

# Réseau international sur l'éthique — 1<sup>er</sup> juin 2021

## Éthique du numérique et de l'intelligence artificielle

Jean-Gabriel Ganascia

Professeur Sorbonne Université – directeur de l'équipe ACASA du LIP6

Ex-président du COMETS (comité d'éthique du CNRS), Président du comité d'éthique de pôle emploi, membre du CNPEN (comité national pilote d'éthique du numérique), président du comité d'orientation du CHEC (cycle des hautes études de la culture), président de l'AFAS (association française pour l'avancement des sciences)

### Préambule

En préalable, pour essayer de traiter cette question, je crois utile de commencer par préciser le sens des mots, tout au moins, sinon de les définir, car que ce soit dix minutes ou une vie n'y suffirait pas, tout du moins de les approcher. Nous verrons ensuite si ce n'est ce qu'est l'éthique du numérique, du moins ce qu'elle n'est pas.

### Numérique

Commençons par le plus simple : le numérique. Le mot est ancien, mais le sens que nous lui attribuons ici apparut très récemment. Auparavant, lorsque j'ai commencé, on parlait d'informatique, éventuellement de calcul numérique, de méthodes numériques ou d'art numérique, mais le terme conservait un sens restreint. Désormais, il recouvre l'ensemble des applications des technologies de l'information. On peut se demander quand le basculement de signification s'opère. Selon moi, cela tient à la diffusion du web dans la population, au début du XXI<sup>e</sup> siècle. Ce qui importe pour nous, c'est que cela recouvre différents champs des technologies de l'information :

- L'intelligence artificielle,
- Les interactions entre humains et ordinateurs,
- La robotique,
- L'internet des choses, etc.

## Éthique

Si le numérique se définit relativement facilement, il en va tout autrement de l'éthique. Commençons par rappeler que le terme vient de **ἦθος** en grec qui signifie mœurs, coutumes, habitudes. Initialement, l'éthique tient à la tradition : on fait ce que ses pères ont fait... D'ailleurs, la première attestation du mot ἦθος en grec se retrouve chez Homère où elle désigne un lieu familier aux animaux où ils retournent régulièrement. Plus tard, pour les philosophes, en particulier pour Aristote, le mot ἦθος convoque aussi, d'un côté, le **désir** et la **volonté** et, d'un autre côté, la rationalité requise pour délibérer et choisir les actions qui satisfont nos désirs et réalisent nos volontés. L'éthique correspond à la philosophie pratique, à savoir à la réflexion sur le bon agir. Implicitement, s'il y a éthique, c'est qu'il y a **liberté**, sinon, il n'y aurait pas de réflexion. Toujours selon Aristote, l'habituation y prend une part importante. En dépend l'acquisition des vertus éthiques comme la justice, la tempérance, la prudence — c'est-à-dire la capacité à délibérer efficacement à partir d'une simple opinion — et la bienveillance. On retrouve cette tension entre *deux composantes de la morale*, l'une qui vient de la société, de la tradition, de l'autorité des anciens, l'autre de l'individu, de sa volonté, de son aspiration propre, chez beaucoup de philosophes. Citons par exemple Bergson, *Les deux sources de la morale et de la religion*.

On conçoit que l'éthique du numérique ait partie liée à la rationalité et à la volonté. Et, c'est tout particulièrement le cas dans les interactions humains-machines puisque celles-ci doivent aider les humains à soumettre les machines à leurs désirs en évitant qu'elles ne les piègent, voire que certains les utilisent pour les piéger. En revanche, a priori, on imagine difficilement que *l'éthique du numérique repose sur les traditions*, car rien n'est plus comme avant dans le monde numérique et la sagesse de nos ancêtres nous paraît désormais bien inutile. Pourtant, cette signification n'est pas à négliger. On sait l'importance des facteurs culturels dans l'appropriation des technologies. En témoigne le concept d'*affordance*, à savoir l'invitation des objets à être utilisés de telle ou telle façon selon le contexte. Cette notion introduite par un psychologue, Gibson, dans les années cinquante, a été reprise par Don Norman dans les années quatre-vingts comme fondement à sa réflexion sur les interfaces graphiques. De même on sait l'importance des études d'usage dans le cycle de conception des technologies du numérique. Enfin, le numérique transforme les relations interhumaines et donc les habitudes, les coutumes, à savoir les mœurs.

### Pourquoi y-a-t-il un problème ?

Le fonctionnement des machines isolées pose des problèmes spécifiques, parfois d'ordre mathématique, d'autres fois d'ordre physique ou technique ; il n'y a rien là d'éthique : rien d'autre qu'une consécution d'opérations matérielles qu'il s'agit de régir au mieux.

