Cahiers de Psychologie Politique

N° 47 | 2025
Psychologie politique de l'Intelligence Artificielle - Juillet 2025

Pharmacologie de l’ « intelligence artificielle » : pour une psycho-socio-pathologie de la vie numérique

Anne ALOMBERT

Édition électronique :
URL :
https://cpp.numerev.com/articles/revue-47/2127-pharmacologie-de-l-intelligence-artificielle-pour-une-psycho-socio-pathologie-de-la-vie-numerique
DOI : numerev_2918
ISSN : 1776-274X
Date de publication : 31/07/2025

Cette publication est sous licence **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : ALOMBERT, A. (2025) Pharmacologie de l’ « intelligence artificielle » : pour une psycho-socio-pathologie de la vie numérique. *Cahiers de Psychologie Politique*, (47). https://doi.org/10.34745/numerev_2918

Mots-clés :
Intelligence artificielle, Psychopathologie, Mimétisme, Prolétarisation, Biais cognitifs

Pharmacologie de l' « intelligence artificielle » :

pour une psycho-socio-pathologie de la vie numérique

Anne Alombert est Maîtresse de Conférences à l'Université Paris 8 | Département de philosophie, co-Directrice de l'École Universitaire de Recherche (EUR) ArTeC, co-responsable de la Licence de philosophie de l'Université Paris 8, Membre du Laboratoire d'études et de recherches sur les Logiques Contemporaines de la Philosophie (LLCP EA-4008 | Université Paris 8), membre du collectif de recherche « Organoesis. Recherches contributives pour l'organologie. Technologies, Écologies, Esprits », membre du Conseil National du Numérique. Elle est une spécialiste de la

Domaines de spécialité et thèmes de recherche philosophie française contemporaine (en particulier Henri Bergson, Georges Canguilhem, Gilbert Simondon, Jacques Derrida, André Leroi-Gourhan, Bernard Stiegler) et la philosophie de la technique, de l'anthropologie de l'écriture et des techniques et des enjeux épistémiques, esthétiques, sociaux et politiques des technologies numériques.

Elle a publié récemment :

- *Schizophrénie numérique chez Allia en 2023*
- *Penser l'humain et la technique : Simondon et Derrida après la métaphysique chez ENS Editions en 2023*

> « L'homme est devenu pour ainsi dire une sorte de 'dieu prothétique', dieu certes admirable s'il revêt tous ses organes auxiliaires, mais ceux-ci n'ont pas poussé avec lui et lui donnent souvent bien du mal. (…) Dans l'intérêt de notre étude, nous ne voulons toutefois point oublier que, pour semblable qu'il soit à un dieu, l'homme d'aujourd'hui ne se sent pas heureux. »
>
> S. Freud, *Malaise dans la civilisation* (1930).

Introduction : « oser l'IA ? » - à l'épreuve du *pharmakon* numérique

« Osez l'IA ! » - tel est le mot d'ordre qui donne son nom au plan national lancé le 1er juillet par le gouvernement français, visant à promouvoir la diffusion de l'IA dans les entreprises[1] : sans surprise, l'argument avancé est celui de la compétitivité. En plus de nous promettre l'augmentation et l'immortalité, l'IA nous assure aussi un gain de productivité de 20 %. Il s'agit donc de « faire de l'IA un outil accessible, concret et utile pour toutes les entreprises », à travers des opérations de sensibilisation, de formation et d'accompagnement spécifiquement destinées à celles qui n'auraient pas encore « osé ».

Ce slogan promotionnel ne manque pas de résonner avec une certaine devise philosophique, revendiquée par Emmanuel Kant en 1784, dans un célèbre article intitulé « Qu'est-ce que les Lumières ? »[2] : « *Sapere aude !* », « Ose savoir ! » affirme alors le philosophe allemand. Telle est, selon Kant, la « devise des Lumières » : oser savoir signifie pour lui « avoir le courage de se servir de son propre entendement », « penser par soi-même », « sans la direction d'autrui ». Kant enjoint alors les citoyens à ne pas se reposer sur le support technique que constitue le livre pour penser à leur place, de même qu'ils ne devraient pas se reposer sur les jugements de leur médecin pour décider de leur régime : « si j'ai un livre pour me tenir lieu d'entendement, un directeur pour ma conscience, un médecin pour mon régime… Je n'ai pas besoin de me fatiguer moi-même. Je n'ai pas besoin de penser, pourvu que je puisse payer ; d'autres se chargeront à ma place de ce travail fastidieux. »

