

Dominique Cardon, *A quoi rêvent les algorithmes ?* (2015)

Fiche de lecture



Luigi Russolo, *Dynamisme d'une automobile* (1913)

Né en Italie au début du XXe siècle, le futurisme se distingue par sa fascination pour la modernité et une recherche de l'expression picturale de la vitesse. Le *Dynamisme d'une automobile* de Russolo témoigne de l'attrait des futuristes pour les machines et leur représentation en mouvement.

**

Dominique Cardon est un sociologue français qui s'intéresse aux différents usages d'internet : la gouvernance de Wikipedia ou internet comme espace d'expression publique qu'il développe dans *La démocratie internet. Promesses et limites*. Dans *A quoi rêve les algorithmes ?*, il se livre à une analyse sociologique des algorithmes et du *big data*.

Comprendre la révolution des calculs

L'objet de ce livre est de montrer comment les techniques du *big data* font progressivement advenir une « *société des calculs* ». Dominique Cardon cherche à dépasser l'analyse technique des algorithmes pour révéler le projet politique qu'ils véhiculent.

Deux dynamiques soutiennent le développement de la « *société des calculs* » : **la numérisation progressive de la société**, qui prend la forme du *big data*, dont le défi reste encore aujourd'hui le traitement, et **le développement des algorithmes** dont l'objet est justement d'ordonner et de donner du sens à une quantité infinie et brute de données.

Le classement de l'information sur le web est aujourd'hui au cœur des enjeux sociaux, éthiques et politiques du calcul algorithmique. Le réel est aujourd'hui façonné et orienté par le calcul algorithmique en partie via ce classement de l'information. Nos choix se retrouvent orientés par les catégories et les conventions que les algorithmes produisent. Progressivement, ils imposent leur propre hiérarchie des valeurs, façonnant ainsi « *les cadres cognitifs et culturels de nos sociétés* ».

1. Quatre familles de calcul numérique

On peut distinguer les algorithmes en quatre familles, selon qu'on les trouve à côté, au dessus, dans et en dessous des données numériques.

Les mesures d'audience, qui servent à mesurer la popularité des sites en comptabilisant les clics des internautes **se placent à coté des données**. Ces mesures d'audience, dont l'utilité est avant tout publicitaire, permettent de recueillir des informations sur le profil sociologique des internautes. Cette technique traditionnelle va progressivement se voir remplacer **par une mesure qui s'intéresse moins aux profils des individus qu'à leur comportement, en enregistrant leurs traces sur le web, en dessous des données numériques**.

On retrouve **au dessus des données, les outils de classement de l'information**, comme l'algorithme *Page Rank*, servant à **mesurer l'autorité des sites**, au moyen des liens hypertextes échangés. Il s'agit d'**apprécier la force sociale** d'une page en partant du principe que plus celle-ci est citée par d'autres pages web, plus son autorité scientifique se renforce. Inspiré du classement des revues scientifiques, ce système qui se veut avant tout **méritocratique** est cependant détourné ; en premier lieu par le marché du *Search Engine Optimization* qui permet aux acteurs d'améliorer leur classement sur Google. De plus, **la mesure d'autorité, en centralisant les acteurs, se transforme rapidement en mesure de popularité, figeant les positions et excluant des acteurs moins connus**. Définissant la position des sites sur les pages de recherche, cette mesure est devenue une source de revenus pour Google, **certaines entreprises payant pour se retrouver en tête de liste**.

Les likes ou autres moyens d'expression sur les réseaux sociaux, positionnés **à l'intérieur du web**, vont permettre de **mesurer la réputation des internautes**. Le ratio entre le nombre de personnes que l'on connaît et le nombre de personnes dont on est connu permet d'évaluer l'influence sur les réseaux sociaux.

Au dessous du web, les algorithmes fonctionnant sur la technique du *machine learning*, enregistrent les traces laissées par les internautes lors de leur navigation. Il s'agit de s'intéresser aux **conduites réelles des internautes** afin d'orienter le contenu publicitaire qui leur est proposé. La technique des **cookies tiers** permet de mutualiser les cookies entre différents sites afin que la trace de l'internaute ne se perde pas lorsqu'il change de page : celui-ci se voit proposer de **la publicité qui reprend les contenus des sites qu'il a précédemment consultés**.

2. La révolution dans les calculs

En s'intéressant aux individualités et en rejetant toute construction holiste, ces nouveaux calculs répondent au mouvement de singularisation des sociétés. Tout comme les instruments des politiques néolibérales, **il s'agit moins de représenter le réel que d'agir sur lui**. Les big data tendent ainsi à faire du monde un laboratoire dans lequel **le réel est quantifiable dans toute sa singularité**.

