



HAL
open science

Recension de Cassou-Noguès (Pierre), La Bienveillance des machines. Comment le numérique nous transforme à notre insu, Paris, Éditions du Seuil, 2022

Thierry Ménissier

► **To cite this version:**

Thierry Ménissier. Recension de Cassou-Noguès (Pierre), La Bienveillance des machines. Comment le numérique nous transforme à notre insu, Paris, Éditions du Seuil, 2022. Implications philosophiques, 2023, pp.329. halshs-04108411

HAL Id: halshs-04108411

<https://shs.hal.science/halshs-04108411>

Submitted on 27 May 2023

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.

Recension de Cassou-Noguès (Pierre), *La Bienveillance des machines. Comment le numérique nous transforme à notre insu*, Paris, Editions du Seuil, 2022, 329 p.

Thierry Ménissier, IphiG, Université Grenoble Alpes

Paru dans *Implications Philosophiques*, Avril 2023

C'est un ouvrage philosophique original que propose Pierre Cassou-Noguès avec *La Bienveillance des machines*. Certes, le domaine qu'il aborde n'est pas quelque chose de révolutionnaire : la philosophie des êtres artificiels, artefacts informatiques, robotiques et numériques, non nécessairement anthropomorphes, connaît déjà quelques solides références¹. Mais le ton adopté par l'auteur aussi bien que sa méthode d'investigation ouvrent de suggestives perspectives. D'une part, il s'agit d'un ouvrage philosophique où le « je » de la première personne du singulier joue un rôle important pour signifier l'engagement de l'auteur dans son propos, pour unifier ses observations et pour coordonner ses expériences de pensée ; de l'autre, la fiction, sous la forme des six courts récits qui accompagnent chaque chapitre, constitue un mode d'intuition sensible pour l'analyse philosophique, elle confère un contenu sensible ou « remplit » des concepts qui, sans cela, demeureraient vides. Les narrations fictionnelles, estime plus précisément l'auteur, permettent en effet de cerner les imaginaires qui sous-tendent leurs usages, et il est à ce propos possible de mobiliser également le concours de nombreuses œuvres littéraires. Ce procédé, l'auteur l'avait déjà expérimenté dans plusieurs ouvrages : *Une histoire de machines, de vampires et de fous*, 2007 ; *Technofictions*, 2019 ; *Virusland*, 2020. L'art de la fiction met donc en scène des formes de vie originales et, de ce fait, stimule l'étonnement philosophique afin d'explorer des expériences nouvelles.

L'ouvrage propose de considérer l'idée que les technologies numériques, souvent rendues invisibles, puissent être conçues pour rendre les humains heureux. Par suite, il convient d'admettre la possibilité que le numérique « nous transforme à notre insu ». Ce sous-titre légèrement anxigène formule moins une prévention dystopique qu'il n'en appelle à considérer avec « l'inquiétante étrangeté » qu'elles méritent (et pour reprendre les mots avec lesquels on traduit l'essai consacré par Freud à ce qu'il nommait l'*Unheimliche*²) les interactions humaines avec des artefacts contemporains. L'auteur n'évoque même pas la servitude librement consentie à l'ère du « capitalisme de surveillance » critiqué par Shoshana Zuboff³, mais soutient la thèse que les machines sont conçues dans le dessein de veiller sur les humains de manière bienfaisante : selon le néologisme proposé, elles nous « bienveillent » (p. 17, 87, etc.). La contrepartie de cette sollicitude *by design* réside dans le fait que, en fonction de leur pouvoir croissant d'automatisation (ou de leur tendance à l'autonomisation), leur influence tend à s'étendre sans limitation claire ni définissable *a priori*. On peut dès lors émettre le soupçon

¹ Voir par exemple ces tentatives en matière d'éthique algorithmique, robotique et artificielle : Colin Allen, Iva Smit et Wentell Wallach, "Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches", *Ethics and Information Technology*, 2005, 7(3), p. 149-155 ; Mark Coeckelbergh, "Can we trust robots?", *Ethics and Information Technology*, 2012, 14, p. 53-60 ; David J. Gunkel, *Robot Rights*. Boston, The MIT Press, 2018 ; Martin Gibert, *Faire la morale aux robots. Une introduction à l'éthique des algorithmes*. Paris, Flammarion, 2021.

