



La Bienveillance des machines :
comment le numérique nous
transforme à notre insu
de Pierre Cassou-Noguès

Roch Delannay

Publié le 17-11-2022

<http://sens-public.org/articles/1683>



Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA
4.0)

Résumé

La Bienveillance des machines n'est ni l'ouvrage d'un technophile, ni d'un technophobe. Pourtant, les machines, les robots, les algorithmes (d'intelligence artificielle) et les réseaux y occupent une place prépondérante. Il ne s'agit pas d'étudier les aspects politiques, économiques ou encore techniques qui s'y rapportent, même si ces dimensions sont légitimement représentées, mais de déplacer la question des rapports que l'humain entretient aux machines à un niveau ontologique. L'objectif de cet ouvrage est d'étudier ce que l'auteur nomme « syndromes technologiques » à travers une thèse originale : « les technologies contemporaines induisent différentes transformations de nos subjectivités. » Ces syndromes permettent d'aborder différentes notions allant de la confiance que l'utilisateur · trice accorde aux informations données par les machines jusqu'à celle de la surveillance dont nous faisons l'objet afin de démontrer, premièrement, le changement de nature de la machine et l'apparition d'un règne machinique (au même titre que le règne animal ou végétal) et, deuxièmement, la variation de notre forme de vie (au sens de Wittgenstein) pour mettre en évidence la transformation du sujet en sujet digital.

Abstract

La Bienveillance des machines is neither the work of a technophile nor of a technophobe. However, machines, robots, algorithms (of artificial intelligence) and networks occupy a preponderant place in it. It is not a question of studying the political, economic or technical aspects that are related to them, even if these dimensions are legitimately represented, but of shifting the question of the relationship that humans have with machines to an ontological level. The book studies what the author calls "technological syndromes" through an original thesis: "contemporary technologies induce different transformations of our subjectivities." These syndromes allow to approach different notions -- from the trust that the user grants to the information given by the machines to the one of the surveillance of which we are the object -- in order to demonstrate, firstly, the change of nature of the machine and the appearance of a machinic reign (in the same way as the animal or vegetal reign) and, secondly, the variation of our form of life (in the sense of Wittgenstein) to highlight the transformation of the subject into a digital subject.

La Bienveillance des machines : comment le numérique nous transforme à
notre insu

Mot-clés : Intelligence artificielle, Algorithmes, Ontologie, Technologie, Subjectivité

Keywords: Artificial intelligence, Algorithms, Ontology, Technology, Subjectivity

Table des matières

Introduction	5
Synthèse de l'ouvrage	6
Résumé	6
Les thèmes abordés	6
Commentaire	12
Bibliographie	12

La Bienveillance des machines : comment le numérique nous transforme à notre insu

Roch Delannay

Introduction

Pierre Cassou-Noguès, professeur au département de philosophie de l'Université Paris 8 Vincennes - Saint-Denis et co-éditeur de la revue SubStance, vient de signer un brillant ouvrage aux éditions du Seuil : *La Bienveillance des machines : comment le numérique nous transforme à notre insu* (2022). Ce n'est pas la première fois que l'auteur aborde la question des machines et des robots : avec la philosophie comme prisme initial, Pierre Cassou-Noguès dénoue les noeuds cybernétiques en convoquant différentes pratiques et disciplines : des sciences de l'information jusqu'aux mathématiques en passant par les arts et l'usage de la fiction. Nous pouvons citer quelques-unes de ses oeuvres majeures comme *Les Démons de Gödel* (2007), *Les rêves cybernétiques de Norbert Wiener* (2014), *Le Sujet Digital* (2022), *Techofictions* (2019), une brève nouvelle dans l'ouvrage *Le Comportement des Choses* (Quinz 2021) - nous y reviendrons ultérieurement -, et la co-réalisation du Web documentaire *Bienvenue à Erewhon* (« Bienvenue à Erewhon » 2019) avec Stéphane Degoutin et Gwenola Wagon.

En parallèle de ses réflexions sur les rapports entre sujets et machines, l'auteur développe une méthodologie basée sur la fiction expliquée dans l'introduction de *La Bienveillance des machines*. Il s'agit pour le philosophe de dépasser les méthodes employées par les autres domaines scientifiques, comme les sciences computationnelles ou les sciences économiques et sociales, et d'offrir littéralement une nouvelle perspective aux lecteur · trice · s. L'auteur rattache le « discours du philosophe [à] un autre domaine, la fiction, le possible » car « la question de la vérité empirique n'y a pas de sens » et que « le philo-

sophe ne sait rien : il invente » (cf. introduction du livre). L'objectif d'une telle méthode est de « redonner à la philosophie une portée spéculative, métaphysique si l'on veut, et en renforcer la portée critique en lui permettant d'envisager le réel à la lumière du possible¹ ».

