



**HAL**  
open science

## Du contact éthique à l’heure des dispositifs “ autonomes ”

Fabrice Métais

► **To cite this version:**

Fabrice Métais. Du contact éthique à l’heure des dispositifs “ autonomes ”. Revue française d’éthique appliquée, 2018. hal-02102909

**HAL Id: hal-02102909**

**<https://hal.science/hal-02102909>**

Submitted on 17 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.

Fichier auteur, dernière version avant publication dans la *Revue française d'éthique appliquée* vol. 5, no. 1, 2018.

<https://www.editions-eres.com/ouvrage/4222/un-monde-d-automatisation>

Pour citer cet article : Métais, F. (2018). Du contact éthique à l'heure des dispositifs « autonomes ». *Revue française d'éthique appliquée*, 5(1), 54-66.

## **Du contact éthique à l'heure des dispositifs « autonomes »**

Fabrice Métais

Fabrice Métais, Maître de Conférences

Aix Marseille Univ, CNRS, PRISM (Perception, Representations, Image, Sound, Music)

UMR 7061, Marseille, France.

[metais@prism.cnrs.fr](mailto:metais@prism.cnrs.fr)

---

Résumé : Nous proposons d'évaluer les enjeux éthiques liés aux nouvelles technologies automatiques – et notamment celles dont le comportement peut être qualifié d'« autonome » – à partir d'un croisement entre la phénoménologie levinassienne du rapport à l'*autre en tant qu'autre*, d'une part, et une pensée de la technique comme anthropologiquement constitutive et constituante, d'autre part. Qu'en est-il de la responsabilité éthique du sujet lorsque les dispositifs automatiques calculent à l'avance et ainsi conditionnent les possibilités d'agir et de prendre soin ? Qu'en est-il de la dimension matérielle du *pour-l'autre* quand les corps sont pris dans des systèmes dédiés à l'anticipation de leurs besoins ? Lorsqu'ils s'introduisent au cœur de la corporéité relationnelle, les dispositifs automatiques menacent-ils ce fil signifiant par lequel le sujet et autrui sont en contact, ou, au contraire, nous promettent-ils de nouvelles modalités du sens ?

Mots clés : Levinas, phénoménologie, responsabilité, matérialité, machines autonomes, intelligence artificielle

Abstract : Ethical stakes emerging from new automatic technologies can be addressed from the crossing between, on the one hand, the levinassian phenomenology of the relationship to the *other as other*, and on the other hand, an idea of technology as being anthropologically

constitutive. What about the ethical responsibility of the subject when automatic devices compute in advance and thereby format the spectrum of possibilities of acting and taking care ? What about the very concreteness of the *for-the-other* when bodies are engaged in systems dedicated to the anticipation of their needs ? When they enter the heart of relational embodiment, do the automatic devices threaten this significant bound through which the subject and the other are in touch, or, on the contrary, do they promise new opportunities of sense-making ?

Keywords : Levinas, Phenomenology, responsibility, materiality, autonomous devices, artificial intelligence

---

Si l'automatisme n'est pas une invention du XXI<sup>e</sup> siècle, nous assistons bien aujourd'hui à l'agrégation et à la convergence d'un ensemble de faits techniques qui placent les technologies automatiques – et en particulier celles de l'informatique – au cœur d'un faisceau de questionnements éthiques dont il nous appartient de nous saisir. Pour contribuer à l'exploration de ces enjeux, nous proposons de les resituer dans la perspective d'un dialogue entre, d'une part, une approche anthropologique et philosophique du rapport homme/technique, et, d'autre part, une phénoménologie de l'expérience éthique et du rapport à autrui : celle développée dans l'œuvre d'Emmanuel Levinas.

Dans un premier moment, il s'agira donc de replacer l'émergence des dispositifs automatiques contemporains dans la continuité d'une relation très ancienne – originaire, en quelque sorte – entre humain et technique : les propriétés fonctionnelles de l'outil saisi prolongent celles du corps, et sont ainsi constitutives des modalités par lesquelles le monde est vécu, et l'autre rencontré. Il s'agira ensuite d'interroger la spécificité de l'impact que pourraient avoir ces technologies sur la dimension éthique de l'expérience et du sens : d'abord en ce qui concerne l'emprise de l'algorithme sur le régime du possible et la liberté de choix, puis au regard de l'engagement des dispositifs autonomes dans la dimension matérielle du contact éthique. Enfin, en guise de conclusion, il apparaîtra que les craintes que font planer les technologies automatiques sur l'enjeu humain du sens profitent d'une sorte d'oubli quant au caractère vécu de l'expérience sensible : on fera de la réduction phénoménologique le geste d'une résistance face à la mécanisation du vivant.

## 1. Informatique et corporéité relationnelle

Dans cette première partie, nous voudrions décrire – au point de rencontre entre la phénoménologie levinassienne et la paléanthropologie de Leroi-Gourhan – l'évolution d'une *corporéité relationnelle* originellement augmentée par la technique. Il s'agira de montrer que la dynamique d'*externalisation* correspond à l'affirmation d'un écart signifiant entre le sujet et la pointe de l'outil : l'automatisme et l'algorithme<sup>1</sup> s'insèrent entre le sujet et l'autre et imposent une modalité computationnelle au sein du contact éthique. Aussi, on se demandera si les technologies automatiques contemporaines – quand elles se démarquent par des propriétés d'adaptabilité et d'autonomie – ne portent pas à sa limite le paradigme de l'outil saisi.