Comme souvent, les difficultés surgissent à la jonction, aux bords, sur les marges. En l'occurrence, il s'agit des liens entre humains et machines. Nous venons de voir que le

numérique prend un sens nouveau au moment où les technologies de l'information se sont massivement disséminées dans la société. On conçoit que ce soit à ce moment-là que les questions d'éthique du numérique commencèrent très sérieusement à se poser, même si certaines, comme Norbert Wiener les avaient pressenties.

**Les risques sont alors de deux ordres.** En premier lieu, **on craint que les machines prennent la place des humains**, qu'elles les gouvernent ou plus exactement, pour reprendre les craintes exprimées par Georges Bernanos en 1947, qu'elles les contraignent à vivre sur un rythme qui n'est pas le leur et, qu'en conséquence, elles les asservissent en en faisant des êtres à leur image. Sans aller jusqu'à prendre au sérieux l'hypothèse d'une domination des humains par des machines qui décideraient à leur place, comme le manifeste régulièrement Elon Musk, il convient de s'assurer que nous demeurerons toujours libres et responsables de nos actes.

En second lieu, il se pourrait que **les machines deviennent des instruments au service de nouvelles puissances** qui s'imposeraient par leur entremise, en passant outre aux pouvoirs traditionnels, par exemple aux institutions des régimes représentatifs. À titre d'illustration, les réseaux sociaux risquent d'exercer une censure automatique sur les informations que nous diffusons à la place des gouvernements qui ne parviennent plus à assumer le pouvoir de censure qui leur incombait auparavant. La question devient alors éminemment politique.

## Questions éthiques

Comme nous l'avons vu, les questions éthiques se posent à la jonction, dans un sens et dans l'autre, autrement dit,

1. Des humains vers les machines, à savoir lors de la commande des machines, et, en retour,
2. Des machines vers les humains

### Humains vers machines: commandes

Concernant la commande des machines, nous distinguerons une *commande volontaire*, par la parole, la voix ou le geste, d'une *commande involontaire*, par exemple par les émotions qui nous trahissent, par notre posture ou par l'activité de notre cerveau. Dans le cas de **commandes volontaires**, on s'assurera de la fidélité des commandes finales aux intentions de ceux qui les ont exprimées. Ainsi, pour la commande par la parole, il faut s'assurer que le résultat soit conforme à ce qui a été dit. Cela fait l'objet des nombreuses évaluations essentielles à la fois pour la sécurité et pour la conformité aux normes éthiques des systèmes.

Dans le cas de **commandes involontaires**, les choses paraissent plus délicates. Prenons l'exemple de la reconnaissance faciale ou de la reconnaissance d'émotions. Lorsque cela se fait à notre insu, par exemple dans la rue, cela pose indéniablement des questions dont on doit débattre. En même temps, on ne se plaindra pas d'avoir une machine qui utilise la reconnaissance faciale pour vous authentifier et déverrouiller automatiquement votre écran d'ordinateur. De même, la reconnaissance d'émotions inquiète, sauf, par exemple, si dans une voiture, un dispositif détecte que vous avez sommeil et vous évite un accident...

## Retours des machines

Pour les retours des machines, on distinguera **trois éventualités**, selon que ceux-ci

- *affectent l'opérateur*, par exemple dans les dispositifs dits haptiques ou à retour d'effort, indispensables pour la chirurgie ou pour les instruments musicaux,
- *portent sur le monde extérieur*, par exemple sur la commande d'un robot ou
- transmettent à des tiers des informations personnelles ou collectives.

Dans le premier cas, il convient de procéder à des évaluations scientifiques sur la portée de ces retours. Soulignons en le bénéfice que l'on en tire, par exemple lorsque, dans le domaine scientifique, on visualise des molécules pour donner une intuition de phénomènes microscopiques.

Pour le deuxième, cela rejoint toutes les questions d'éthique relatives aux automatismes, comme par exemple pour la voiture ou les armes dites autonomes. Dans ce cas, il faut s'assurer que l'opérateur humain assume toujours la responsabilité de ses décisions et donc qu'il comprenne parfaitement l'action qui répond à ses commandes.

Enfin, dans le troisième cas, cela peut poser des problèmes considérables, surtout en cas de commandes involontaires, par exemple lors de reconnaissance d'émotions, voire d'intentions avec les interfaces cerveau-ordinateur (ICO).