² Sigmund Freud, *L'inquiétante étrangeté* (1919), in *Essais de psychanalyse appliquée* trad. par M Bonaparte et E Marty, Paris, Gallimard, 1971. Cette même catégorie est employée par Tyler Reigeluth pour qualifier certains effets du déploiement de l'intelligence artificielle, voir Tyler Reigeluth, « Le rapport magique à l'Intelligence Artificielle, ou comment vivre avec l'aliénation technique », *Quaderni. Communication, Technologies, Pouvoir*, n°105, 2022, p. 35-52.

³ Shoshana Zuboff, *L'Âge du capitalisme de surveillance : le combat pour un avenir humain face aux nouvelles frontières du pouvoir* (2019), trad. B. Fomentelli et A.-S. Homassel, Paris & Veules-les-Roses, Éd. Zulma, 2022.

qu'elles manifestent un « intérêt » à étendre sans fin leur pouvoir d'action, et même qu'elles utilisent les humains pour y parvenir, l'auteur reprenant à son compte l'intuition autrefois exprimée par l'écrivain Samuel Butler⁴.

Comment un tel phénomène est-il possible ? Le premier chapitre décrit un « syndrome du thermomètre », évocateur selon l'auteur de la tendance actuelle de nombreuses personnes à vouloir être renseignées sur leur état intérieur par des artefacts estimés capables de le faire, ce qui implique de prêter par principe à de tels artefacts le pouvoir de savoir ce qu'il en est mieux que les personnes affectées elles-mêmes. Ainsi autorisées par leurs usagers, les machines « s'accaparent notre expérience même, nos états d'âme » (p. 35), Et parce que nous leur permettons de prendre une mesure objective de notre état, nous avons tendance à considérer que ces « dolorimètres » manifestent la capacité de nous débarrasser de la partie désagréable de notre expérience. Ce qui vaut pour les états de corps apparaît également pertinent pour les états de pensée, l'auteur faisant pour l'établir référence à plusieurs fictions littéraires de diverses époques⁵. Cette tendance est accrue par le fait que notre vie psychique ainsi que l'état de notre santé mentale se voient confiés au pouvoir des machines (chapitre 2 : « L'esprit plat »). De la sorte, les questions qui se posent au philosophe sont de savoir s'il y a de ce fait un contrôle des humains par les machines, et si ces derniers peuvent encore être considérés comme des sujets à part entière. L'auteur n'apporte pas de réponse catégorique à ces questions, mais les instruit en discutant les thèses de Michel Foucault sur le biopouvoir⁶ : cette orientation donne à penser qu'à défaut de contrôle, les machines exercent une surveillance constante des humains. En contrepartie, il s'interroge sur la capacité des machines à comprendre ce qu'est la pensée humaine, en convoquant de manière ingénieuse des fictions relatives à des « machines à lire les rêves » ou encore certaines expérimentations artistiques attentives à la possibilité émergente que certaines machines puissent participer à l'activité onirique (p. 92-102).