Synthèse de l'ouvrage

Résumé

La Bienveillance des machines n'est l'ouvrage ni d'un technophile, ni d'un technophobe. Pourtant, les machines, les robots, les algorithmes (d'intelligence artificielle) et les réseaux y occupent une place prépondérante. Il ne s'agit pas d'étudier les aspects politiques, économiques ou encore techniques qui s'y rapportent, même si ces dimensions sont légitimement représentées, mais de déplacer la question des rapports que l'humain entretient aux machines à un niveau ontologique. L'objectif de cet ouvrage est d'étudier ce que l'auteur nomme « syndromes technologiques » à travers une thèse originale : « les technologies contemporaines induisent différentes transformations de nos subjectivités. » Le parcours de lecture en six chapitres alterne deux écritures : récit littéraire, fiction, et écriture universitaire. Exceptée l'introduction, chaque chapitre est précédé d'une petite histoire. Ces histoires racontent les syndromes technologiques et sont une porte d'entrée aux thèses détaillées dans chaque chapitre. Ces deux styles donnent à voir les syndromes sous deux prismes qui se complètent, laissant le soin au lectorat de s'approprier les informations qu'ils contiennent. Ces syndromes permettent d'aborder différentes notions allant de la confiance que l'utilisateur·trice accorde aux informations données par les machines jusqu'à celle de la surveillance dont nous faisons l'objet afin de démontrer, premièrement, le changement de nature de la machine et l'apparition d'un règne machinique (au même titre que le règne animal ou végétal) et, deuxièmement, la variation de notre forme de vie (au sens de Wittgenstein) pour mettre en évidence la transformation du sujet en sujet digital.

Les thèmes abordés

Un récit intitulé « Mon Oncle » permet d'introduire le syndrome du thermomètre lors du premier chapitre : « le syndrome du thermomètre consiste à

1. <https://pierrecassounogues.org/index.php/bio/>, consulté le 31 août 2022.

confier à une machine la tâche de décrypter ce qui devrait relever d'une expérience intérieure, ou en première personne ». L'auteur commence à décrire ce syndrome au moyen d'un appareil simple, un thermomètre, avant de créer un pont avec les neurosciences et les algorithmes d'intelligence artificielle, le principe du syndrome restant identique : l'individu se fie à une machine pour exprimer de manière non ambiguë ce qui, a priori, relève d'une expérience sensible et subjective vécue « à la première personne ». Apparaît alors une « indétermination de notre forme de vie » puisqu'une extériorité, une machine en l'occurrence, connaîtrait mieux l'individu que l'individu lui-même. Cette indétermination, cette « transformation de nos subjectivités », n'est pas anodine. Il ne s'agit pas simplement de confier aux machines le soin de lever une ambiguïté et de confirmer une sensation mais de leur confier le soin de penser à la place du sujet : ce n'est plus le sujet qui détient la vérité, c'est la machine qui la lui fournit. Or, Pierre Cassou-Noguès nous indique que d'un point de vue technique, les machines n'ont pas la capacité de réaliser une telle prouesse : aucune machine ne peut décoder un cerveau et toute la complexité des pensées qui le traversent. Ce phénomène est rendu possible grâce à « un contexte beaucoup plus large [fait de] croyances implicites, [d']une idéologie » et de l'acceptation sociale de la technique et de l'argument scientifique. Ainsi, selon le degré d'autorité accordé ou d'assujettissement, la machine, ou l'entité propriétaire de cette machine, détient une vérité de la pensée du sujet et opère un nivellement de la pensée que l'auteur définit comme un « esprit plat ». Le chapitre suivant sera quant à lui dédié à la diffusion du syndrome du thermomètre au moyen des algorithmes d'intelligence artificielle.

Une nouvelle fiction est insérée avant le deuxième chapitre : celle-ci nous conte une comparaison entre chiromancie et neurosciences (neuromancie). Le parallèle entre les deux pratiques se construit sur la croyance qu'une vérité fondamentale est ancrée, « imprimée » dans notre corps et que seuls des experts, pour ne pas dire des prophètes, sont en mesure de déchiffrer. C'est ce que nous offre la machine (et son propriétaire) à travers le syndrome du thermomètre et la croyance que nous lui accordons : une lecture de nos pensées que nous ne serions pas capables d'exprimer. Toutefois, Pierre Cassou-Noguès nous rappelle que les machines n'ont pas la capacité de penser par elles-mêmes : les machines et les algorithmes ne pensent pas, ils appliquent des calculs à des données, pour les transformer et en faire émerger de nouvelles données, des informations. Le cas de Cambridge Analytica (2016) montre l'utilisation d'un algorithme qui dresse le profil psychologique des utilisateurs selon le modèle