### 1.1 Relation éthique et *externalisation*

Emmanuel Levinas (1906-1995) est d'abord phénoménologue, et sa pensée de l'éthique est ancrée dans la description de l'expérience vécue. Plus précisément, c'est comme critique à l'égard des phénoménologies husserlienne et heideggérienne<sup>2</sup> que naîtra et ne cessera de s'affirmer dans son œuvre l'urgence d'une préoccupation à l'endroit de *l'autre en tant qu'autre* – sensibilité face au visage d'autrui et à l'altérité radicale qu'il révèle, d'une extériorité irréductible aux mouvements centripètes de la conscience. Aussi, pour Levinas, l'éthique ne se résout pas dans l'abstraction des concepts, mais dans la concrétude d'un engagement corporel. Entre jouissance égoïque et appel au don, c'est par le fonctionnement et la matérialité des corps – dans l'agencement d'une *corporéité relationnelle* – que le sujet et autrui entrent en contact – se parlent, se touchent, se soignent. Mais cette *corporéité relationnelle* n'est pas seulement celle des corps biologiques. Comme tendent à le montrer les études paléanthropologiques de Leroi-Gourhan, la technique est en quelque sorte une caractéristique zoologique de l'espèce humaine : l'outil saisi est le prolongement fonctionnel du squelette, prolongement du corps au-delà de lui-même dans le fonctionnement et la matérialité de l'objet technique. Cette augmentation du corps au-delà de lui-même, Leroi-Gourhan la nomme *externalisation* (Leroi-Gourhan, 1964a). Ainsi, le corps humain – corps de *cyborg*, hybride organique et inorganique – est originellement augmenté par les dispositifs : depuis le silex taillé jusqu'aux technologies automatiques contemporaines. D'ailleurs, certaines descriptions levinassiennes font signe, elles aussi, vers le rôle des choses et des

---

1 On distingue bien d'abord l'automatisme, d'une part, et l'algorithme, d'autre part. L'automatisme est la "qualité des appareils et installations qui fonctionnent sans intervention humaine" (d'après la définition du dictionnaire Larousse en ligne, <http://www.larousse.fr/dictionnaires/francais/automatisme/>, consulté le 18 octobre 2017). L'algorithme est un "[ensemble] de règles opératoires dont l'application permet de résoudre un problème énoncé au moyen d'un nombre fini d'opérations. Un algorithme peut être traduit, grâce à un langage de programmation, en un programme exécutable par un ordinateur" (d'après la définition du dictionnaire Larousse en ligne, <http://www.larousse.fr/dictionnaires/francais/algorithme/>, consulté le 18 octobre 2017). Les dispositifs qui vont spécifiquement nous intéresser ici déploient de manière automatique des logiques algorithmiques.

2 Et, plus largement, comme critique à l'égard de toute la tradition philosophique occidentale.

outils dans le contact éthique : ils participent eux aussi des modalités fonctionnelles par lesquelles le soin et le don se réalisent. Donner son pain, accueillir dans sa demeure, partager son vêtement : autant de situations concrètes dans lesquelles le *toucher-l'autre* passe par la médiation d'un objet technique. Ainsi, penser aujourd'hui, avec la philosophie levinassienne, l'enjeu éthique de l'automatisation, c'est suivre – au cœur de la *corporéité relationnelle* – les mouvements contemporains de la dynamique d'*externalisation*, soit l'insertion du programme et d'une efficacité algorithmique au cœur du dispositif *cyborg* par l'intermédiaire duquel le sujet et l'autre entrent en contact.

## 1.2 D'une qualité computationnelle de l'écart

L'évolution du système technique, depuis les premiers hominidés jusqu'à nos jours, réalise une prodigieuse diversification des possibilités d'agir, de percevoir et d'être affecté par l'extériorité. En suivant, avec Leroi-Gourhan, la dynamique d'*externalisation*, on repère – depuis le silex taillé jusqu'aux machines informatiques – différentes étapes qui correspondent à l'augmentation d'un écart entre la main du sujet et la pointe de l'outil (Leroi-Gourhan, 1964b). Depuis l'outil manipulé en motricité directe (par exemple, le *chopper* ou le couteau), puis en motricité indirecte (par exemple, l'arc et la flèche, la flexion du bois de l'arc opérant comme réserve intermédiaire d'énergie externalisée), jusqu'à l'*externalisation* du processus moteur par le recours à une source d'énergie externe (motricité animale, moulins, machine à vapeur, etc.) : l'écart entre le sujet et le monde qu'il constitue ne cesse de grandir et de se diversifier dans ses modalités. Quand la main déclenche un processus programmé – dernier stade d'*externalisation* identifié par Leroi-Gourhan en 1964 – c'est l'efficacité d'une mémoire qui s'insinue entre le sujet et la pointe de l'outil. Aussi la médiation technique – qui sépare, autant qu'elle unit, le sujet et son monde – n'est bien sûr pas à concevoir comme la dénaturation d'une immédiateté qui aurait à être synonyme d'authenticité. L'écart<sup>3</sup> – distance, délai – est originaire, positivement créateur de sens, constitutif du monde propre de l'humain. Il est l'hybridation, au cœur de l'expérience, de la présence et de l'absence. Entre le sujet et autrui, l'expérience éthique est la traversée d'une non-immédiateté – spatiale, temporelle : elle ne saurait se résoudre dans un pur présent sans épaisseur, sans distance et sans délai. Parce qu'elle traverse un tel écart, elle prend le sens – phénoménologiquement – d'une urgence, d'une impatience, d'un problème<sup>4</sup>.

---

<sup>3</sup> Ou la *différance* si on s'autorise ici à reprendre le mot de Derrida.

<sup>4</sup> Par exemple, dans une communication érotique épistolaire, le délai nécessaire à l'acheminement du courrier est constitutif de la volupté de l'expérience. De même, lorsque je reçois un appel à l'aide par téléphone, la possibilité du soin n'est pas immédiate : aussi l'expérience prend le sens d'une urgence.

La question face à laquelle nous nous trouvons aujourd'hui est celle de savoir si les technologies numériques, par les spécificités de leurs fonctionnements, nous mèneraient à devoir considérer une « qualité » nouvelle de cet écart : quand un traitement automatique d'information participe de l'efficacité du geste, devra-t-on penser, aux côtés de la spatialité et de la temporalité, une « qualité » proprement computationnelle de l'écart séparant le sujet et le monde qu'il constitue ?

En effet, on pourrait d'abord penser que les technologies numériques ne font que poursuivre – sans y ajouter de nouvelle dimension – le mouvement de l'*externalisation* : elles participent ainsi – mais sans plus – à l'augmentation des distances atteignables et à la suspension des gestes dans des mémoires. En un sens, un courriel est bien comme une lettre papier qui, simplement, voyagerait plus vite<sup>5</sup>. Mais si les technologies numériques devaient être porteuses d'un enjeu anthropologique nouveau, c'est au contraire parce que la propriété spécifique de leur fonctionnement – c'est à dire la computation sur des symboles – pourrait produire un changement de qualité – et non plus seulement de quantité – dans l'écart signifiant qui sépare, autant qu'il unit, le sujet et son monde, le sujet et autrui.