Conviant alors à la discussion Gabriel Tarde, de nouveau Sigmund Freud et les thématiques de l'économie de l'attention développées par Yves Citton⁷, l'auteur aborde dans le chapitre 3 les questions relatives à la synchronicité et à la désynchronisation entre le temps vécu par les humains et le rythme imposé par les machines auxquelles ceux-ci se connectent. L'économie psychique des « masses complexes » recouvre de considérables enjeux en termes de rapport à la qualité des informations, compte tenu du caractère intellectuellement pauvre des contenus des clichés pourtant véhiculés de manière virale par les réseaux sociaux et les plateformes ludiques. Cette notion de cliché est empruntée par l'auteur à Norbert Wiener⁸ et désigne l'appauvrissement des contenus d'informations sous l'effet de leur diffusion massive. Leur diffusion s'explique selon l'auteur par le fait que, comme avec ce qui se passe pour les mots d'esprit dans la vie psychique selon Freud, « *la stabilité des clichés qui circulent dans le réseau est celle des rêves récurrents* » (p. 146, souligné par l'auteur). Parallèlement au risque encouru d'une désinformation massive par les *fake news*, l'installation d'un processus

⁴ Samuel Butler, *Erewhon ou De l'autre côté des montagnes* (1872), trad. V. Larbaud, Paris, Gallimard, 1981.

⁵ Sont notamment invités par l'auteur : Edgar Allan Poe (« Double assassinat dans la rue Morgue », 1841), Auguste Villiers de l'Isle-Adam (*L'Eve Future*, 1886), André Maurois (*La Machine à lire les pensées*, 1937), ou encore Jean-Philippe Toussaint (*La Télévision*, 1997).

⁶ Michel Foucault, *Naissance de la biopolitique*. Cours au Collège de France, 1978-1979, Paris, Gallimard & Editions du Seuil, 2004.

⁷ Voir Gabriel Tarde, *Les Lois de l'imitation* (), Paris, Les empêcheurs de penser en rond, 2001 ; Sigmund Freud, « Psychologie des foules et analyse du moi » in *Essais de psychanalyse*, trad. Y. Bourguignon, Paris, Payot, 1981 ; Yves Citton, *Pour une écologie de l'attention*, Paris, Editions du Seuil, 2014 ; *Médiarchie*, Paris, Editions du Seuil, 2017.

⁸ Wiener, Norbert, *Cybernétique et société. L'usage humain des êtres humains* (1952), préface R. Le Roux, trad. P.-Y. Mistoulon, Paris, Editions du Seuil, 2014.

d'insignifiance structurelle apparaît préoccupante, puisque, ainsi appréhendés, « les clichés constituent les obsessions du réseau » (p. 152).

Le chapitre 4 (« Le travail zombie ») aborde plusieurs aspects de la transformation de la catégorie de travail sous l'effet des usages numériques. Rouvrant aujourd'hui le dossier créé par les mots du PDG de la première chaîne TV française, Patrick Le Lay en 2004 sur le « temps de cerveau disponible » pour les annonceurs publicitaires, il reprend plusieurs hypothèses formulées, dès les années 2000, par les précurseurs de la philosophie sociale et politique de l'essor d'internet, telles que celles du capitalisme cognitif⁹, du « travail immatériel »¹⁰, ou encore du « travail bio-matériel »¹¹. Enrichie par les analyses sur le « *digital labor* »¹², l'auteur construit la figure du zombie issue de la culture populaire en catégorie philosophique, susceptible de donner à comprendre le genre de transformations qui affectent aussi bien l'attention des usagers des médias que la production de valeur marchande *via* les transformations de l'économie numérique. Cette figure, Cassou-Noguès l'affectionne ainsi qu'en témoigne son ouvrage intitulé *Mon Zombie et moi. La philosophie comme fiction* (2010). Dans ce nouvel ouvrage, devient zombie « un humain transformé dont est supprimée l'individualité, et qui n'est plus que le jouer d'[un] désir étranger » (p. 179).

