



HAL
open science

Il faut se défier de la “ pensée magique ”

David Monniaux

► **To cite this version:**

David Monniaux. Il faut se défier de la “ pensée magique ” : 5 questions à David Monniaux. Gilles Rouet. Algorithmes et décisions publiques, CNRS Éditions, 2019, 978-2-271-12466-1. hal-02043272

HAL Id: hal-02043272

<https://hal.science/hal-02043272>

Submitted on 10 Apr 2019

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.

Il faut se défier de la « pensée magique »

5 questions à David Monniaux (directeur de recherche au CNRS, laboratoire VERIMAG, Grenoble)

1) Les algorithmes aident les responsables publics à prendre des décisions concernant les citoyens. N'y a-t-il pas un risque de déresponsabilisation de ces décideurs publics avec la généralisation de procédures peu maîtrisées ?

Dans l'idéal, les pouvoirs publics spécifient clairement ce que l'algorithme doit produire en sortie en fonction des données qu'on lui présente, et les spécialistes proposent un algorithme répondant à cette spécification, ou, dans un dialogue, pointent d'éventuelles ambiguïtés ou impossibilités dans celle-ci, ou encore des conséquences logiques désagréables que le décideur n'aurait pas vues. Le pouvoir politique se concentre alors sur le « quoi » et non le « comment ».

En pratique, le pouvoir politique peut trouver commode de se défausser sur l'algorithme et ses concepteurs, de la même façon qu'il peut trouver commode d'imputer une décision déplaisante à l'« administration ». Gouverner, c'est souvent s'attribuer les succès et imputer les échecs et dysfonctionnements aux subalternes.

Le danger est peut-être encore plus grand avec les algorithmes basés sur l'apprentissage, c'est-à-dire ceux qui sont obtenus en faisant étudier un grand nombre de cas par un modèle numérique (par exemple, un réseau de neurones) dans l'espoir que celui-ci généralise les cas étudiés vers le cas général, dont il n'y a donc pas de spécification à proprement parler. En effet, il est alors tentant de dire que les résultats du modèle sont « scientifiquement établis », alors que la constitution et l'utilisation d'un modèle numérique sont choses délicates, pour lesquelles y a des choix à faire, éventuellement contestables et dont il faudrait pouvoir discuter. Il faut se défier de la « pensée magique » !

La transparence des algorithmes et des logiciels (ce qui inclut, pour ceux basés sur l'apprentissage, les paramètres issus de l'apprentissage) est une réponse partielle à ces questions. En effet, une condition nécessaire d'un débat politique est que l'on sache ce dont on parle. Or, souvent, on parle de logiciels ou de services « boîte noire », au fonctionnement connu uniquement de leurs fournisseurs.

2) Le jugement moral, l'intuition, la créativité, l'affect, entre autres, fondent une grande partie des relations humaines et de la communication. La généralisation des procédures algorithmiques semble installer les relations de service dans un cadre « déshumanisé ». La conséquence de cette évolution n'est-elle pas une perte de sens à la fois pour le décideur et pour le citoyen ?

Lorsque nous avons affaire à une administration, nous ne désirons pas forcément que l'employé fasse appel à son jugement moral, son affect ou sa créativité, mais plutôt qu'il applique les lois et règlements. Le fonctionnaire qui emploie son jugement moral, cela peut vouloir dire, comme on l'a vu dans un célèbre exemple étatsunien, le refus de l'inscription d'un mariage homosexuel. Moins grave et plus proche de nous, de nombreux employés administratifs attribuent d'autorité aux femmes mariées le nom de leur conjoint même quand elles demandent explicitement à ce que ne soit pas le cas, au mépris de la loi mais selon leur idée personnelle que c'est l'ordre naturel des choses. La supériorité de la relation humaine n'est donc pas claire.

À mon avis, le principal danger est l'application de procédures excessivement rigides, dont il est très difficile de sortir si l'on n'est pas dans le cas courant prévu par les concepteurs du logiciel. C'est déjà le cas

avec des procédures administratives gérées par des humains, et ce sera pire si l'on ne prévoit pas systématiquement la possibilité d'avoir recours à un opérateur humain « superviseur » pour régler les cas compliqués. Il faut également se rappeler que tous les citoyens n'ont pas forcément l'habitude des procédures automatisées, n'ont pas forcément accès à Internet dans de bonnes conditions, et que la dématérialisation imposée des procédures est donc fortement discriminatoire envers certaines populations.

