

**Hannah Fry, Hello world: how to be human in the age of the machine, London, Penguin Random House, 2019.**

Esin Ekizoglu

► **To cite this version:**

Esin Ekizoglu. Hannah Fry, Hello world: how to be human in the age of the machine, London, Penguin Random House, 2019.. 2020, pp.149-154. hal-02900742

**HAL Id: hal-02900742**

**<https://hal-univ-paris.archives-ouvertes.fr/hal-02900742>**

Submitted on 16 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Esin EKIZOGLU\***

**Hannah FRY, *Hello world: how to be human in the age of the machine*, London, Penguin Random House, 2019.**

Comment être humain à l'ère des machines ? L'ouvrage d'Hannah Fry<sup>1</sup> dresse le tableau de la vie quotidienne des hommes et des femmes du XXI<sup>e</sup> siècle. Autour des premiers mots de l'ouvrage, « Hello world ! », marquant le début des interactions entre humain et machine, H. Fry organise son propos en sept chapitres afin de mettre en exergue les enjeux que constituent les algorithmes dans une société où les nouvelles technologies d'information et de communication sont si prégnantes. En effet, en tant que simple individu vivant dans cette société, est-ce que nous connaissons vraiment les propriétés de ces algorithmes dont la présence change notre vie ? Quels sont les compromis que nous sommes menés à faire avec ces entités mathématiques complexes qui agissent sur nos vies privées, protections, équités et sécurités ?

Le premier chapitre, « Power », s'intéresse essentiellement à la force des algorithmes sur nos vies. L'auteure tente d'abord de définir ce qu'est un algorithme en s'appuyant essentiellement sur l'instant où la frontière entre le contrôleur et le contrôlé devient imperceptible. En effet, selon elle, un algorithme marque « le début d'un partenariat, un voyage partagé de possibilités où l'un ne peut exister sans l'autre » (p. IX). Il a le pouvoir caché de changer lentement et subtilement les règles liées à la vie de son contrôleur, c'est-à-dire l'être humain. Aucun algorithme, aucun objet technique n'est jamais bon ou mauvais en soi. Le pouvoir des objets techniques qui fonctionnent grâce à des algorithmes dépend de la façon dont ces objets sont utilisés. L'utilisation du GPS<sup>2</sup> est un bon exemple puisque « inventé pour lancer des missiles nucléaires, le GPS permet désormais de livrer des pizzas » (p. 3).

Largement alimenté par les notes des conversations privées,

---

\* Architecte diplômée d'état (DE), urbaniste, doctorante en architecture sous la direction d'Elizabeth Mortamais au sein du Laboratoire EVCAU à l'École d'Architecture Paris Val-de-Seine.

<sup>1</sup> Hannah Fry est professeure agrégée de mathématiques à l'University College London.

<sup>2</sup> *Global Positioning System*

des reportages et des références scientifiques, le deuxième chapitre, « Data » éclaire les lecteurs sur les champs d'action des données issues des algorithmes. Quelle est la capacité d'action des algorithmes ? Que devons-nous savoir lorsque nous utilisons les ressources numériques créées par les algorithmes ? H. Fry, en s'appuyant sur trois études en psychologie<sup>3</sup> concernant les comportements humains face aux algorithmes, émet l'idée qu'un algorithme pourrait déduire la personnalité de ses utilisateurs à partir de leurs simples interactions avec les technologies numériques et connectées (p. 46). Une fois les personnalités des utilisateurs connues, celles-ci deviennent l'objet de manipulation émotionnelle. Penser être immunisé contre l'influence des algorithmes est une erreur étant donné que, selon l'auteure, personne n'est à l'abri de la collecte de données privées et de la manipulation de celles-ci (p. 50).

Il est très fortement conseillé par H. Fry de faire attention aux données recueillies et formulées par les algorithmes. Les données agissent fortement et indéniablement sur le quotidien des êtres humains. Nous pouvons tout de même nous interroger sur la fiabilité de celles-ci. Le troisième chapitre, « Justice », prouve aux lecteurs que les algorithmes, malgré leurs apparences objectives, ne sont pas infaillibles. En étudiant les algorithmes utilisés dans le système judiciaire américain, l'auteure montre de quelle manière les erreurs de calcul peuvent perpétuer les inégalités.