Ainsi, selon l'auteur, l'utilisateur des médias travaille à son insu et contribue à véhiculer des clichés, à tel point qu'« une société pleinement numérique ne connaîtrait plus que des clichés » (p. 187). Des développements approfondis s'attachent à déterminer si ceux-ci se situent au-delà ou en dehors de toute forme d'idéologie, ou s'ils en constituent une nouvelle forme, une forme originale d'idéologie vide où le marché de l'attention ne renvoie qu'à lui-même (p. 187-195). Parce qu'il s'interroge sur les paradoxes de l'économie contributive, l'auteur, en s'appuyant sur la réflexion de Christian Fuchs¹³, se demande s'il ne s'agit pas d'une pure et simple, quoique nouvelle, exploitation de la force de travail par l'espionnage des comportements numériques et la captation de l'attention des usagers, qui, de fait, fournissent un travail non rémunéré capable d'accroître sensiblement la valeur des produits et des services, et par-là d'augmenter les inégalités sociales et économiques. Il conclut que, dès sa conception, dans le cadre de l'économie numérique, « le travail zombie accroît la désirabilité de l'objet » (p. 201). L'auteur n'affine pas ses analyses comme il aurait pu le faire avec des développements adossés à ce qui a été identifié comme le capitalisme affectif, où les émotions deviennent la matière même de la production donnant lieu à des échanges marchands¹⁴. Ce chapitre s'achève sur un paradoxe sans surprise : certes, la société algorithmique et numérique tend à faire disparaître le travail salarié, pour autant elle n'en finit pas avec l'exploitation capitaliste qu'on connaissait dans le monde industriel, mais la reconduit autrement.

Le chapitre 5 évoque la tentation contemporaine d'entourer l'existence humaine des soins que pourraient lui apporter une assistance personnalisée sous la forme des artefacts domestiques. La robotique sociale, envisagée sous l'angle du besoin d'assistance, offre l'occasion de dépasser le préjugé selon lequel l'utilité fonctionnelle des artefacts dans la sphère

⁹ Yann Moulier-Boutang, *Le capitalisme cognitif. La Nouvelle Grande Transformation*, Paris, Editions Amsterdam, 2007.

¹⁰ Antonio Negri et Michaël Hardt, *Empire*, trad. D.-A. Canal, Paris, Exils, 2000.

¹¹ Eugene Thacker, *Biomedica*, Chicago, University of Minnesota Press, 2004.

¹² Antonio Casilli, *En attendant les robots. Enquête sur le travail du clic*, Paris, Editions du Seuil, 2019.

¹³ Fuchs, Christian, *Digital Labour and Karl Marx*, Londres, Routledge, 2014.

¹⁴ Illouz, Eva, *Les Sentiments du capitalisme*, Paris, Editions du Seuil, 2006 ; (dir.), *Les marchandises émotionnelles. L'authenticité au temps du capitalisme*, Paris, Premier Parallèle, 2019.

domestique représente leur unique critère d'adoption : en réalité, leur plus-value réside dans le fait qu'ils contribuent à la relation sociale de manière plus large, au-delà des fonctions spécifiques pour lesquelles on les a conçus. L'analyse de la manière dont les robots Nao et Pepper sont présentés par leurs fabricants, notamment, indique que « le robot compagnon est un être nouveau, qui n'existait pas auparavant et dont la fonction est, si l'on peut dire, de « bienveiller » l'humain : veiller à ce que l'humain se sente bien. » (p. 225, souligné par l'auteur). Tandis qu'on a souvent tendance à estimer que les robots sont acceptés parce qu'ils sont étroitement assignés à des fonctions domestiques comme le sont les appareils électroménagers, il s'agit là d'un véritable principe de constitution de la robotique sociale.