Outre la possibilité d'accéder à un opérateur humain, certaines difficultés peuvent être évitées par de bons choix en matière d'interface utilisateur, travail souvent négligé. Par exemple, des lycéens ont eu de graves déboires avec ParcoursSup en raison de la mauvaise conception de l'interface de l'application de connexion.

3) Les algorithmes amélioreraient nos quotidiens, mais en utilisant des données liées à nos actions passées. Comment dans ce contexte faire reconnaître nos désirs et en particulier nos éventuels souhaits de rupture alors que le futur que les algorithmes nous proposent est enchaîné à notre passé ?

Vous parlez ici des algorithmes de suggestion basés sur l'apprentissage d'un profil d'utilisateur, du style « si ce client a par le passé commandé des albums de *heavy metal*, on va lui proposer d'autres albums de *heavy metal* ». Il me semble qu'il ne s'agit pas d'une nouveauté fondamentale. L'amateur de musique « classique » s'abonnait à un magazine de musique « classique » qui lui proposait des enregistrements de musique « classique ». Suivant ses options politiques, l'électeur s'abonnait au *Figaro* ou à *l'Humanité*, qui lui proposait des informations et des analyses conformes à sa vision du monde. Une hypothèse qui a beaucoup circulé depuis dix ans est que ce phénomène de « bulle informationnelle » (ne voir que des choses, des points de vue, des gens similaires à ceux que l'on a appréciés) est attisé par les algorithmes de suggestions. En même temps, Internet donne accès à une plus grande variété de points de vue (par exemple, presse internationale) que ne le permettaient les moyens classiques.

La « justice prédictive » et les assurances posent des problèmes similaires, mais plus graves. On se base alors sur l'idée que le comportement passé d'un individu ou son adéquation à un « profil » permet de prédire son comportement futur. Ce raisonnement date de bien avant l'informatique : on considère qu'un homme violent par le passé risque de l'être à l'avenir, etc. Comment articuler ceci avec l'idée que chaque individu peut s'amender, et qu'on n'est pas enfermé *ad vitam* dans ses comportements passés ? Algorithmiquement, on peut bien sûr former des modèles attribuant plus de poids aux actions récentes que passées. Il s'agit là de problèmes relevant plus de la psychologie que de l'informatique, et pour lesquels je ne me sens donc guère compétent.

4) La généralisation des systèmes algorithmiques et du Big Data semble aboutir à installer une primauté de l'information sur la communication. Ce nouveau paradigme fonde en partie un modèle économique particulier dans lequel les producteurs des informations participent de manière souvent inconsciente à la création de valeur. Cette valeur profite-t-elle finalement aux citoyens ?

On parle en effet beaucoup de la production gratuite d'information et de valeur ajoutée par les utilisateurs de services. Par exemple, le simple fait que de très nombreux conducteurs et passagers de véhicules aient leur téléphone portable allumé sur un segment de route permet de détecter qu'il y a un embouteillage ; une hausse des recherches sur le nom d'une maladie pourrait permettre de détecter une épidémie avant les « remontées de terrain » des réseaux sanitaires. Il s'agit là d'actions inconscientes.

On peut aussi demander aux utilisateurs de fournir un petit travail. Par exemple, afin d'éviter que des « robots » n'accèdent à certains services, on demande de justifier de leur qualité d'humain en accomplissant une petite tâche « Captcha » réputée difficile pour un logiciel, notamment reconnaître un objet sur une photographie, ou encore reconnaître du texte imprimé passant mal la reconnaissance automatisée de caractères. Ce faisant, les utilisateurs entraînent gratuitement des algorithmes par apprentissage, qui pourront par la suite traiter automatiquement plus d'images et de textes, au profit des détenteurs des systèmes (notamment des grands opérateurs tels que Google).