Mais l'enquête visant à analyser, dans leur globalité, les enjeux phénoménologiques de l'insertion d'une couche computationnelle dans la corporéité constituante dépasse largement le cadre de cette contribution<sup>6</sup>. Nous voudrions ici concentrer notre intérêt sur une disposition particulière de certaines technologies automatiques contemporaines – et qui va outre la dimension proprement computationnelle de leur fonctionnement : une certaine aptitude à développer, dans une phase donnée de leur fonctionnement, un comportement que l'on peut qualifier d'« autonome » ou d'« adaptatif ».

### 1.3 Dispositifs autonomes

C'est à la croisée de plusieurs faits techniques qu'émergent des dispositifs dits « autonomes », « adaptatifs », « intelligents ». Parmi ces faits, il faut relever d'abord le développement considérable, depuis le milieu du XXe siècle, de l'ingénierie logicielle. On notera en particulier l'apport des échanges avec les recherches en intelligence artificielle : la technologie algorithmique permet aujourd'hui – grâce à des logiques et des méthodologies très variées – d'appréhender des problèmes, des situations très diverses : depuis le plus

---

<sup>5</sup> De même, à un premier niveau, peu de choses distinguent une communication téléphonique par liaison analogique ou par liaison numérique : il s'agit toujours de communication orale à distance. *Idem* pour la transmission d'image : la télévision numérique permet certes de diffuser plus de chaînes mais il s'agit toujours de transmettre à distance des flux vidéos.

<sup>6</sup> Pour une telle étude, on pourra notamment consulter (Bachimont, 2000).

abstrait des jeux combinatoires (comme le jeu de go, par exemple) jusqu'à certaines tâches les plus directement en prise avec le monde, comme la vision par ordinateur, la reconnaissance de la parole naturelle, le pilotage automatique de drones en environnement complexe, etc. Ainsi, par exemple, les logiciels dits à « apprentissage machine » permettent de résoudre certains problèmes originaux – comme certaines tâches d'analyse automatique d'image – grâce à une stratégie originale qui consiste en l'occurrence à construire, à partir de larges bases de données, des modèles prédictifs, sans avoir recours à un modèle explicatif.

Second fait technique déterminant : la « loi de Moore » ne s'étant pas massivement démentie jusqu'à présent, une puissance de calcul considérable est aujourd'hui disponible, dispersée autour du globe, dans les serveurs, les ordinateurs personnels, les téléphones portables, etc. L'augmentation exponentielle de la puissance de calcul disponible permet d'exécuter des algorithmes toujours plus gourmands, opérant sur des systèmes de données toujours plus riches.

Troisième fait technique remarquable : à l'ère du *big data*, les propriétés du monde et des êtres qui le peuplent sont quantifiées et enregistrées de manière massive. De très larges bases de données sont alors disponibles et nourrissent les algorithmes à apprentissage, participant ainsi de leur efficacité<sup>7</sup>.

Quatrième fait technique : la collecte et la circulation de ces données s'appuient sur l'infrastructure de larges réseaux numériques interconnectés à l'échelle de la planète : câbles de cuivre, fibres optiques, liaisons hertziennes, etc. Ces réseaux permettent également de partager la puissance de calcul (*cloud computing*).

Enfin, cinquième fait technique remarquable : l'efficacité computationnelle des technologies informatiques est reliée au monde par un panel toujours plus large de capteurs<sup>8</sup> – caméras, microphones, antennes et puces RFID, instrumentation diverse – et d'actionneurs – dispositifs moteurs, diffusion de contenus, etc. Cette amplification de l'interface entre la machine et son extériorité fait l'augmentation de la prise par laquelle l'efficacité computationnelle embrasse la matérialité du monde.

---

<sup>7</sup> On peut citer l'exemple tout à fait caractéristique des algorithmes dédiés à la reconnaissance automatique de la parole : c'est tout particulièrement grâce à la disponibilité de très larges banques d'enregistrements de voix – accompagnés parfois de leurs retranscriptions textuelles – que les algorithmes ont pu « apprendre » à distinguer de mieux en mieux, dans le flux audio, les différents phonèmes prononcés et ainsi acquérir finalement un niveau de robustesse suffisant pour participer à l'interface homme/machine.

<sup>8</sup> Le *capteur* est ici un « organe qui élabore, à partir d'une grandeur physique, une autre grandeur physique, souvent de nature électrique, utilisable à des fins de mesure ou de commande ». (d'après la définition du dictionnaire Larousse en ligne, <http://www.larousse.fr/dictionnaires/francais/capteur>, consulté le 18 octobre 2017)

La congruence de ces cinq faits (ou ensembles de faits) techniques<sup>9</sup> confère aux dispositifs automatiques contemporains une capacité certaine d'anticipation et d'adaptation à différentes situations, à différents environnements. À partir des informations qu'ils manipulent, par les raisonnements qu'ils produisent, et grâce à leurs organes<sup>10</sup> perceptifs et actifs, ils s'engagent dans le monde physique et prennent des décisions, de telle sorte que leur comportement peut être appréhendé comme « autonome ». Dès lors, on peut s'interroger : le paradigme de l'outil saisi<sup>11</sup> n'atteint-il pas ici une sorte de crise ? Quand la dynamique d'*externalisation* correspond à l'affirmation d'un écart signifiant de non-immédiateté entre le sujet et le monde, les dispositifs « autonomes » ne quittent-ils pas la main humaine de manière plus radicale que la flèche de l'archer ou la lettre de l'écrivain ? Quand ils s'insèrent entre le sujet et autrui, ne compromettent-ils pas l'effectivité du contact éthique, ne menacent-ils pas ce qui relève – pour Levinas – d'une dimension proprement humaine de l'existence ? Ou au contraire, ces dispositifs nous promettent-ils des modalités nouvelles de ce contact, des modalités nouvelles du sens ? Comment saurons-nous distinguer une technique qui menace d'une technique qui crée ?