Il existe deux types d'erreurs qu'un algorithme peut commettre selon Richard Berk<sup>4</sup>. Ce sont les faux positifs et les faux négatifs. Afin de mieux comprendre ces deux erreurs, H. Fry reprend l'exemple de Berk fondé sur le film culte *Star Wars* de George Lucas. Dès le départ, l'algorithme calcule ses données en considérant qu'il y a deux sortes de résultats : « des bons et des méchants » (p. 72). L'algorithme demande qui ce sont les Darks Vadors (les méchants) et qui sont les Lukes Skywalker (les bons). Laisser un Dark Vador se libérer est une sorte

---

<sup>3</sup> Michal KOSINSKI, David STILLWELL et Thore GRAEPEL, « Private traits and attributes are predictable from digital records of human behavior ». *Proceedings of the National Academy of Sciences*, vol. 110, n° 15, 2013, p. 5802 ; Wu YOUYOU, Michal KOSINSKI et David STILLWELL, « Computer-based personality judgments are more accurate than those made by humans », *Proceedings of the National Academy of Sciences*, vol. 112, n° 4, 2015, p. 1036 ; Sandra MATZ, Michal KOSINSKI, Gideon NAVE, David STILLWELL, « Psychological targeting as an effective approach to digital mass persuasion », *Proceedings of the National Academy of Sciences*, vol. 114, n° 48, 2017, p. 12714–12719 .

<sup>4</sup> Il est professeur de criminologie et de statistique à l'Université de Pennsylvanie et pionnier dans le domaine de la prédiction de la récidive.

d'erreur, connue sous le nom de faux négatif. Cela se produit chaque fois que l'algorithme ne parvient pas à identifier le risque qu'une personne pose. L'incarcération de Luke Skywalker, en revanche, est un faux positif. C'est à ce moment que l'algorithme identifie incorrectement quelqu'un comme une personne à haut risque (p. 72). Comment persuader les individus, les juges de faire preuve de recul lorsqu'il est question d'utiliser ces algorithmes (p. 76) ? Les recherches récentes de ProPublica<sup>5</sup> sur l'algorithme COMPAS<sup>6</sup>, utilisé par les juges américains, mettent en évidence à quel point il est facile de commettre des erreurs<sup>7</sup>.

Malgré les probabilités d'erreurs, les algorithmes sont toujours utilisés. Le quatrième chapitre, « Medicine », aborde l'utilité des algorithmes dans le cadre médical. Le principe de « ressemblance » mis en avant par les mathématiques permet en effet de sauver des vies. Un des exemples fructueux de cette utilisation peut être vu dans l'algorithme élaboré par l'équipe Google Brain. Celui-ci, fondé sur le principe de travail sur les ressemblances, parvient à dépister la plus grande cause mondiale de rétinopathie diabétique, une maladie pouvant être évitée par un dépistage précoce (p. 113). Par ailleurs, dans ce chapitre, H. Fry met en évidence l'équilibre induit, et ce dans toutes les facettes de la vie, par l'introduction d'un algorithme : équilibre entre intimité et bien public, individu et population, mais également entre différents défis et priorités (p. 132). Cependant, et malgré cet équilibre, l'auteure montre que l'utilisation des algorithmes en vue de l'amélioration de nos systèmes de santé est une voie sinueuse, d'autant plus quand la concurrence économique mène à la surestimation des avantages et la minimisation des risques. H. Fry en appelle ainsi à davantage de réflexion quant aux objectifs économiques des acteurs manipulant les algorithmes.

L'utilisation du principe des ressemblances par les algorithmes ouvre également différents horizons aux êtres humains. Elle permet de débiter une réflexion sur la liberté de l'être humain des obligations qui

---

<sup>5</sup> ProPublica est un organisme à but non lucratif basé à New York et gestionnaire du site Internet ProPublica.org. L'organisme se décrit comme créateur de nouvelles indépendantes spécialisées en journalisme d'enquête d'intérêt public. L'étude de ProPublica sur COMPAS est disponible sur : <https://www.propublica.org/>.