Les techniques numériques s'intéressent **d'avantage aux évènements isolés qu'aux catégories structurantes forcément moins précises**. Le système publicitaire d'enchères en temps réel (*real-time bidding*) qui met en vente le profil d'un utilisateur lambda en fonction de ses recherches précédentes répond à cette logique. La géolocalisation fait désormais office de seul outil de catégorisation.

Le modèle scientifique traditionnel se retrouve lui aussi obsolète. En effet, **les corrélations n'ont désormais plus besoin de causes**. Tant qu'on enregistre les activités individuelles avec une précision parfaite, la recherche des causes devient superflue. On travaille désormais **les corrélations sans chercher de modèle explicatif a priori**. L'algorithme se sert du comportement d'un premier groupe pour prédire le comportement d'un autre groupe aux caractéristiques similaires. Les cadres de la société sont désormais déterminés par le bas, en partant des comportements individuels pour en inférer ensuite les attributs qui les rendent statistiquement probables.

3. Les signaux et les traces

Si le développement du *big data* a permis l'accès à un nombre quasi infini de données, leur traitement fait aujourd'hui face à plusieurs limites. Tout d'abord, les données brutes, non biaisées, n'existent pas. Sortis du contexte de leur production, ces chiffres n'ont souvent aucun sens. Les données prennent une signification en fonction des intérêts et des buts de ceux qui les utilisent. Les mêmes données de l'assurance maladie prennent un sens différent selon qu'elles sont utilisées par un assureur ou une association de malades.

Les machines de l'IA et du *deep learning* sont aujourd'hui moins intelligentes que statistiques dans le sens où **leur développement passe davantage par l'assimilation d'énormes masses de données que par l'acquisition d'une capacité d'adaptation à des situations différentes**. Par exemple, le développement de la traduction automatique est d'abord passé par un travail de conceptualisation cherchant à reproduire le raisonnement formel d'un cerveau humain au sein d'un traducteur automatique. Face aux résultats peu convaincants, une autre direction est prise : **à la place d'un raisonnement abstrait, la machine apprend mot par mot**, groupe de mots par groupe de mots, la correspondance entre deux textes déjà traduits par d'autres (par exemple les traducteurs des institutions européennes). La machine ne traduit pas mais estime la traduction en la comparant aux autres traductions qu'elle a en mémoire. **La machine ne « comprend » rien mais « apprend » beaucoup.** *« L'innovation des big data réside dans ce passage des règles abstraites vers la statistique des contextes. L'enjeu n'est plus d'apprendre aux machines une grande théorie appliquée à peu de données, mais de multiplier les petites théories en demandant à beaucoup de données contextuelles de sélectionner la ou les meilleures d'entre elles ».*

« Un algorithme fonctionne véritablement lorsqu'il parvient à épouser si étroitement le milieu dans lequel il intervient que les comportements des acteurs se règlent sur ses verdicts et que les principes qu'il met en œuvre nourrissent leur représentations ». Comme c'est le cas de *Page Rank* ou de la notation d'hôtels sur *Tripadvisor*, la conduite de l'utilisateur devient **régulière et prévisible**. Dénotant **un processus de socialisation**, elle révèle **un habitus sur lequel se construit un modèle**. *« Vue depuis les algorithmes, la société ne repose plus sur de grands systèmes de détermination, mais elle est une sorte de microphysique des comportements et des interactions que des capteurs placés à bas niveau savent décoder ».* La prédiction des algorithmes se fonde sur une analyse précise des actions passées d'un individu et sur le pari que celui-ci ne déviera pas de sa trajectoire.

Les individus tendent à produire eux-mêmes leurs propres données à travers des **pratiques de mesure de soi** (*quantified self*). Ainsi, cette quantification, cette expression chiffrée, ne serait qu'**une nouvelle forme de liberté d'expression**. La vie privée n'a ainsi pas disparu mais est devenue plus individuelle, « privatisée ». On passe d'une conception de la vie privée comme protection à **la vie privée vue comme une liberté**, les individus souhaitant contrôler ce qu'ils publicisent. Cet arbitrage qui se veut rationnel est néanmoins **biaisé par le déséquilibre de la relation entre individus et plateformes** au positionnement monopolistique.

Face à la critique récurrente du biais des algorithmes dans la hiérarchisation des informations, nait la revendication, récemment portée par le Conseil national du numérique et le Conseil d'État¹, d'**une loyauté des plateformes à l'égard de leurs utilisateurs**. Ces derniers doivent connaître les choix « éditoriaux » qui président à la hiérarchisation ou la sélection d'une information sur une autre dans le moteur de recherche Google ou le fil d'actualité Facebook.