## 2. Éthique sans contact

L'éthique levinassienne est avant tout descriptive, logée dans le creux d'une phénoménologie. Aussi, elle ne relève pas d'un caractère prescriptif : elle ne donne pas de règle pour discriminer *a priori* ce qui est bien et ce qui est mal. Ainsi, on ne se demandera pas si telle ou telle technologie automatique pourrait avoir une incidence bonne ou mauvaise pour telle ou telle situation. La question à laquelle nous mène la philosophie levinassienne est celle de savoir si ces technologies – en interposant une couche computationnelle entre le sujet et le visage d'autrui – pourraient menacer le déploiement, au sein de l'expérience vécue, du régime éthique – soit la dimension proprement humaine du sens. Il ne s'agirait pas, à proprement parler, d'un problème éthique, mais de la question de la disparition potentielle de l'éthique même<sup>12</sup>. Sans verser dans le catastrophisme, ou ce qui pourrait ressembler à une

---

9 Sans doute la liste pourrait être plus longue : on pourrait notamment penser à l'apport des accumulateurs lithium-ion qui participent à la mobilité des dispositifs, et donc à la collecte de données géo-localisées.

10 Par analogie avec le domaine de l'anatomie, on désigne ici par *organe* : « [dans] une machine, [un] ensemble de pièces assurant une fonction particulière (commande, guidage, transmission, assemblage, etc.). » (d'après la définition du dictionnaire Larousse en ligne, <http://www.larousse.fr/dictionnaires/francais/organe>, consulté le 18 octobre 2017)

11 Il faut signaler ici que ledit « paradigme de l'outil saisi » – soit la prise en considération de la prolongation fonctionnelle du corps dans l'outil – n'est pas le seul point de vue envisageable pour aborder les problématiques de la co-individuation homme-technique. En particulier, on peut penser que l'abord du numérique et de l'automatisme comme constituant un *milieu* technique au sein duquel l'individu psychique et social se réalise, pourrait être un cadre d'analyse particulièrement fécond pour étudier les problématiques éthiques liées à ces technologies. (Petit, 2013)

12 Que cette dernière question soit, ou non, elle-même une question éthique relèverait, en quelque sorte d'une éthique de second ordre.

technophobie, nous cherchons ici à dégager un point de vue original pour éclairer à nouveaux frais certaines peurs et certains espoirs associés à l'émergence de ces technologies spécifiques.

Afin de poursuivre notre enquête, nous voudrions repérer deux fronts par lesquels les machines automatiques contemporaines pourraient corrompre la dimension éthique de l'expérience : premièrement, par leur emprise potentielle sur le régime du possible et la liberté de choix, et deuxièmement, par l'engagement des automates dans la dimension matérielle du contact entre le sujet et autrui.

## 2.1 La bonté automatique

La capacité à discriminer parmi une multitude de choix possibles est une des caractéristiques principales de nos technologies algorithmiques. Or, l'engagement de l'automatisme dans le régime du possible peut avoir des conséquences plus ou moins directes sur des choix liés à des enjeux éthiques. C'est le cas bien sûr des dispositifs explicitement conçus pour l'aide à la décision : pour les diagnostics et prescriptions médicales<sup>13</sup>, pour les procédures pénales<sup>14</sup>, pour les transactions financières<sup>15</sup>, pour le recrutement en entreprise<sup>16</sup>, etc. C'est le cas encore lorsque des dispositifs « autonomes » ont à prendre « par eux-mêmes » des décisions pouvant impacter des vies humaines : par exemple lorsqu'une voiture sans pilote est sur le point d'être impliquée dans un accident. C'est le cas également lorsque – en lien avec des fins extérieures (commerciales, politiques, idéologiques, etc.) – un algorithme propose à l'utilisateur, à partir de ses choix précédents, un ensemble personnalisé de possibilités pré-calculées.

À partir de la philosophie levinassienne – du moins dans un premier degré de lecture – il pourrait sembler que l'*externalisation*, dans des dispositifs automatiques, de la capacité de choix entre en contradiction avec l'érection d'une subjectivité responsable. En effet, dans la pensée levinassienne de la subjectivité, éthique et liberté sont unies dans une tension signifiante : le visage ne révèle le sens éthique de son appel qu'en laissant une trace dans

---

13 Dès les années 1970, des systèmes experts, ou systèmes à bases de règles (*Mycin*, *Caduceus*, etc.) sont conçus pour identifier des maladies en liant des symptômes et des règles.

14 Voir, par exemple, les expérimentations menées, au printemps 2017, dans les cours d'appel de Rennes et de Douai avec un logiciel nommé *Prédicite*.

15 Publié le 22 mai 2017, le *rapport du Conseil de Stabilité Financière (Financial Stability Board – FSB) et du Comité sur le système financier mondial (Committee on the Global Financial System - CGFS) sur les nouvelles formes d'intermédiation de crédit portées par les Fintech*, mettait en garde les acteurs du domaine contre de potentiels risques liés à l'utilisation de l'intelligence artificielle par les établissements financiers.

<https://acpr.banque-france.fr/sites/default/files/medias/documents/cgfs-fsb-report-on-fintech-credit.pdf> consulté le 24 octobre 2017.

