

Des algorithmes peuvent-ils être éthiques ?

Encore discrets il n'y a que 20 ans, les algorithmes sont désormais devenus **des objets du quotidien**. Peuplant le monde à nos côtés, l'espace qu'ils occupent et les actions qu'ils y mènent ne sont pas sans **susciter des questionnements et des réflexions quant à leur éthique**.



Cet article a été rédigé par Aurélien Grosdidier, chercheur indépendant fondateur de *Latitude77*, un espace de réflexion et de prospective à la confluence de l'anthropologie, de la sociologie et de la philosophie appliquées au questionnement technoscientifique et aux enjeux numériques.

Lors de cette exploration, **la première difficulté rencontrée est ontologique**: les algorithmes sont idéels, l'éthique relève d'une pratique. À cela **s'ajoute une difficulté culturelle**: rares sont les éthiciens qui s'attachent aux différences entre les algorithmes et le code informatique, tout aussi rares sont les développeurs qui font la distinction entre l'éthique et la morale. L'absence de telles distinctions pourtant très basiques empêche de gravir la première marche vers la moindre réflexion.

Ces deux univers se croisent rarement, et s'ils se rencontrent, ce n'est la plupart du temps que lorsqu'ils s'opposent et nous contraignent à devoir *réagir*. Arriver à leur ouvrir un espace de questionnement mutuel implique de donner à voir une complexité qui permet aux expertises multiples de faire entrer leurs regards complémentaires en synergie. **Une réflexion interdisciplinaire est nécessaire**. Appuyée sur les disciplines existantes, elle permet d'éclairer les détails. Je crois que ce serait une erreur de s'arrêter là. Comme l'écrivait avec une pointe d'ironie Gaston de Pawlowski à propos du scientisme:

"démontons et classons minutieusement tous les rouages de notre montre. Il serait bien étonnant qu'au terme de ce processus nous ne sachions pas enfin l'heure qu'il est".

La nature pervasive des technologies desquelles il est question incite à un questionnement systémique, c'est à dire **au-delà des disciplines.**

L'enjeu éthique des données

Commençons pragmatiquement par ce que nous pouvons chacun expérimenter: les données. Elles sont de toutes sortes: nombres, textes, images, sons, vidéos. En très grande quantité, en temps réel, et de nature variée, celles qu'on appelle désormais **#bigdata** sont indifféremment produites par des objets connectés ou par chacun de nous, intentionnellement aussi bien qu'à notre insu. Il est courant de penser que ces *traces* ne sont que des réductions binaires du réel, signifiants orphelins de signifiés, anonymisantes généralisations qui condamnent les **doubles numériques** et les traitements

dont ils sont les objets à n'être que de pâles imitations du monde qui nous entoure et dans lesquelles nous pourrions être enfermés. Rien n'est plus faux: les données manipulées par les algorithmes tirent précisément leur valeur de leur contextualisation grâce aux **métadonnées**:

"Results without context are meaningless".

Ces dernières, toutes aussi numériques, souvent moins réglementées, renseignent sur les données elles-mêmes parfois d'ailleurs si bien qu'elles dispensent de les **regarder en détail**. Par conséquent, celles que nous laissons et que **d'autres** exploitent permettent d'envisager un changement de paradigme de compréhension de (et d'action sur) nos **variabilités interindividuelles**, bien loin des statistiques historiques, ce qui n'est pas sans poser un **certain nombre de défis**.

Le problème que posent ces données n'est pas lié à une impuissance qui résulterait de ce qu'elles sont une réduction du monde, mais précisément à la puissance qui permet aux traitements desquelles elles sont l'objet de réduire le monde à leurs mesures, éventuellement avec notre **complicité**. C'est parce que ces données nous décrivent, ainsi que le monde qui nous entoure, que leur traitement est un enjeu éthique, aussi bien que par l'usage qu'il est possible d'en faire, que par la réalité que cet usage fait advenir et qui concerne chacun d'entre nous.

Algorithmes et au-delà

Un **algorithme** est une

"suite finie et non ambiguë d'opérations ou d'instructions permettant de résoudre un problème ou d'obtenir un résultat".

C'est un objet mathématique, un processus logique comme il en existe depuis bien avant le traitement automatisé de l'information. Autrement dit, une recette de cuisine est un algorithme, tout comme l'est la procédure d'évacuation d'un lieu public, la stratégie d'une entreprise, un code civil, pénal, ou encore une constitution. Ces algorithmes-là sont mis en œuvre par des humains.

