temps, ils ont jeté les bases de l'eugénisme. En effet, Galton, cousin de Darwin, voyait dans le traitement des données sur les populations un outil d'action politique au service d'un darwinisme social, dont un des objectifs serait de favoriser les individus selon leur valeur intrinsèque. Dans *Genetics in the Madhouse*, l'historien des sciences Théodore Porter montre comment la génétique est née des travaux de ces premiers statisticiens eugénistes à partir de la grande quantité de données d'hérédité accumulées à cette époque par les hôpitaux psychiatriques.

Que ce soit pour la recherche ou pour les décisions politiques, stratégiques ou opérationnelles, la construction d'infrastructures de données n'est jamais neutre et résulte toujours d'un nombre incalculable de décisions, d'arbitrages, d'actions individuelles et collectives, de renégociations permanentes, qui sont souvent occultées par la « vision » portée par un projet. Dans nos routines quotidiennes, on utilise Google, Facebook, Spotify ou d'autres services numériques comme on ouvre le robinet pour puiser quelques centilitres au réseau d'approvisionnement en eau. Qui sait quelles controverses oubliées ont accompagné en leur temps la mise en place de ces réseaux plus anciens : l'eau potable, l'électricité, les égouts, le chemin de fer ?

Une des tâches de la sociologie consiste à débusquer les nombreux processus sociaux à l'œuvre dans la conception, le déploiement et les usages de dispositifs techniques qui finissent par aller de soi. C'est ce que souligne Susan Leigh Star dans « L'ethnographie des infrastructures ». Un article de Clément Marquet traite des enjeux politiques locaux de l'implantation, depuis les 20 dernières années, en banlieue nord de Paris, de la plus grande concentration en Europe de *data centers*, centres de traitement de données informatiques, assemblages massifs de matériel informatique lourd, de composants

architecturaux austères et de dispositifs de sécurité ostensibles, qui tranchent avec le métaphorique nuage dont ils sont la matérialisation concrète.

D'autres travaux récents en sociologie mettent ainsi en évidence des aspects largement ignorés de la production des données et du fonctionnement des algorithmes. Jérôme Denis et Samuel Goëta montrent à quel point l'*open data*, la publication de données par les organisations, notamment les administrations publiques, nécessite une quantité de travail considérable, à différentes étapes d'un processus complexe (identification, extraction, nettoyage, redéfinition) impliquant des transformations organisationnelles. De même, Antonio Casilli, dans *En attendant les robots*, souligne les multiples formes de travail hautement précaire que cache la nouvelle économie qui met en avant, depuis quelques années, les algorithmes comme principal progrès technique.

Les conditions de travail des chauffeurs Uber commencent aujourd'hui à être bien documentées, ce qui est beaucoup moins visible, ce sont les modalités de mise en place des méthodes d'intelligence artificielle, qui ne reposent pas seulement sur des algorithmes très performants, mais aussi sur un vaste prolétariat de travailleurs et travailleuses à l'intelligence tout à fait humaine, très majoritairement implanté/es dans des pays où le droit du travail est très peu développé. Leur rôle est d'entraîner les algorithmes sur des micro-tâches très répétitives telles que cliquer, dans une suite d'images, sur celles qui contiennent des tomates, ou vérifier la retranscription des caractères des plaques d'immatriculation apparaissant sur une série de vidéos.

D'une manière générale, les visions véhiculées par les algorithmes de la « science des données » ou de l'« intelligence artificielle » sont souvent plus proches de l'utopie scientifique ou de la dystopie façon *Meilleur des mondes*, que de la prise en compte réaliste et éclairée des processus sociaux qui accompagnent le déploiement de ces avancées techniques, processus sociaux qu'étudie la sociologie, avec ou sans algorithmes. Comment des problèmes deviennent-ils prioritaires dans l'agenda de l'innovation ? Comment, pour répondre à ces problèmes, des standards émergent-ils parmi de nombreuses solutions concurrentes ? Dans l'application de ces standards, quelles procédures met-on en place ? Et pour contourner les lourdeurs, les biais ou les angles morts de ces procédures, quels petits ou grands arrangements s'inventent-ils au quotidien ? Autant de questions dont dépendra la société de demain, avec des machines, mais surtout avec des humains.