16 On peut citer l'exemple de l'entreprise *Unilever* qui utilise des algorithmes - notamment de reconnaissance faciale (sourire) - pour pré-sélectionner ses candidats à l'embauche.

l'horizon phénoménologique du sujet, une trace dans l'horizon des possibles. La liberté du sujet n'est pas seulement en opposition avec l'altérité d'autrui : « [autrui] ne limite pas mais promeut ma liberté, en suscitant ma bonté » (Levinas, 1990b). L'appel d'autrui fait entrer la liberté du sujet dans le régime éthique du sens. Et, inversement, si la « bonté [est] toujours plus ancienne que le choix », la liberté est bien la caisse de résonance nécessaire dans laquelle l'appel éthique se laissera seulement entendre. L'apparition du visage tient à la révélation du commandement « tu ne tueras pas » – soit « l'impossibilité de la possibilité » (Levinas, 2004) : l'obligation éthique se détache sur un fond de possibilités factuelles. « L'expression que le visage introduit dans le monde ne défie pas la faiblesse de mes pouvoirs, mais mon pouvoir de pouvoir » (Levinas, 1990a). L'éthique du visage n'est pas en elle-même phénoménale mais elle se révèle en laissant sa trace dans le phénomène, dont la texture est le régime du possible. Si, comme nous allons le voir, s'insèrent entre le sujet et l'autre des dispositifs qui formatent et filtrent le possible, alors l'engagement de la subjectivité dans la dimension éthique du sens pourrait en être affecté. Avec Levinas, la liberté n'est pas en elle-même une valeur éthique, mais c'est une modalité nécessaire à l'engagement de la subjectivité dans le régime éthique du sens. Absolument dépourvue de liberté, la subjectivité resterait transie dans un mode paradoxal de la responsabilité : responsabilité sans possibilité d'agir, comme si la mortalité d'autrui était prise dans une fatalité absolue.

Mais, dans la grande diversité des situations dans lesquelles l'humain et l'intelligence artificielle composent dans l'appréhension du possible, la subjectivité reste, le plus souvent, impliquée. L'emprise de l'algorithme sur la liberté ne se résout pas dans une alternative de type tout ou rien. Mais de manière plus subtile, les dispositifs automatiques semblent affecter le rapport au possible dans sa qualité même. Par exemple, lorsqu'un praticien, engagé dans un processus de décision, utilise un programme de type « système expert », qui s'appuie sur la manipulation de règles connues, il est en mesure de saisir le résultat produit et de faire sien le raisonnement logique qui y aura mené. Il peut alors s'orienter dans le possible *en connaissance de cause*. En revanche – parce qu'ils franchissent, dans leur fonctionnement, un niveau d'abstraction supplémentaire, parce qu'ils produisent un modèle prédictif sans construire de modèle explicatif – les algorithmes de type réseaux de neurones opèrent, en quelque sorte, comme des boîtes noires, et produisent des résultats sans les accompagner d'une justification saisissable. Ainsi, cet exemple illustre comment, en changeant de procédé algorithmique, on change la qualité même du rapport au possible : en l'occurrence en passant de programmes à systèmes de règles à des programmes statistiques, on se prive d'un

raisonnement explicite dans la production du résultat et une sorte d'opacité cognitive s'introduit entre le sujet et son choix.

L'étude exhaustive des modalités par lesquelles les technologies automatiques contemporaines affectent la qualité du rapport entre le sujet et le régime du possible ne peut être menée ici. Nous voudrions plus spécifiquement ici nous intéresser aux dispositifs « autonomes » et mettre en perspective la manière avec laquelle, à travers eux, le sujet et autrui peuvent être encore en contact.

Nous voudrions considérer le cas de la plateforme *Moral Machine*<sup>17</sup> développée par le *Massachusetts Institute of Technology*. Accessible au plus grand nombre via Internet, la *Moral Machine* se présente comme une enquête en ligne – voire une sorte de jeu : face à un ensemble de scénarios où une voiture « autonome » va inexorablement provoquer un accident où des vies humaines (ou animales, parfois) vont être impactées, le sujet doit chaque fois décider parmi deux choix possibles lequel est le moins inacceptable. Par exemple, vaut-il mieux que la voiture écrase un piéton, ou qu'elle tue son passager ? Vaut-il mieux qu'elle écrase une femme enceinte ou deux personnes âgées ? Un enfant qui a traversé la route illégalement ou cinq adultes qui ont traversé légalement ? etc. Le principe du test reprend en quelque sorte le fameux *dilemme du tramway* – il n'y a que des mauvaises réponses – sauf qu'ici on ne demande pas au sujet de se projeter comme engagé en première personne dans une situation de dilemme moral, mais il lui est demandé d'exprimer son avis sur ce qu'une voiture « autonome » devrait faire : quel devrait être son choix parmi les différentes possibilités offertes ? Mais s'il reprend dans son principe le *dilemme du tramway*, le dispositif et l'expérience sont ici totalement différents : en effet, ici la personne qui répond aux questions n'est pas face à un problème purement fictionnel et abstrait. Bien au contraire, ici chaque réponse participe activement à l'apprentissage machine d'un algorithme destiné à être réellement embarqué sur des véhicules autonomes, et qui donc est destiné à être réellement impliqué dans des décisions impliquant la mort d'autrui (Noothigattu & al., 2017). À travers une interface web<sup>18</sup> l'internaute participe – par ses choix – à une opération de *crowdsourcing* dont l'objectif est d'« apprendre » à un algorithme à « résoudre » les dilemmes moraux<sup>19</sup>. L'intelligence artificielle – mais doit-elle encore, dans ce cas, porter ce nom ? – fait ici

---

17 <http://moralmachine.mit.edu/> consulté le 20/07/17. (Bonneton & al., 2016)

18 Éventuellement une interface au design ludique – conformément aux pratiques du *digital labor* dénoncées par Dominique Cardon et Antonio Casilli. (Cardon & Casilli, 2015)

19 Bien sûr, dans ce cas, on pourrait difficilement parler d'« apprentissage » ou de « résolution » : contrairement à des tâches de reconnaissance d'image ou de reconnaissance vocale, face à un dilemme moral, il n'y a pas de compétence spécifique à acquérir.