OCEAN basé sur six paramètres, modèle qui n'est pas sans rappeler celui du SONCAS enseigné dans les filières commerciales². Cet exemple révèle un aplatissement et une réduction de l'individu à un paramètre, un archétype, en lieu et place de toute la complexité qui le constitue, et en fait un « esprit plat » sans profondeur. La présentation de cet exemple introduit la diffusion du syndrome du thermomètre à travers Internet et de l'instrument référant pour la réception et la diffusion d'informations sur les réseaux : le téléphone intelligent. L'auteur aborde ici la question de l'assujettissement de l'utilisateur · trice à son téléphone portable, aux multiples applications installées et aux propriétaires (fabricants) de ces machines. Cette diffusion à travers les réseaux et les instruments mobiles – un complexe créé à partir du *hardware* et du *software* – déploie un nouveau système de surveillance afin que les machines nous bienveillent, afin qu'elles puissent être à l'écoute et répondre à nos syndromes technologiques en « déterminant nos contenus de pensée ». Ce faisant, l'auteur avance une nouvelle hypothèse d'un dépassement du panoptique de Bentham et Foucault vers ce qu'il nommera synhapticon dans le dernier chapitre.

Après un bref rappel historique de l'évolution de l'informatique et du terme d'intelligence artificielle des années 1950 à nos jours, débute le troisième chapitre sur la désynchronisation que provoque la diffusion du syndrome du thermomètre à travers les téléphones « intelligents ». La désynchronisation dont il est question est double : premièrement, elle est relative à la temporalité de l'individu (individuelle) et, deuxièmement, elle concerne la désynchronisation des consciences (sociales). Les deux sont issues du changement de medium de masse, du passage de la télévision à Internet. Face au modèle économique qui émerge avec l'infosphère (économie de l'attention), la temporalité de l'individu éclate devant la multitude de flux d'informations et se calque sur un mouvement extérieur, celui, infatigable, du numérique et d'Internet. Du point de vue social, une autre désynchronisation apparaît, celle des consciences. Là où la télévision permettait une synchronisation des consciences (en suivant

2. Le principe de ces deux modèles est de décrire une représentation simplifiée de l'individu basée sur des archétypes. Chaque individu porte en lui l'ensemble des archétypes mais un (ou deux) peuvent être plus forts que les autres. L'acronyme SONCAS correspond aux termes : sécurité, orgueil, nouveauté, confort, argent et sympathie et permet de définir la relation entre acheteur et vendeur, ce dernier adaptant son discours en fonction de l'archétype présumé majeur chez le client. De la même manière, l'auteur décrit le modèle OCEAN comme un outil pour dresser le profil des utilisateurs et ainsi adopter un comportement précis en fonction dudit profil.

un programme déterminé par les diffuseurs), Internet, par la disponibilité des contenus qu'il propose, désynchronise ce flux de conscience : les individus ne sont plus collés à un unique flux et à un rythme imposé par la chaîne mais au rythme qui leur convient. Un des effets de la désynchronisation qui aurait pu être observé est celui d'un éclatement de la sphère sociale, par dispersion des valeurs et des idées. Or, une convergence sociale est toujours présente, malgré la démultiplication des flux et des contenus. Pour expliquer ce phénomène, l'auteur introduit la notion de cliché tel que définie par Wiener dans un système d'informations :

Norbert Wiener appelle *cliché* un message, qui, dans un certain contexte, est plus probable qu'un autre, un message donc qui se répète avec quelques variantes et ne contient plus beaucoup d'informations. [...] Les clichés engagent une forme de pensée commune, qui a bien la position de ce que Tarde appelle « l'opinion » relativement à la presse, mais ce que l'on peut difficilement encore associer à ce terme. Les messages qui circulent, cette sorte de pensée commune donc, n'ont plus la forme d'énoncés linguistiques mais s'expriment dans des images et des formules qu'il est impossible de dissocier. *Plutôt que d'opinion, il faudrait parler de rêve commun.* (Cassou-Noguès 2022)

Les clichés deviennent alors « l'obsession du réseau » et sont largement diffusés par les plateformes du fait de leur popularité. Il s'étendent et se propagent en touchant de nouveaux utilisateurs · trices. L'auteur conclut ce chapitre sur la « dictature du cliché » et sur sa potentielle capacité à représenter le vrai.