<sup>6</sup> COMPAS, acronyme de *Correctional Offender Management Profiling for Alternative Sanctions*, est un outil de gestion de cas et d'aide à la décision.

<sup>7</sup> Aaron M. BORNSTEIN, « Are Algorithms Building the New Infrastructure of Racism? », *Nautilus*, n° 55, 2017, disponible sur : <http://nautil.us/issue/55/trust/are-algorithms-building-the-new-infrastructure-of-racism>.

Keith KIRKPATRICK, « It's not the algorithm, it's the data », *Communications of the ACM*, vol. 60, n° 2, 2017, p. 21-23.

demandent de la motivation et de l'attention. Le cinquième chapitre, « Cars », se concentre sur les possibilités de libérer l'être humain d'une partie de ses obligations grâce aux algorithmes, notamment grâce à ceux qui sont formés d'après le théorème de Bayes. Le chapitre prend les voitures autonomes comme exemple majeur. L'auteure met l'accent sur deux positions contradictoires concernant ce théorème et ses finalités. La première position permet de faire comprendre que le théorème de Bayes<sup>8</sup> décrivant la probabilité d'un événement est basé sur la connaissance préalable des conditions qui pourraient être liées à cet événement et qu'il va à l'encontre de la conviction profondément ancrée que la science moderne nécessite objectivité et précision (p. 143). Ce théorème permet de tirer des conclusions sensées à partir d'observations sommaires, de données en désordre, incomplètes et approximatives<sup>9</sup>. À partir de cela, de nombreuses entreprises entreprennent la construction de voitures autonomes, ne nécessitant donc aucun conducteur, en affirmant pouvoir anticiper tous les risques afférents à leur mise en circulation. Paul Newman<sup>10</sup> le confirme en disant : « C'est beaucoup, plusieurs millions de lignes de code, mais je pourrais encadrer le tout [...] » (p. 149).

La seconde position, émanant de chercheurs, voit dans cette utilisation industrielle du théorème une bétise. Pour Jack Stilgoe<sup>11</sup> la nature malicieuse de l'être humain rend l'utilisation des algorithmes dans la conception des véhicules sans conducteur dangereuse. Les individus, d'après lui, ne constituent pas seulement les éléments passifs du paysage, mais restent des « agents actifs » imprévisibles (p. 151). La rhétorique de l'autonomie des transports consiste à ne pas changer le monde. En revanche, créer un robot et lui permettre d'être aussi bon qu'un humain dans un environnement en changement perpétuel n'a pas de sens (p. 152).

Pour faire la synthèse de ces deux positions, l'auteure s'appuie sur les propos de Lisanne Bainbridge selon qui le maintien d'une attention visuelle et efficace sur une source d'information est impossible au-delà d'une demi-heure, même pour un individu très motivé (p. 159).

---

<sup>8</sup> Les informations complémentaires sont disponibles en lien : [https://arbital.com/p/bayes\\_rule/](https://arbital.com/p/bayes_rule/).

<sup>9</sup> Sharon B. McGRAYNE, *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy*, Londres, Yale University Press, 2012.

<sup>10</sup> Il est professeur de robotique à l'Université d'Oxford et fondateur d'Oxbotica, une entreprise qui construit des voitures sans conducteur et les teste dans les rues de Grande-Bretagne.

<sup>11</sup> University College de Londres.

En reprenant la référence de Bainbridge<sup>12</sup>, l'auteure indique que les humains sont vraiment « bons » pour comprendre les subtilités, pour analyser le contexte, et appliquer l'expérience, mais qu'ils ont une mauvaise attention pour la précision, la cohérence et la conscience de leur environnement. Ils ont, en bref, précisément l'ensemble des compétences dont sont dénués les algorithmes (p.164). C'est pourquoi H. Fry ouvre un nouveau champ de réflexion : pourquoi ne pas suivre l'exemple du logiciel de recherche médicale où l'expertise humaine est complétée par les machines ? Laisser les compétences de la machine compléter celles de l'homme ne pourrait-il pas faire progresser les capacités des deux (p. 165) ?