l'interface entre, d'une part, l'humain, *les humains* – ceux pour qui seulement la responsabilité face à la mort d'autrui peut avoir un sens<sup>20</sup> – et, d'autre part, la pointe de l'outil lorsque, engagée dans la phase « autonome » de son fonctionnement, elle devra s'orienter au sein du possible. Aussi, lorsqu'on prend part à une telle collecte de données, peut naître une sorte de malaise. Ce malaise, c'est précisément la trace du visage d'autrui, le commandement « tu ne tueras pas », l'impossibilité du meurtre, révélé à même le possible. Lorsqu'on nourrit de ses choix une telle « *Moral Machine* », chaque clic a une incidence sur une décision aux conséquences létales. La virtualité des situations présentées dans l'enquête est en prise directe avec la mortalité réelle de futurs accidents de la route. C'est une responsabilité réelle qui s'engage par l'entremise de réponses à des dilemmes fictifs. La technologie engage la responsabilité éthique à travers une hybridation du virtuel et du réel. À travers cette médiation technique complexe – une enquête internet collectant des données pour nourrir une intelligence artificielle embarquée sur des véhicules autonomes circulant dans l'espace public – le sujet et autrui sont bien encore en contact, saisis dans une intrigue ayant pour enjeu la mort et la responsabilité. À ce point, plusieurs problèmes technologiques/éthiques sont à étudier : celui du traitement statistique du dilemme moral quand la responsabilité éthique – dans la philosophie levinassienne – est éminemment individuelle, celui de la temporalité de la décision et le fait que l'intelligence artificielle consacre une sorte de préméditation de l'acte légal, celui encore lié au fait que l'algorithme appelle l'abstraction et l'explicitation de catégories sensées être significatives pour la décision morale (âge, sexe, passager/piéton, classes sociales, etc.) Mais si l'on concentre ici notre intérêt sur un enjeu propre à l'intelligence artificielle, il est intéressant de considérer qu'un tel dispositif est en mesure de produire des décisions – empreintes de la responsabilité humaine – pour des situations pour lesquelles aucune subjectivité n'aura été explicitement interrogée. Pour un algorithme de type réseau de neurones, l'extrapolation, à partir d'un apprentissage sur un nombre fini de cas, vers des situations inédites sera, de plus, marquée par l'opacité cognitive, dont nous avons parlé plus haut, liée au niveau d'abstraction franchi par ce type de programmes. Pour une décision prise par le véhicule « autonome », on ne pourra donner d'autre justification que celle qui associe, d'une part, la logique propre au fonctionnement de l'algorithme à, d'autre part, l'ensemble de données fini et déterminé à partir duquel il aura été « formé ». Il faut noter également que grâce à un tel système, les concepteurs de la « *moral machine* » pourraient se dégager – dans une certaine mesure – de la responsabilité, puisqu'ils ne font que fabriquer un

---

20 Mais prenons ici la précaution de ne pas préjuger de ce que pourrait être la responsabilité éthique pour les individus d'autres espèces animales.

dispositif apprenant, soit, en quelque sorte, une boîte vide<sup>21</sup>. C'est une masse d'utilisateurs anonymes qui – via ses réponses au test en ligne – porte alors une responsabilité « massivement partagée » pour tous les accidents mortels provoqués par les voitures autonomes (du moins pour ceux « pris en charge » par l'intelligence artificielle embarquée). Ainsi, ce dispositif nous permet de concevoir comment un véhicule « autonome » – soit une machine associant l'algorithme à un engagement concret dans le monde physique – peut s'interposer entre le sujet et la vulnérabilité d'autrui. Il montre comment l'expérience éthique se trouve saisie dans une logique computationnelle. Il montre comment, dans la continuité du mouvement d'*externalisation*, les dispositifs « autonomes » modifient les modalités de l'engagement subjectif dans le possible, modifient la qualité de cette caisse de résonance dans laquelle l'appel d'autrui se laissera seulement entendre. Aussi, on peut encore s'interroger : le contact éthique – le lien de sens entre le régime du possible et l'interdit du meurtre – n'est-il pas finalement rompu par cette épaisseur technologique ? La liberté et la responsabilité subjective ne se trouvent-elles pas noyées dans la masse des données et dans les statistiques, dans les calculs, dans la ludification des interfaces, dans l'abstraction des catégories ? Ou est-ce que, finalement, la participation – aussi infime soit-elle – à un meurtre collectif reste absolument insupportable ?

## 2.2 Corps sans gravité

Si les technologies automatiques contemporaines devaient interférer avec le régime éthique du sens, en troublant les modalités de sa réalisation concrète, ce pourrait être également en contaminant la dimension matérielle de l'existence subjective et du rapport à autrui. Pour la subjectivité levinassienne, le don – le geste même de l'éthique – s'apparente à un arrachement (Levinas, 1990b) : sacrifice pris dans l'épaisseur du moi pour être donné à l'autre. Aussi, c'est la dimension matérielle de la vie qui peut seulement être sacrifiée : celle d'une existence qui s'affirme dans un rapport à l'extériorité vécu comme besoin et consommation, jouissance dans le comblement même du manque, celle d'un corps qui doit fournir l'effort de soutenir son propre poids, comme enchaîné à lui-même par la gravité. À travers son œuvre (Levinas, 2002, 2004, 1990a, 1990b), Levinas dessine une articulation phénoménologique/éthique dans laquelle la jouissance et l'effort font le matériau du rapport entre l'être et l'autre. « [Le] donner n'a de sens que comme un arracher à soi [...] ; mais s'arracher à soi malgré soi n'a de sens que comme s'arracher à la complaisance en soi de la jouissance ; arracher le pain de sa

---

21 Bien sûr, le dispositif apprenant est loin d'être absolument neutre, notamment parce qu'il fonctionne à partir d'un ensemble de catégories qui devraient être choisies *a priori*. En ce sens, la responsabilité des concepteurs ne saurait être totalement déagée.

bouche » (Levinas, 1990b). À l'heure où les dispositifs automatiques s'insinuent entre le sujet et son monde, il faut noter, dans ses descriptions du rapport à l'extériorité comme besoin, travail, et don, la place déterminante que Levinas attribue à la main. À la fois, d'une part, elle est *tâtonnement* (Levinas, 1990a), elle s'engage, à l'aventure, dans les interstices de l'être, elle maîtrise l'élément et le ramène vers la bouche, elle est, de manière éminente, un organe de la jouissance. Mais, d'autre part, précisément parce qu'elle ne coïncide pas avec la bouche, elle maintient la chose maîtrisée à une distance essentielle de la consommation. Elle « ajourne la jouissance », elle maintient cet écart par lequel seulement la chose pourra entrer dans le langage et se faire le véhicule concret du contact éthique. Ainsi, la dimension matérielle de l'existence humaine relève d'une juste distance entre le sujet et le monde – celle du besoin. La main, à la fois, traverse cette distance pour combler le manque, mais tout autant elle la maintient pour pouvoir répondre à la demande d'autrui. Quand la médiation de dispositifs automatiques s'impose, entre le sujet et l'extériorité, dans la continuité de la main – voire à la place de la main – on peut s'interroger sur l'avenir de cette juste distance : n'y a-t-il pas un risque qu'elle devienne trop grande et que toute jouissance s'y perde ? Ou, au contraire, ne va-t-elle pas devenir trop faible, vers une coïncidence immédiate de la bouche et de l'aliment, rendant le don impossible ?