Pierre Cassou-Noguès ajoute à cette « dictature du cliché », lors d'un nouveau chapitre (4), la notion de « travail zombie », notion qui se démarque des travaux de Casilli sur le *digital labor*³. L'objectif de cette partie est de « comprendre la production de valeur [de l'internaute] devant son écran »,

3. Le *digital labor* tel que défini par Antonio Casilli et Dominique Cardon rassemble deux formes de production de valeur dans les espaces numériques. La première est relative au travail des usagers sur les différentes plateformes sur lesquelles ils naviguent. Chaque clic, chaque usage d'une fonctionnalité particulière peut faire l'objet d'une tâche et d'une valeur ajoutée. Le deuxième objet concerne ce que Shoshana Zuboff (2019) nomme *Le Capitalisme de surveillance*, donc la production de valeur ajoutée à partir des traces numériques que laissent les usagers sur les plateformes.

au même titre que celle du téléspectateur devant sa télévision. L'auteur commence donc par revenir sur le médium qu'est la télévision et développe la production de « temps de cerveau disponible » du téléspectateur et le travail zombie que celui-ci réalise pour rendre son cerveau disponible, valeur qui sera ensuite revendue comme produit par les chaînes de télévision. Ensuite, c'est au tour du médium numérique de faire l'objet d'une analyse ; l'auteur donne à voir les évolutions du travail zombie depuis la télévision jusqu'aux multiples flux et canaux (plateformes) accessibles grâce à Internet. Dans les deux cas de figure, l'utilisateur travaille à se transformer en produit, à créer, à travers la production de temps de cerveau disponible, une valeur ajoutée à un produit : « la désirabilité ». En conclusion de ce chapitre, l'auteur projette l'idée d'une société sans travail en s'appuyant sur les travaux de Marx, et y entrevoit un décalage, où n'est pas anticipée l'aliénation du loisir (rendu disponible par l'automatisation) dans un rapport à la machine et qu'il « puisse être réinvesti dans la production d'une marchandise : du temps de cerveau disponible. Ce qui, à la fois, empêche l'émancipation intellectuelle du travailleur zombie et reproduit la logique de l'exploitation au sein même du loisir : le travailleur zombie produit de la valeur, en rendant ces marchandises désirables, et cette valeur ajoutée lui est extorquée ».

Un récit de quelques pages, précédant le cinquième chapitre, traite des habitudes que nous avons, nous êtres humains : un ensemble de règles plus ou moins complexes que nous suivons pour construire nos quotidiens. Ces habitudes sont la porte d'entrée aux algorithmes de prédiction que l'auteur regroupe sous l'appellation de robots bienveillants. « La bienveillance ou la fable du corbeau et du robot », titre de ce nouveau chapitre, entend donner un visage à cette machine qui nous bienveille : celle du robot compagnon. L'intérêt de l'auteur se porte sur Pepper et Nao qui « constituent pour [lui] les meilleurs représentants actuels de l'idée de robot compagnon ». Ce sont des robots « humanoïdes et bienveillants », c'est-à-dire qu'ils ne travaillent pas, ne sont pas assignés à une tâche particulière hormis celle de veiller à notre bien-être : ils nous « bien-veillent ». Leur apparence dépasse la vallée de l'étrange de Mori de manière à ce que l'imaginaire du robot domestique soit accepté dans nos quotidiens et nos habitudes, plutôt que l'imaginaire d'un robot inquiétant véhiculée depuis un certain temps. Ces robots provoquent un changement dans le rapport que nous entretenons avec la machine, au-delà du traitement accordé à de simples outils, des bouts de plastiques et des composants électroniques ; ce sont des affects et de nouvelles conventions

morales qui sont mises en jeu. Le pendant de cette bienveillance est celui de la surveillance : pour veiller à notre bien-être, les robots compagnons doivent surveiller leurs propriétaires, ce qui techniquement se traduit par la captation et le traitement de données (personnelles).

Le sixième chapitre, « la sensibilité synhaptique », traite de nos sensibilités et de leur « fonctionnement » afin de mettre en évidence la caractéristique particulière du numérique : la non-réciprocité (exemple de non-réciprocité et de réciprocité : je peux voir sans être vu, imaginaire de l'homme-invisible, mais je ne peux pas toucher sans être touché). L'ensemble des modifications dont il est question dans cette partie dirige le lecteur · trice vers le dépassement du panoptique de Foucault, où, à l'ère des données, le *panopticon* devient un *synhapticon* gardé et surveillé par une machine. La machine morcelle le sujet et le corps, elle les découpe en données fragmentées à travers une surveillance discontinue dont elle assemble les morceaux pour recomposer le sujet. La figure d'évasion du système de surveillance se transforme : de l'homme invisible elle passe à celle du *hacker*, non plus celui qui échappe au regard du gardien mais à celui qui circule sans laisser de traces numériques. Toutefois, dans ce passage du panoptique au synhaptique, se pose la question du sens accordé au mot « sujet », car « l'instance qui rend le sujet perceptible, l'instance qui l'interpelle, n'est plus un humain (gardien de prison, maître d'école ou psychiatre), mais une machine, un algorithme ».