Toutefois, même en acceptant ce principe, la présence de tels artefacts ne va pas sans poser de très sérieux problèmes aux humains, en fonction de leurs détournements toujours possibles. Ainsi, discutant une controverse récente, l'auteur se demande s'il est réellement scandaleux du point de vue moral de détourner dans un sens sexuel un robot compagnon initialement non conçu pour cette activité¹⁵. Le trouble qu'il est possible de ressentir concerne également le cœur de leur activité : en effet, le statut aujourd'hui conféré à ces machines, déployées dans le sens d'un accompagnement de soins pouvant aller jusqu'à la thérapie psychique, n'est pas clair : l'artefact est à la fois objet transitionnel et une « sorte de vampire [...] sous des dehors attendrissant » aspirant irrésistiblement les données privées de leurs propriétaires (p. 245). Si l'on peut adresser une remarque critique à cette intéressante analyse de la robotique sociale, il aurait été intéressant (puisque les robots sont de véritables espions enregistrant et transmettant les informations privées *via* le numérique) de l'approfondir dans une perspective politique à partir des thèses de Shoshana Zuboff relatives au « capitalisme de surveillance »¹⁶, tout en élargissant son horizon au dialogue entre les humains et les non-humains (naturels et artéfactuels) tel que l'envisage Bruno Latour¹⁷ (puisque'ils font désormais partie de l'environnement humain, ils en sont des agents comme le sont également les animaux).

Enfin, le sixième chapitre, examinant plusieurs dispositifs expérimentaux ou déjà accessibles au grand public, explore audacieusement la transformation de la sensibilité : le sens du toucher, initialement évoqué pour la philosophie moderne à travers les expériences mentales autrefois proposées par Descartes¹⁸, apparaît aujourd'hui passible d'être modifié dans ses propriétés mêmes par la circulation des informations numériques. Ce qui est bouleversé avec les nouveaux dispositifs haptiques (i.e. qui concernent le sens du toucher) est la possibilité de toucher à distance sans que la chose ou la personne que l'on touche perçoivent ce contact, et de nouvelles combinatoires apparaissent entre les sens du toucher, de la vue et de l'ouïe (ainsi dans plusieurs dispositifs évoqués p. 270-280, rendant possibles des embrassades ou une relation sexuelle à distance et favorisant une étrange expérience qualifiée de téléhaptique). De même, l'expérience des écrans démultipliés modifie la perception au point de permettre d'ériger le

¹⁵ L'auteur rappelle qu'en 2015, le hacker japonais Toshiaki Miyachi a créé la controverse en partageant une vidéo équivoque, ce qui a conduit l'entreprise Softbank, qui louait ces robots à des particuliers, à ajouter une clause au contrat que signent ses clients de non-utilisation dans des activités sexuelles ou indécentes (p. 237-238).

¹⁶ Shoshana Zuboff, *L'Âge du capitalisme de surveillance*, *op. cit.*

¹⁷ Bruno Latour, *Nous n'avons jamais été modernes. Essai d'anthropologie symétrique*, Paris, Editions La Découverte, 1991.

¹⁸ René Descartes, *La Dioptrique* (1637), Premier discours, in *Œuvres*, Paris, Librairie philosophique J. Vrin, tome VI, p. 83-84.

régime synhaptique en contrepoint du panoptique (p. 295-306)¹⁹. Au-delà des problématiques de surveillance et de contrôle, ce qui se joue dans la relation avec les artefacts est l'émergence de nouvelles formes de sensorialité ainsi qu'une modification dans les manières qu'ont les humains d'être au monde, notamment à travers la mise en question, déjà évoquée plus haut, de la notion de sujet. L'auteur explique, analyse des dispositifs technologique à l'appui, que la « réflexivité du toucher » évoquée notamment par Merleau-Ponty et Derrida²⁰ est peut-être en train de prendre un sens nouveau grâce aux technologies haptiques. Il est dommage à ce propos que l'auteur n'envisage pas explicitement les nouvelles métronomies du moi permises par le « *Quantified self* »²¹ : par exemple, que fait apparaître le processus de subjectivation qu'on pourrait nommer techno-ascétique, par référence aux pratiques aujourd'hui massivement développées dans les salles de sport à travers les applications algorithmiques et numériques ?