Si, souvent, on a associé le numérique avec l'idée d'immatériel, force est de constater aujourd'hui que – sans même parler de la matérialité propre du *hardware* – les dispositifs informatiques s'engagent – par une interface toujours plus riche de capteurs et d'actionneurs – dans une action efficace directement appliquée sur le monde physique. La pointe de l'outil répondant à une dynamique programmée, le rapport du sujet à l'extériorité relève alors d'une activité essentiellement cognitive : de contrôle, de vérification, de sélection, de validation. Nicholas Carr, dans son ouvrage *The Glass Cage : Automation and Us* (Carr, 2014), s'emploie à dénoncer cette couche d'automatisme – depuis les systèmes d'assistance à la navigation jusqu'aux robots industriels – qui se développe tout autour de l'humain et qui modifie – à outrance, selon l'auteur – la matérialité du rapport au monde. Les dispositifs opèrent une simplification, un lissage de la granularité essentielle du réel, et affectent ainsi sa profondeur *élémentale* – ils nous relèguent à une trop grande distance de cette « qualité pure » (Levinas, 1990a) qui fait la chair de l'existence subjective. Face à la planéité des écrans, l'aventure de la main et son expertise incarnée sont court-circuitées. Les dispositifs d'assistance réalisent une *externalisation du tâtonnement* et placent le sujet au poste de contrôle – ainsi isolé de la dimension matérielle de l'extériorité. Aussi – sauf à penser une

jouissance de la pure cognition – le sujet se trouve privé de la saveur de ses aliments, et se vide de cette épaisseur à partir de laquelle le don peut seulement avoir un sens.

Mais, à l'inverse, on peut concevoir que – lorsque les dispositifs tournent leurs capteurs et actionneurs vers le corps même du sujet et s'emploient à répondre à ses besoins de manière automatique – c'est la structure même du manque qui pourrait être troublée, et cette distance entre la main et la bouche – distance du don et de la parole – anéantie. Collectant des données au quotidien, l'intelligence artificielle saisit l'individu dans toutes les dimensions de son être – ses gestes, ses goûts, ses relations, son langage, son histoire, ses rythmes, ses projets – et apprend ainsi à anticiper ses mouvements, ses besoins, ses buts. Dans une connectivité généralisée – s'appuyant sur des protocoles que les géants industriels fixent – elle entoure le sujet : le programme domotique communique avec celui du véhicule autonome, avec celui de la *smart city*, avec les réseaux sociaux, etc. de sorte qu'une appréhension globale de l'individu permet de dégager, pour chaque situation, une intelligence contextualisée de ses objectifs et de ses attentes. Loin du paradigme de l'outil saisi, c'est ici le sujet qui est saisi dans un cocon automatisé, adaptatif, qui – dans une dynamique dialectique, en quelque sorte – comble les besoins qu'il crée et structure un système d'habitudes. Comme pris dans un exosquelette généralisé, qui soutient le poids avant même qu'il ne pèse, le sujet n'a plus à prendre en charge la gravité de son existence propre : plus de paresse, plus de fatigue, tout besoin est comblé avant de n'avoir pu signifier comme un manque. L'*ego* coïncide exactement avec la matière qui le nourrit : impossible alors d'arracher le pain de sa bouche puisqu'il est toujours déjà avalé.

Ainsi, on peut concevoir deux modalités par lesquelles la dimension matérielle de l'existence humaine, et sa réalisation dans le régime du besoin, pourraient être affectées : soit par excès de distance, soit, au contraire, par absence de distance, entre le sujet et les jouissances promises de l'extériorité. L'idée suivant laquelle l'humain pourrait, grâce aux dispositifs automatiques, se débarrasser du poids des contraintes matérielles est bien celle qui, dès 1960, dans le contexte de l'exploration spatiale, inspire Clynes et Kline lorsqu'ils imaginent qu'un corps *Cyborg* devra laisser l'humain « libre d'explorer, de créer, de penser et de ressentir » (Clynes & Kline, 1995, notre traduction). C'est elle aussi sans doute qui nourrit le fantasme trans-humaniste d'une conscience téléversable dans les systèmes informatiques, au profit d'une existence de pure pensée, libérée des contingences matérielles, et notamment de la mortalité. Mais notre but n'était sans doute pas ici de donner gage à ces projections aventureuses. Et il faudra bien d'abord imaginer que si les technologies automatiques

modifient le rapport à la matière tel que nous le connaissons, sans doute seront-elles à l'origine de nouvelles modalités de la jouissance et du don, de l'effort et de la possession, dans des formes hybrides d'abstraction et de contrétude<sup>22</sup>.

### 3. Conclusion

Dans la dynamique de co-évolution qui unit l'espèce humaine et le système technique, les dispositifs automatiques contemporains ouvrent sans aucun doute des voies nouvelles pour constituer le monde et rencontrer autrui. Cependant, nous avons aussi remarqué qu'ils pourraient entrer en conflit avec les modalités par lesquelles notre existence prend – selon Levinas – un sens proprement éthique.

Dans une version noire de ce scénario, populaire aujourd'hui dans l'imaginaire collectif, l'automatisme pourrait contaminer de manière définitive le régime du possible et la matérialité de l'existence de sorte que les rapports entre humains ne pourraient plus avoir le sens de la responsabilité ou du soin. Toute vie serait réduite à des processus automatiques dans lesquels le sens serait lui-même réduit à des calculs, des règles, des symboles, des traces et des données, des rythmes et des signaux. Mais nulle place dans ce monde pour le désir et la volupté : monde sans espoir dans lequel toute dimension humaine aurait été avalée par le grand algorithme. Le tout fonctionnerait comme une machine globale, sans extériorité : le temps serait plat, et l'espace serait toujours déjà parcouru. Cette vision du futur – pour le moins sombre – motif central de nombreuses œuvres de fictions, structure aujourd'hui, pour une large part, notre rapport à la partie technique de notre être.