Dans un contexte numérique, les algorithmes s'affranchissent en partie des humains, mais ne deviennent performatifs qu'après plusieurs étapes successives. D'abord leur traduction en code informatique, lisible par l'homme. Ensuite, la transformation de ce code informatique en code qu'une machine peut exécuter. Enfin, l'exécution de ce code exécutable, sous forme d'une instance qui opère sur une *infrastructure* en lien avec un contexte opérationnel spécifique: entrées, sorties, interactions. À ces étapes préalables succède une quatrième: la maintenance du code et de l'*infrastructure* qui permet de l'exécuter. Cette dernière étape fréquemment oubliée est **pourtant absolument essentielle** sous peine de **prise de risques considérable**.

Un algorithme, en lui-même, n'est capable de rien d'autre que de nous permettre — au mieux — de saisir une partie de l'intention de son concepteur. Une métaphore simpliste serait de mettre en parallèle l'algorithme et le plan d'un bras robotisé. Le code exécutable serait le bras robotisé lui-même. L'instance serait le bras robotisé allumé, l'*infrastructure* serait l'ensemble de ce qui assure l'alimentation électrique du robot. Le contexte opérationnel serait l'action qu'il mène sur une chaîne de montage. La maintenance serait l'ensemble des procédures visant à inscrire l'action du robot dans la durée.

Le concept de chien n'aboie pas, le plan d'un robot ne fabrique rien. Quiconque souhaiterait comprendre ce que la chaîne de montage réalise, en particulier pour s'assurer de la conformité de l'assemblage réalisé vis-à-vis d'un quelconque critère, ne saurait se satisfaire de regarder les plans

du robot. Autrement dit, **restreindre le questionnement éthique au champ algorithmique reviendrait à se focaliser sur un objet connu depuis longtemps, et simultanément à exclure le traitement automatique de l'information (TAI) puisque ce dernier est *exclusivement en aval* de l'algorithme.**

Le traitement automatique de l'information

C'est principalement ce TAI qui, par ses effets potentiels, pose aujourd'hui question. **Considérer une éthique numérique impose donc d'élargir le champ de vision au-delà des algorithmes**, à chacune des différentes étapes des chaînes ontologique et chronologique dans lesquelles il est impliqué: de l'intention qui préside à son élaboration en amont, à la réalisation de code informatique, celle du code exécutable, celle de l'exécution, celle du contexte d'exécution et celle de la maintenance. **Ce sont par conséquent les logiciels et leurs infrastructures spécifiques aussi bien techniques qu'humaines qu'il s'agit de questionner, et ce dans leurs contextes.**

Le logiciel et les infrastructures

Réaliser un logiciel est très **loin d'être trivial**, et relève la plupart du temps de l'expérimentation itérative qui ajoute des couches de complexités les unes **au-dessus** — plus rarement à côté — des autres, mais trop rarement en **remplacement les unes des autres**. Les comprendre est bien plus difficile encore puisqu'il s'agit d'en **saisir la complexité révélée en totalité, puis de la réduire**. Les logiciels me semblent toujours échapper ne serait-ce qu'en partie à la compréhension : faute de temps, de moyens, de compétences, de volonté, voire simplement à cause d'aléas. J'en veux pour preuve les **multiples bugs** par **exemple** et la **difficulté et les compromis nécessaires** à leur prévention, ou encore la difficulté de **suivi** des **projets** de développement, qui impactent non seulement nos pratiques personnelles mais aussi **nos échanges**, et même la **production des savoirs**.

Cette complexité qui échappe même aux spécialistes qui disposent pourtant des *codes sources* me semble donc de très loin hors de portée de toute compréhension de l'écrasante majorité de ceux auxquels les systèmes sont accessibles, le plus souvent sans accès au code source. C'est important d'en ouvrir le fonctionnement, mais **de loin insuffisant** pour évaluer leur qualité, leur conformité, et donc à plus forte raison leur éthique. Il me semble qu'il n'est d'ailleurs possible de considérer comme éthique un système technique qu'en faisant l'inventaire de toutes ses trahisons possibles pour les exclure une à une. Penser que cette démarche laborieuse soit faisable relève de la spéculation, et qu'elle soit effectuée de l'hypothèse. Elle est tension: sa réalité est toujours à conquérir.

Il me semble par conséquent que l'ouverture puis l'étude de systèmes ne peuvent qu'aboutir à une prise de conscience de leurs limites, sachant que ces limites seront toujours sous-estimées. Construire la connaissance suppose la destruction des croyances. L'acquérir suppose de consentir à payer un prix si élevé que, collectivement, je ne suis pas certain que nous soyons prêts à assumer le coût. Sans compter qu'il faudrait, déjà, que nous manifestions un minimum d'intérêt vis-à-vis des systèmes qui nous entourent, et dont le marketing renforce la perception animiste en ne faisant leur promotion qu'à travers des pratiques, des "solutions" qui les rendent très difficiles à penser en tant qu' "objets techniques" et "problèmes potentiels".