Il ne faut cependant pas oublier qu'à côté de cette captation par des services commerciaux, il existe des services non commerciaux dont les utilisateurs sont invités à contribuer bénévolement. Le plus connu est Wikipédia. Il s'agit là d'une forme de bénévolat particulièrement intéressant, car il n'exige pas un engagement horaire précis, au contraire de la plupart des engagements associatifs : l'utilisateur contribue quand il le peut et sur ce qu'il veut.

5) Une nouvelle « culture décisionnelle » basée sur des algorithmes et des dispositifs intégrés d'acquisition de données de toutes sortes peut certainement améliorer le fonctionnement des services publics, mais avec quelles conséquences, finalement, pour la démocratie et les libertés individuelles ?

Le principal danger pour la démocratie est, à mon avis, la possibilité pour le pouvoir politique de se cacher derrière des processus techniques présentés comme scientifiques, objectifs et inéluctables, alors que leur conception et leur opération impliquent des choix politiques, ainsi que des choix d'analyse sociale, psychologique ou économique, qu'il faudrait pouvoir discuter. À ce sujet, on ne peut que s'inquiéter de la méconnaissance par les citoyens, mais aussi apparemment par certains décideurs politiques, de certains faits statistiques de base : par exemple, que les résultats d'un algorithme « précis à 99,9 % », appliqué à repérer un phénomène très rare dans la population, seront en grande majorité constitués de faux positifs.

Plus généralement, même si un dispositif technique était bien conçu, il se poserait la question de son acceptabilité sociale, notamment en cas de controverse. À notre époque, la parole de l'expert est remise en cause : on lui attribue, à tort ou à raison, des liens avec telle ou telle industrie, tel ou tel parti politique... Ses analyses sont mises sur un pied d'égalité avec celles d'amateurs, qui peuvent avoir un certain succès médiatique. Un raisonnement scientifique trop long ou faisant appel à trop de prérequis sera incompréhensible pour la population. Ceci conduit à des débats où l'on oscille d'une position béate (« la méthode des algorithmes permettra de détecter les terroristes ») à l'hostilité, voire au néo-luddisme, sans que l'on ne puisse vraiment discuter des avantages ou inconvénients de telle ou telle approche.

Les libertés individuelles peuvent en effet être mises en danger par la collecte massive d'informations. Tandis que dans les régimes totalitaires des décennies passées cette collecte avait un grand coût humain, son automatisation la rend à la fois indétectable et indolore ; elle est souvent associée à des services très pratiques (communication, cartographie, paiement...). Nous en sommes déjà d'ailleurs au point que certains envisagent la suppression du paiement en numéraire afin de rendre les flux financiers totalement traçables, et qu'on décrit comme comportement suspect le fait d'éteindre son téléphone portable, voire de ne pas en porter. Le gouvernement chinois met en place, dit-on, un dispositif de notation des citoyens, avec divers inconvénients (interdiction de voyager en train ou en avion...) pour les moins bien notés.

Les questions politiques et éthiques de tels dispositifs sont vertigineuses. Faut-il faire confiance aux protections légales et réglementaires sur le droit d'accès à telle ou telle donnée ? Divers scandales ont montré que des personnes obtenaient l'accès à des informations (vidéosurveillance, etc.) auxquelles elles n'avaient pas droit ; si cet accès est techniquement et discrètement possible, il est probable qu'on en abusera officieusement. Le citoyen doit-il conserver une part de vie privée, ou la vie privée a vécu ? Est-il éthique de noter le comportement de chacun à chaque instant ? Et, même si cela l'était, quels mécanismes de correction fournir pour gérer les inévitables erreurs dans les informations collectées et les conclusions tirées ?

Enfin, sous un angle plus technique, il faut rappeler qu'une fois que les données ont été collectées, il faut s'assurer de leur sécurité. Les affaires Wikileaks, Snowden, EternalBlue, avec des fuites d'information détenues par des services secrets, les annonces de piratage de bases de données de mots de passes entières sur de grands sites, nous montrent que cela n'est pas chose facile. Même si nous collectons des informations pour la bonne cause et avons mis en place des garde-fous afin d'éviter leur mésusage en interne, nous ne pouvons forcément exclure que des tiers malintentionnés y aient accès, que ce soit à des fins crapuleuses ou à des fins de manipulation politique ou de chantage.