## Difficultés d'une éthique pour le numérique

Il convient toutefois d'insister sur la difficulté de la réflexion éthique sur le numérique qui tient en grande partie à l'impossibilité à anticiper toutes les applications que l'on peut faire d'une technologie et à décider a priori de ce qui est problématique. Les multiples **comités d'éthique du numérique** qui se sont constitués ces dernières années ont souvent essayé d'établir des régulations très lourdes en se fondant sur des grands principes juridiques, voir en dressant des analogies entre la bioéthique et l'éthique du numérique. Or, régulations se sont révélées à la fois contraignantes et inefficaces. De plus, en imposant des normes et des

règles, elles passent outre les institutions démocratiques et la souveraineté. De plus, au plan personnel, elles ont tendance à évacuer la réflexion éthique. Rappelons le:

- L'éthique n'est pas le droit; les comités d'éthique ne sont pas là pour légiférer
- L'éthique n'est pas la technique non plus; nous ne sommes pas là pour porter un diagnostic sur la viabilité des technologies

Dans le même temps qu'ils légifèrent, ces comités conçoivent des normes rigide, ils ne dénoncent pas les projets très choquants de grands acteurs du numérique, par exemple de Marc Zuckerberg ou d'Elon Musk sur les interfaces cerveau ordinateur. Rappelons que Mark Zuckerberg travaille depuis plusieurs années sur des ICO qui liraient dans nos pensées afin d'identifier nos désirs et « fluidifier » notre relation avec son réseau social. De son côté, Elon Musk aspire à fabriquer des implants cérébraux au moyen desquels on mettra les cerveaux en ligne sur le réseau. En tant que tels, ces projets paraissent bien effrayants puisqu'ils se proposent de faire intrusion dans notre intimité et, par là, de supprimer toute intimité. Comment ne pas les condamner au plan éthique?

Pour autant, il ne faudrait pas condamner toutes les applications des ICO, loin de là. Ainsi, il existe des travaux potentiellement très utiles portant sur l'utilisation de ces technologies pour la rééducation de patients cérébro-lésés suite à un accident vasculaire cérébral ou à un traumatisme crânien.

La difficulté avec les projets de régulation tient à trois choses :

1. En premier lieu, ces projets imposent des règles à la fois aux États et aux individus. Du fait de leur statut de normes, ces règles se présentent comme fondamentales, au point que l'on ne peut y déroger ; elles échappent alors à la souveraineté des États, c'est-à-dire à leur capacité à se doter eux-mêmes, par leurs institutions propres, de leurs lois. De même, au plan éthique, cela restreint la liberté des individus.
2. . En deuxième lieu, on se trouve dans une situation paradoxale où l'on affirme que l'éthique se fonde sur des droits ou des principes fondamentaux. Or, la réflexion éthique est antérieure à la loi ; la loi prend en considération des considérations éthiques, même si elle répond souvent à d'autres impératifs d'ordre politique ou économique.
3. Enfin, alors que beaucoup de comité se fondent sur les principes d'éthique de la biologie et de la médecine. Or ce qui advient dans l'univers numérique n'a rien à voir avec les questions éthiques posées par la biologie ou la médecine ; les notions d'autonomie, de justice et surtout de bienfaisance ou de non-malfaisance n'ont pas grand sens.

Au reste, les grands principes généraux qui ont servi de guide aux comités de régulation **entrent en conflit les uns avec les autres** : comment être transparent sans faire intrusion dans l'intimité des individus ? Comment assurer la sécurité de tous tout en respectant la vie privée (nous avons un excellent exemple de ce problème avec les questions de traçage) ? Comment refuser toute discrimination tout en adoptant une politique inclusive ? Etc. Or, toute la difficulté tient à l'arbitrage entre des différents principes. Et, de cela, les comités soi-disant d'éthique ne disent rien.

Pour conclure, sur ce registre, et même s'il y aurait beaucoup à dire, voici une citation de Vladimir Jankelevitch qui devrait nous donner à réfléchir : *le mal, c'est la disjonction des vertus, c'est d'avoir une vertu sans les autres... le mal c'est d'être courageux sans être sincère, sincère sans être juste, etc. Le mal, c'est la dissociation des valeurs, une valeur dissociée des autres devient méchante, devient mauvaise...*

Je crains que, trop souvent, dans les comités d'éthique aient tendance à privilégier un point sur les autres...

## Références

- Ganascia J.-G., *Servitudes virtuelles*, Éditions du Seuil, Collection Sciences Ouvertes, 2022. Cet ouvrage a reçu le prix du livre FIC 2022, mention recherche universitaire.
- *Ce matin, maman a été téléchargée*, Buchet-Chastel, 2019 (roman écrit sous le nom de plume de Gabriel Naéj)
- Ganascia J.-G., *Le mythe de la Singularité : faut-il craindre l'intelligence artificielle ?*, Éditions du Seuil, Collection Sciences Ouvertes, 2017. Cet ouvrage a obtenu le prix Roberval grand public.
- Ganascia J.-G., *Intelligence Artificielle : vers une domination programmée ?*, Le Cavalier Bleu, Collection Idées reçues, 2017