S'il était possible de questionner l'éthique des différentes étapes du cycle de vie d'un logiciel, il faudrait certainement commencer par l'éthique de l'intention qui préside à sa création. Cette intention est jusqu'à présent humaine (et par ailleurs parfois ancrée dans un référentiel **contestable**), mais pourrait **devenir machinique**, ce qui rendrait encore plus difficile voire impossible le questionnement. Concernant la conception des algorithmiques, il faut noter **un effort significatif et prometteur** de la part

de l'*Institute of Electrical and Electronics Engineers (IEEE)*, reste à voir si les points soulevés peuvent être résolus sans porter atteinte à la viabilité économique des logiciels. L'implémentation de ces algorithmes sous forme de code informatique reste toutefois réalisée par des humains, lesquels, s'ils ne sont pour la plupart pas **malveillants**, sont parfois néanmoins **soumis à des pressions**. Une fois réalisés, ces logiciels, même résultant d'une intention et d'une réalisation éthique – donc a fortiori lorsque ce n'est **pas le cas** – peuvent néanmoins être détournés par les humains qui les utilisent, tout comme ils peuvent interagir avec d'autres TAIs, déjà existants ou encore à venir et donc par **nature imprévisibles**. Enfin, la création, la maintenance et le recyclage des logiciels peut impliquer des contraintes, en **particulier sociales**, pour le moins discutables.

Au-delà des algorithmes, ce sont les logiciels et leurs infrastructures qu'il s'agit de questionner. Aussi bien pour eux-mêmes, c'est-à-dire en tant qu'objets techniques, que vis-à-vis de nous, c'est à dire en tant qu'instigateurs de la culture qui découle des pratiques qu'entraînent leurs usages, que pour les intentions qu'ils implémentent et les externalités qu'ils produisent. Ce questionnement doit embrasser tout leur cycle de vie, de leur conception à leur retrait, et s'étendre à l'écosystème de leurs interactions entre eux et avec nous, aussi bien sur les plans individuel et collectif.

Le contexte d'exécution

Le contexte d'exécution des TAIs est primordial. **En ce qu'ils sont automatiques et performatifs, ils s'inscrivent dans un rapport de forces entre ce qui est et ce qui devrait être et qu'ils font par conséquent advenir.** L'espace défini par ces rapports de forces peut échapper à nos perceptions, c'est d'ailleurs le cas dans la plupart des usages: le propre de l'outil adéquat étant de nous dispenser d'avoir à exercer la force nous-mêmes, ne serait-ce qu'en partie.

Ces rapports de forces peuvent être implicites, s'ils sont partagés par les concepteurs et les utilisateurs des TAIs. Dans ce cas-là, ils sont réductibles à des points aveugles individuels et/ou culturels. **Ils peuvent être dissimulés**, principalement pour des raisons stratégiques. Par exemple, les *traces* numériques que nous laissons sont exploitées selon des modalités et pour des finalités qui ne sont **pas toujours explicites**. **Ils peuvent être visibles**, mais **sans nécessairement générer de réaction**, piégés que nous sommes par nos propres **biais cognitifs**. **Ces rapports de forces peuvent être explicites**, mais sans pour autant donner à **entrevoir l'écosystème duquel ils font partie**. Puisqu'ils y sont internes et donc biaisés, ils ne peuvent d'ailleurs nous le révéler qu'en partie, c'est à nous d'en prendre conscience par une distanciation critique.

Ces rapports de forces que sont les TAIs mettent en tension différents acteurs animés d'intérêts différents. Il peut s'agir d'autres TAIs sur lesquels il s'agirait de **prendre l'ascendant** ou qu'il s'agirait de **compromettre**. Il peut également s'agir d'individus ou d'acteurs collectifs. Le cas des TAIs *cryptographiques* est symptomatique. La **cryptographie** est partagée par les individus, les groupes, et les gouvernements. Or, l'effet levier propre à la technologie numérique démultiplie les possibilités de l'individu potentiellement au-delà des possibilités d'un groupe, quel que soit sa taille. Par conséquent, les sphères individuelle, collective et étatique ont en partage un socle technologique si proche que les acteurs correspondants peinent à organiser leurs rapports de forces sans mettre à mal les droits et les devoirs les uns des autres, car leurs propres subsistances sont potentiellement antagonistes et donc affaire de compromis. Des frictions sont inévitables et appellent à de nouveaux espaces de négociation pour arbitrer des situations inédites.