Et sans aucun doute, sur le plan politique et social, des luttes sont à mener pour que les machines automatiques ne fassent pas le lit de nouvelles injustices. De ces luttes nous connaissons déjà plusieurs visages : d'une incitation à la destruction de la machine (comme les célèbres mouvements *luddites*), à l'encouragement frénétique de son accélération (*accélérationnisme*), en passant par le développement d'une nécessaire expertise *mécanologique* (Simondon, 2001 ; Stiegler, 1998) à même de guider le design des dispositifs pour une co-individuation harmonieuse.

Mais sur le plan anthropologique/existential, nous pensons que la vision mortifère d'une humanité asservie par l'automatisme est entretenue par une confusion. Celle-ci remonterait aux commencements de l'intelligence artificielle (son programme fort, ou *strong AI*), quand certains chercheurs ont cru – et laissé croire – qu'ils pourraient créer une conscience à partir

---

<sup>22</sup> Comme par exemple la technologie de l'impression a donné naissance à l'argent papier.

de dispositifs automatiques : alors l'humain s'est compris comme étant lui-même un dispositif automatique. Et on pourrait suivre, depuis les années 1950 (et même bien avant si l'on remonte, par exemple, à La Mettrie et son *Homme machine* de 1747 (La Mettrie, 1999)) jusqu'à nos jours, les études scientifiques, les œuvres de fiction, les systèmes philosophiques qui participent à entretenir la chimère d'une homogénéité absolue entre l'humain et la technique : les sciences cognitives computo-représentationnalistes, les discours trans-humanistes et post-humanistes, certaines neurosciences et autres projets de naturalisation de l'esprit, le design de robots anthropomorphes, les synthèses et commandes vocales, etc. C'est une longue et complexe histoire des idées qui façonne le regard que l'humain, face au miroir de la technique, porte sur lui-même.

Ainsi, s'il fallait – à notre niveau, qui est celui de la philosophie – imaginer un geste de résistance face à l'automatisation généralisée de l'être, pour que le visage d'autrui n'y finisse pas définitivement étouffé, alors celui-ci consisterait simplement à réaffirmer – et à célébrer même – le caractère vécu de l'expérience et du rapport à l'autre. Pour autant qu'elle mette en lumière l'irréductibilité du vécu à la logique des algorithmes et au fonctionnement des machines, la réduction phénoménologique est un acte de résistance – s'il en fallait un – face à la mécanisation de la pensée. Il y a un gouffre radicalement profond entre ce que l'on peut observer de la vie consciente et ce que l'on peut vivre (Levine, 1983). La phénoménologie, la philosophie de l'esprit, les sciences cognitives – les artistes, aussi – doivent s'emparer de l'incommensurabilité du vécu et produire de nouvelles représentations afin que l'humain ne s' imagine pas automatisable – et ainsi résiste à son automatisation. Les paradoxes de l'intersubjectivité, et les phénomènes limites tels que l'art, l'amour ou l'événement, doivent devenir les énigmes qui motivent la construction de paradigmes critiques. Il s'agit d'un programme scientifique, philosophique, artistique et finalement humaniste - une responsabilité.

## Bibliographie

- Bachimont, B., « L'intelligence artificielle comme écriture dynamique : de la raison graphique à la raison computationnelle », in Petitot J. & Fabbri P. (dir.), *Au nom du sens*, Paris, Grasset, 2000, pp 290-319.
- Bonnefon, J.-F., Shariff, A., Rahwan, I., « The social dilemma of autonomous vehicles », *Science*, 24 Jun 2016, vol. 352-6293, pp. 1573-1576
- Cardon, D., Casilli, A., *Qu'est-ce que le Digital Labor ?*, Bry-sur-Marne, INA, collection « Etudes et controverses », 2015, 104 p.
- Carr, N., *The Glass Cage : Automation and Us*, New York, WW Norton & Company, 2014
- Clynes, M. E. & Kline, N. S., « Cyborgs and Space », in *The Cyborg Handbook*. Gray C. H., Figueroa-Sarriera H. & Mentor S. (dir.), New York, Routledge, 1995, pp. 29-34
- La Mettrie, J. O., *L'homme-machine*, Paris, Folio, collection « Folio Essais», 1999
- Leroi-Gourhan, A., *Le Geste et la Parole, 1 : Technique et langage*, tome 1, Paris, Albin Michel, 1964a
- Leroi-Gourhan, A., *Le Geste et la Parole, 2 : La mémoire et les rythmes*, tome 2, Paris, Albin Michel, 1964b
- Levinas, E., *Autrement qu'être ou au-delà de l'essence*, Paris, Livre de Poche, 1990b
- Levinas, E., *De l'existence à l'existant*, Paris, Vrin, 2002.
- Levinas, E., *Le temps et l'autre*, Paris, Presses Universitaires de France, 2004
- Levinas, E., *Totalité et infini : Essai sur l'extériorité*, Paris, Livre de Poche, 1990a
- Levine, J., « Materialism and qualia: the explanatory gap », *Pacific Philosophical Quarterly*, 1983, 64: 354-361.

- Noothigattu, R., Gaikwad, S. « Neil » S., Awad, E., Dsouza, S., Rahwan, I., Ravikumar, P., & Procaccia, A. D. (2017). A Voting-Based System for Ethical Decision Making. *arXiv:1709.06692 [cs]*. Consulté à l'adresse <http://arxiv.org/abs/1709.06692> en octobre 2017.
- Petit, V., « Internet, milieu technique d'écriture », in Estrella R. (dir.), *Le social est-il soluble dans le Web ?*, Paris, Hermès, collection « Information, hypermédias et communication », 2013, 293 p.
- Simondon, G., *Du mode d'existence des objets techniques*, Paris, Aubier, 2001
- Stiegler, B., « Temps et individuations technique, psychique et collective dans l'œuvre de Simondon », *Intellectica*, 1998/1-2, 26-27, pp. 241-256