La conclusion de cet ouvrage débute par le récit des Vélib' et de leurs habitudes en région parisienne. Cette fiction, rédigée en 2020 pour un autre ouvrage, raconte le « comportement des choses » (Quinz 2021). Elle sort du prisme anthropocentré et octroie aux objets, et plus spécifiquement aux machines et aux robots, la capacité d'adopter (comme si elles pensaient) des comportements particuliers. Cette fiction ouvre une nouvelle perspective, celle de l'avènement du règne machinique, où les machines se hisseraient à l'égal des animaux, des végétaux et des champignons. Toutefois, les machines, malgré leur grande capacité de calcul et d'apprentissage, ne peuvent pas vivre indépendamment des êtres humains : elles les surveillent, pour leur bien-être, et ne peuvent pas améliorer leurs services (et les travaux secondaires qu'elles peuvent effectuer) sans eux, sans les surveiller, car c'est par la captation des données (et des informations) qu'elles enrichissent leurs bibliothèques de connaissances, les serveurs. Inversement, les êtres humains, de par leurs syndromes technologiques, transitent entre deux formes de vies, et passent de

l'état de sujet à celui de sujet digital parce qu'ils partagent leur quotidien avec les machines.

Commentaire

La *Bienveillance des machines* est un ouvrage saisissant dont les mots offrent aux lectrices et lecteurs une nouvelle vision, originale, sur les problématiques liées à l'espace numérique, aux algorithmes d'intelligence artificielle et au moissonnage des données (personnelles). Ainsi la machine n'est plus qu'un simple outil de production ou un moyen de générer des richesses pécuniaires, elle intègre également une partie (externalisée) de notre système de pensée. Cependant, il semblerait que la question de la responsabilité, intimement liée à celles de l'autorité et de la croyance évoquées à plusieurs reprises, ne soit pas directement traitée dans ce livre. En effet, à quel niveau la responsabilité du sujet numérique se situe-t-elle dans cette relation avec une machine bienveillante? Il ne s'agit plus, dans l'espace numérique, de croire en une seule institution mais d'être confronté à une multitude d'entités qui préservent leurs algorithmes bien à l'abri des yeux indiscrets, derrière des licences propriétaires et des droits d'utilisation. Quel degré d'autorité le sujet digital peut-il ou doit-il accorder à ces entités? Nous pourrions tourner la question autrement : Est-ce qu'il y a déresponsabilisation du sujet digital lorsque celui ou celle-ci croit en la machine qui le/la bienveille, et donc en l'entité propriétaire de cette machine et des pensées qu'elle génère? Est-ce toujours le cas si le sujet digital à affaire à une machine non propriétaire (sous licence libre par exemple)?

Bibliographie

« Bienvenue à Erewhon ». 2019. <http://welcometoerewhon.com/>.

« Bio ». s. d. Site web personnel. *Pierre Cassou-Noguès*. Consulté le 31 août 2022. <https://pierrecassounogues.org/index.php/bio/>.

Cassou-Noguès, Pierre. 2007. *Les démons de Gödel : logique et folie*. Science ouverte. Paris : Seuil.

———. 2014. *Les Rêves cybernétiques de Norbert Wiener*. Seuil.

———. 2019. *Technofictions*. Paris : Les éditions du Cerf.

La Bienveillance des machines : comment le numérique nous transforme à
notre insu

———. 2022. *La Bienveillance des machines. Comment le numérique nous transforme à notre insu*. Paris : Seuil.

Cassou-Noguès, Pierre, Claire Larsonneur, Arnauld Regnauld, et Sara Touiza, éd. 2022. *Le Sujet Digital*. Traduit par Stéphane Vanderhaeghe, Géraldine Bertres, Hélène Soldano, et Marie Springinsfeld. La Grande Collection ArTEC. Nanterre : Presses universitaires de Paris Nanterre. <http://books.openedition.org/pupo/25517>.

Quinz, Emanuele. 2021. *Le Comportement des Choses*. Dijon : Presses du réel.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism*. New York : PublicAffairs.