Une éthique, en tant que pratique, est contextualisée et donc relative à des intérêts spécifiques qu'il s'agit d'arbitrer. Par conséquent, certains TAIs seront jugés éthiques par certains acteurs, mais non éthiques **par d'autres**. L'éthique, dans ce cas-là, dévoile qu'elle n'est pas liée aux TAIs — quand bien même ils ne sont ni neutres ni dénués d'intentions —, mais bien au regard que l'on porte sur eux. Ils disent les désaccords entre nous. L'automatisation qu'ils permettent nous accule à investir la

responsabilité qu'impose l'exercice la liberté de concevoir des algorithmes. Que les conséquences de cette automatisation nous interrogent devrait nous conduire à questionner la culture depuis laquelle ils émergent et qu'ils contribuent désormais à faire advenir. Or cette culture est issue d'algorithmes tout à fait humains et biaisés qui régissent nos **rapports sociaux** et à **nous-mêmes**. Autrement dit, les TAIs rendent visibles à grande échelle et potentiellement instantanément des divergences *entre nous* et *en nous*. Par conséquent, qu'une éthique unifiée des TAIs soit possible supposerait d'aligner en amont nos intérêts, ce qui supposerait que nous soyons **alignés avec nous-mêmes**. Si ce n'est pas possible, de quelle éthique parle-t-on ?

De l'éthique au possible au Politique

Derrière l'enjeu éthique des TAIs se trouvent donc nos subjectivités et nos incohérences. Si nos regards sont multiples et partiels, se pose la question de définir ce que serait une éthique numérique. Devrait-elle être aussi éthique que nous, c'est-à-dire inclure nos **propres biais**, y compris **celui d'oublier avec le temps ses actions non éthiques** ? Si elle devait être davantage éthique que chacun de nous, qui donc serait en mesure de l'évaluer ? Si elle devait être en pratique moins éthique que nous, quel intérêt aurions-nous à lui déléguer quoi que ce soit ?

L'arbitrage des rapports de forces est nécessaire **même au sein d'une communauté homophile**, donc a fortiori en l'absence d'éthique et d'intérêts partagés. Ces arbitrages font intervenir des institutions qui doivent exister et rendre opérants les espaces de négociations susmentionnés. Ces institutions doivent être communes, c'est-à-dire que leur légitimité doit être reconnue par *tous ceux qui sont concernés par les rapports de force qu'elles arbitrent*. La question devient donc Politique, au sens du vivre ensemble, du contrat social **à penser** au-delà de nos divergences et de nos juridictions.

En 2000, **Lawrence Lessig** écrivait *code is law*, soulignant que le code s'était progressivement établi comme régulateur des comportements en ligne, se substituant de fait à la loi. Fin 2016, un **article de Primavera De Filippi** suggère que le registre public distribué qu'est la **blockchain** pourrait être à l'origine de solutions techniques permettant au code de devenir un moyen d'exprimer la loi (*law is code*). A minima, **un tel horizon pose question**, entre autres car il rend impossible le recours autrement — au mieux — qu'a posteriori, et parce que le droit actuel a cet avantage d'être plastique et de pouvoir s'améliorer **en s'adaptant**, comme le souligne **Antoinette Rouvroy**.

L'automatisation permise par le code informatique est absolue. Elle est au-delà de l'éthique car cette dernière promeut certaines valeurs plutôt que d'autres. Elle est au-delà de la morale car cette dernière partitionne l'espace des possibles en bien et mal. Le code informatique, lui, ne cherche pas, latéralement, à influencer les comportements mais se place **potentiellement au-dessus et au-dessous** de ces derniers, en modelant l'espace des possibles lui-même : il encadre les valeurs, il encadre le bien et le mal, au service d'**intérêts particuliers**. En même temps qu'il ouvre un espace possible, les TAIs en **ferme** un **autre**.

L'automatisation: quelle Politique par quels hommes ?

En creux, ce questionnement de l'éthique du numérique est donc bien principalement — et pas uniquement — le nôtre: la technique est un miroir où se reflètent les divergences de nos intentions particulières, individuelles et collectives. Par conséquent se pose la question de notre volonté à **prendre conscience** de **nos biais**, pour les dépasser.

Le problème est donc humain avant d'être technologique. L'aborder sous l'angle technologique en cherchant à encadrer ses conséquences indésirables dans une marge — **acceptable collectivement**

mais qui ne le sera sans doute (plus?) jamais individuellement — n'est un palliatif. Les TAIs sont des pharmakons. Ils relèvent autant de la médecine dans leurs usages que de la pharmacie dans leurs conceptions. Les serments d'Hippocrate et de Galien devraient par conséquent être transposés dans le monde qu'ils contribuent à faire émerger.

Contact: [Latitude77](#)

Cet article est également disponible sur [Usbek & Rica](#).