Foisonnant et agréable à lire, l'ouvrage de Pierre Cassou-Noguès est riche en cas observés et discutés, curieux d'aspects variés des innombrables pratiques sociales assistées par l'algorithmique, le numérique et la robotique, original dans les angles de vue qu'il propose, enfin souvent éclairant du point de vue heuristique : il contribue à la découverte du monde qui nous entoure en stimulant l'étonnement du lecteur. Sans jamais céder ni au techno-optimisme ni à l'esprit de dystopie, il cerne plusieurs menaces liées à la condition technologique typique des sociétés où l'innovation joue un rôle majeur, et explore les possibles ouverts par cette condition, de manière à la fois informée, critique et inventive. On pourrait toutefois identifier deux manques.

D'une part, on regrette que l'ouvrage n'explore pas la dimension écosystémique de la condition technologique contemporaine : issues du monde industriel d'hier, portées par un imaginaire de l'illimitation des ressources qui promet non seulement de l'assistance mais souvent de l'augmentation machinique, les relations actuelles entre les humains et les artefacts gagnent également à être comprises en relation avec tous les vivants et en regard des nouvelles données environnementales et en fonction d'un mode de vie sobre.

De l'autre, on pourrait également regretter qu'il néglige, dans l'exploration des relations entre les humains et les artefacts, la dimension religieuse ou spirituelle, pourtant bien présente dans le marketing des objets connectés : par exemple, parce qu'est assez clairement accordé le rôle d'agent sociaux aux assistants vocaux dans les spots publicitaires qui les promeuvent, s'exprime actuellement un néo-animisme conférant un sens original au concept marxien de « fétichisme de la marchandise » (en reprenant les analyses de l'anthropologue Fabienne Martin-Juchat, développées dans une étude récente²²). Grâce à sa démarche nourrie

¹⁹ Le panoptique désigne le dispositif carcéral imaginé par Jeremy Bentham en 1791, que Michel Foucault invite à considérer comme un modèle de l'esprit moderne en manière pénale dans *Surveiller et punir. Naissance de la prison*, Paris, Gallimard, 1975.

²⁰ Maurice Merleau-Ponty, *Phénoménologie de la perception*, Paris, Gallimard, 1945 ; Jacques Derrida, *Le Toucher*, Jean-Luc Nancy, Paris, Galilée, 2000.

²¹ On entend par cette expression toutes les manifestations par lesquelles un individu mesure ce qui peut l'être dans son activité corporelle *via* des capteurs variés (capteurs du pouls et de la tension artérielle, de la sudation, du nombre de pas effectués, etc.) et de fait rythme cette activité par des « alertes » qui contribuent à soumettre son existence à la double normativité biologique et technique. Voir Deborah Lupton, *The Quantified Self. A Sociology of Self-tracking*, Cambridge UK & Malden MA, Polity Press, 2016.

²² Fabienne Martin-Juchat, Sur le néo-animisme technologique à l'ère de l'engouement pour l'Intelligence Artificielle, *Quaderni. Communication, Technologies, Pouvoir*, n°105, 2022, p. 53-72.

d'observations personnelles acérées et d'ingénieuses expériences mentales, l'ouvrage de Pierre Cassou-Noguès assigne joyeusement la philosophie à une réflexion à la fois épistémologique, ontologique, éthique et politique sur la condition technologique, que l'on pourra compléter avec la considération de ces deux dimensions.

Thierry Ménissier

Thierry Ménissier est professeur de philosophie politique à l'Université Grenoble Alpes, membre de l'Institut de Philosophie de Grenoble (IPhiG) et actuellement responsable scientifique de la chaire « éthique & IA » au sein de l'institut multidisciplinaire en intelligence artificielle de Grenoble MIAI. Ses travaux concernent la dynamique de l'innovation dans le contexte des transitions et interrogent de quelles manières l'usage des technologies modifie les concepts hérités de la philosophie éthique et politique moderne, tels que l'autonomie, la responsabilité et l'autorité. Derniers ouvrages parus : *Innovations. Une enquête philosophique*, Hermann, 2021 ; *Ethique, politique, philosophie des techniques* (dir.), Librairie Philosophique J. Vrin, 2021.