« Apprenant » d'une multitude de situations passées, les algorithmes reproduisent les situations statistiques les plus fréquentes, sur le comportement des internautes, quitte à donner des résultats faux ou reproduire des situations de discrimination. C'est ainsi que le processus algorithmique est considéré comme « conservateur » en ce qu'il ne fait que reproduire et accentuer un ordre social. *« Sans doute est-ce à travers cette manière d'entériner l'ordre social en reconduisant les individus vers leurs comportements passés que le calcul algorithmique exerce sa domination. Il prétend leur donner les moyens de se gouverner eux-mêmes ; mais, réduits à leur seule conduite, les individus sont assignés à la reproduction automatique de la société et d'eux-mêmes. Le probable préempte le possible. »*

4. La société des calculs

En analysant la société « par le bas », en se fondant sur les comportements individuels, les algorithmes et leurs promoteurs **rejetent toute entreprise de création d'un commun**. Voulant libérer les individus des structures collectives qui gommeraient leurs particularismes, les algorithmes continuent *« de produire des effets de centralité d'autant plus fort qu'ils se sont largement émancipés des cadres nationaux pour devenir globaux »*. Au contraire d'une exacerbation de la liberté, l'algorithme tend à **enfermer les individus dans leurs déterminismes en proposant une méritocratie poussée à son paroxysme**. Ceux occupant une position centrale sont favorisés et ceux à la marge sont oubliés. Les gagnant du mérite numérique se voient conférer des avantages cumulés considérables quitte à faire sécession.

Ce web de la réputation et du mérite tend à valoriser le réseau relationnel comme un puissant vecteur de socialisation, à travers **une multitude de liens faibles**. **La multiplication des identités** entraîne également **la multiplication des actions collectives** autour d'une cause ou d'une pétition. La visibilité devient une nouvelle hiérarchisation du social. Pouvant paraître spontanée, cette nouvelle socialisation est également **une stratégie d'acteurs qui réussissent à détourner à leurs fins les mesures comportementales** comme le montre le brouillage entre contenu social et contenu publicitaire sur les réseaux.

¹ CNNum, « Ambition numérique », rapport remis au Premier ministre, juin 2015 ; et Conseil d'État, « Rapport sur la neutralité des plateformes », juin 2014.

Les apports de l'ouvrage de Dominique Cardon

Le travail de Dominique Cardon poursuit la réflexion engagée par Michel Foucault et ses disciples autour de l'émergence combinée de l'individualisme et des dispositifs de contrôle. Les algorithmes ont tous les attributs de ces **dispositifs de contrôle éloignés mais attentifs, quasiment invisibles**, que décrivait Foucault. Ils s'inscrivent ainsi pleinement dans les concepts de **gouvernementalité**, de gouvernement cybernétique², ou de gouvernement des conduites individuelles³.

Si les algorithmes procèdent « *d'un désir d'autonomie et de liberté, ils contribuent aussi à assujettir l'internaute à cette route calculée, automatique qui s'adapte à nos désirs en se réglant secrètement sur le trafic des autres* ». C'est ainsi un objectif ultime d'orientation sans contrainte que poursuivent les algorithmes. Les instruments du *big data* sont des **nudges**, ces dispositifs au fondement d'un nombre croissant de politiques publiques et qui influencent les comportements des individus à leur insu. Fondant une norme idéale auquel chacun doit tendre, **les algorithmes comme les nudges exacerbent paradoxalement l'individualité et le conformisme des comportements**.

L'usage qui est fait des algorithmes n'est pas indépendant du contexte culturel de leur émergence. En cela, ils doivent être étudiés au prisme d'une anthropologie de la technique comme le proposent André Leroy-Gourlan ou Gilbert Simondon. S'inspirant de ces derniers, le philosophe Bernard Stiegler propose **une lecture sociale et culturelle de la technique**⁴. Elle ne fonctionne comme technique que parce que la société se rend intelligible à son égard, se laisse modeler par son action. **La société sélectionne ainsi les techniques**, et en privilégie certaines sur d'autres. Les techniques ne sont pas isolables des modes de vie et de pensée qui structurent une société : elles sont **le résultat direct de l'action des structures sociales dominantes**.

Les logiques de personnalisation à l'œuvre à travers les structures du *big data* proposent une nouvelle forme de lecture sociale en mettant **la focale sur les comportements individuels**. Elles participent ainsi du **mouvement bien plus large d'autonomisation des individus et d'individualisation de la société**.

² Antoine Picon, *La ville des réseaux, un imaginaire politique*, Manucius Eds, 2014

³ Sophie Dubuisson-Quelet, *Gouverner les conduites*, Presses de Sciences Po, 2016

⁴ Bernard Stiegler, *Dans la disruption. Comment ne pas devenir fou*, Les liens qui libèrent, 2016