Parmi ces attentes se trouvent les prédictions, au cœur du sixième chapitre, « Crime ». S'appuyant toujours sur le principe de ressemblance des algorithmes, H. Fry explique à ses lecteurs comment les algorithmes permettent de prédire les crimes. La prédiction saine dans les cas de crimes n'est possible qu'en considérant les crimes passés comme des cas non isolés (p. 175). Des mises à jour permanentes sont indispensables afin de faire de meilleures prédictions.

Le deuxième grand champ d'attente des humains envers les algorithmes se trouve dans le monde de la création. Le septième chapitre, « Art », suit la logique claire de l'auteure et complète les autres chapitres. S'inspirant des prédictions des crimes, H. Fry invite ses lecteurs à un raisonnement similaire pour la production des arts. Elle essaie de les faire réfléchir sur le principe de ressemblance des algorithmes dans le cadre de nouvelles créations artistiques. Ayant ce but, elle donne deux arguments référencés, mais opposés.

Le premier argument repose sur l'habitude d'imitation humaine dont la référence est Mark Twain. Plus clairement H. Fry explique qu'à chaque fois que les humains n'ont pas suffisamment d'informations pour prendre des décisions par eux-mêmes, ils copient le comportement des individus qui les entourent. Les arts se développent d'après cette logique d'imitation. La grande partie de la créativité humaine, comme les produits des algorithmes de « composition », n'est qu'une combinaison inédite d'idées préexistantes (p. 228). Ainsi les machines peuvent certainement contribuer aux arts.

Néanmoins l'auteure préfère la vision de Léon Tolstoï, d'après lequel la créativité humaine est inatteignable par les machines. Les vrais arts ont à leurs origines la connexion « humaine » liée à l'émotion communicationnelle. Les arts ne sont pas un artisanat, c'est

---

<sup>12</sup> Lisanne BAINBRIDGE, « Ironies of Automation », *Automatica*, vol. 19, n° 6, 1983, p. 775-779.

la transmission du ressenti que l'artiste a vécu<sup>13</sup>.

Que ce soit dans le système judiciaire, les soins de santé, les services de police ou même les achats, se posent des problèmes de confidentialités, de partis pris, d'erreur, de responsabilité et de transparence. En vertu de certains algorithmes existants, les êtres humains sont confrontés à des problèmes d'équité touchant au modèle de société auquel ils aspirent. C'est peut-être précisément le fait de penser les algorithmes comme détenteur d'une sorte d'autorité qui constitue une erreur (p. 235).

L'auteure propose ainsi d'imaginer les impacts que pourrait avoir l'acceptation de l'imperfection par les humains. Les algorithmes pourraient alors commettre des erreurs. Cela ne devrait en aucun cas distraire les humains de la lutte pour les rendre plus précis et moins biaisés partout où ils le peuvent, mais reconnaître que les algorithmes ne sont pas infaillibles, tout comme les humains, pourrait permettre d'amoindrir leur autorité (p. 237).

H. Fry conseille à ses lecteurs d'accepter que les algorithmes soient conçus pour soutenir les humains dans leurs décisions plutôt que de les instruire au sens éducatif du terme (p. 238). Elle ajoute qu'un avenir où l'effet global des algorithmes serait une force positive pour la société n'est possible que si les humains cessent de considérer les machines comme des maîtres objectifs. C'est-à-dire en remettant en cause leurs décisions, en examinant leurs motivations, en exigeant de savoir qui en bénéficiera et en les tenant responsables de leurs erreurs. Enfin, à l'ère des algorithmes, les humains n'ont jamais été aussi importants.

Avec son ouvrage clair et succinct, Hannah Fry, réussit à écrire un récit accessible à un très large public. Tout en posant des questions auxquelles se confrontent les chercheurs, il constitue un très bon exemple de vulgarisation scientifique. Par ailleurs, H. Fry nous prouve par l'intégrité de son ouvrage que les sciences dites « dures » doivent s'intéresser, sans aucun doute, de très près aux sciences humaines, pour ne pas être déconnectées de la réalité et du quotidien de la vie.

---

<sup>13</sup> Léon TOLSTOÏ, *What is Art?*, Londres, Penguin, 1995 [1897].