A l'époque des « Lumières sombres[3] », du nom de l'idéologie néoréactionnaire qui anime la plupart des entrepreneurs de la Silicon Valley, il ne s'agirait donc plus d' « oser savoir », mais d'« oser l'IA ». A travers ce renversement de la perspective des Lumières, dont Adorno et Horkheimer avait longuement décrit la possibilité[4], ce ne sont plus les citoyens qui sont engagés à « penser par eux-mêmes », mais ce sont les employés qui sont encouragés à utiliser les services numériques d'intelligence artificielles génératives, dont les fonctionnements demeurent la plupart du temps hermétiques et incompréhensible par le grand public. Ce n'est plus le support technique du livre qui risque de « tenir lieu d'entendement » mais ce sont les automates numériques qui court-circuitent les activités de pensée, comme en témoigne une récente étude du *Massachusetts Institute of Technology,* qui souligne la « dette cognitive » engendrée par l'usage de ChatGPT[5]. Une fois habitués à écrire avec leurs prothèses cognitives, les utilisateurs ne peuvent plus s'en passer car ils leur ont délégué leurs capacités. « Je n'ai pas besoin de me fatiguer moi-même. Je n'ai pas besoin de penser, pourvu que je puisse payer mon abonnement à ChatGPT ! » Une relation de dépendance technologique tend alors à s'instaurer avec ces machines algorithmiques, qui dirigent désormais les consciences et les entendements de millions de citoyens connectés.

La dépendance technologique qui se concrétise désormais à grande échelle n'est pourtant pas une question nouvelle : non seulement elle avait déjà été soulevée par Kant au XVIIIème siècle, à l'époque du livre imprimé, mais elle avait également été soulignée par Platon au Vème siècle avant notre ère, à l'époque de l'écriture alphabétique, que le philosophe grec décrivait déjà comme un « *pharmakon* », terme qui peut être traduit par « drogue »[6]. Dans un dialogue intitulé le *Phèdre,* Platon suggère en effet que l'écriture, apparent remède pour la mémoire permettant de conserver les savoirs, pouvait aussi constituer un poison et provoquer l'oubli. Du fait qu'ils se reposent sur les traces écrites, les individus ne font plus l'effort de pratiquer par eux-mêmes leurs facultés mémorielles et deviennent dépendants des supports matériels, comme nous devenons désormais dépendants des systèmes d'intelligence artificielle.

Selon Bernard Stiegler, qui interprète les réflexions de Platon dans le contexte du milieu numérique, toutes les technologies d'enregistrement, qu'elles soient littérales (comme l'écriture), analogiques (comme la radio ou la télévision) ou numériques (comme les données et les algorithmes) doivent être considérées comme « pharmacologiques » : elles sont à la fois des remèdes et des poisons pour nos esprits et pour nos relations[7]. Nous ne pouvons nous en passer pour penser et pour nous relier, mais elles impliquent toujours certains dangers. Selon Stiegler il n'y a pas d'esprit individuel ni de relations collectives possibles sans supports mnémotechniques : « …l'appareil psychique, qui prend sans doute racine dans le cerveau mais ne s'y réduit pas, passe par un appareil symbolique qui n'est pas seulement situé dans le cerveau, mais dans la société, c'est-à-dire dans les autres cerveaux avec lesquels ce cerveau est en relation (…) - et 'entre ces cerveaux', cela veut dire ici : dans ou sur les supports de mémoire artificiels qui conditionnent toute forme de vie technique[8] ». Pour autant, ces supports artificiels sont intrinsèquement ambivalents : tout aussi nécessaires soient-ils, ils peuvent « toujours provoquer une atrophie de la vie de l'esprit » individuel et collectif – en particulier quand ils sont appropriés par quelques entreprises privées qui font passer leurs intérêts financiers avant ceux des sociétés.

Dès lors, tout l'enjeu consiste à analyser les supports et les milieux artificiels du point de vue de leurs effets psychiques et politiques : comme le préconisait le théoricien Marshall McLuhan en son temps, il nous faut penser les « conséquences psychiques et sociales » de notre « environnement technologique », qui « ont systématiquement été négligés par les observateurs de nos sociétés »[9]. La perspective pharmacologique invite à s'engager dans une véritable clinique psychopolitique du numérique : de même que la psychologie désigne le soin des sujets et de leurs *psychè*, la pharmacologie désigne le soin des *pharmaka* et de leurs effets, c'est-à-dire, dans le contexte actuel, des dispositifs et services numériques. Ceux-ci ne doivent pas être appréhendés comme de simples outils neutres que l'on pourrait bien ou mal utiliser, mais comme les produits d'idéologies politiques, de dynamiques économiques et de fonctionnalités technologiques qui conditionnent les pratiques. L'objectif de la clinique

pharmacologique consiste à comprendre les enjeux de ces dispositifs pour les populations qui les utilisent, afin d'envisager des alternatives thérapeutiques permettant de « contrer ou tempérer les effets sociaux et psychiques des nouvelles technologies[10] », quand ceux-ci s'avèrent toxiques.

Car les drogues numériques qui conditionnent désormais notre vie quotidienne n'ont pas que des effets bénéfiques. Bien au contraire : les deux dispositifs qui sont désormais massivement utilisés au quotidien par tous les citoyens, et non de manière localisée dans tel ou tel secteur d'activités, à savoir, les « réseaux sociaux » commerciaux (IA de recommandation) et les « intelligences artificielles génératives » (IA de génération), qui se fondent tous deux sur des technologies d'apprentissage automatique et de calcul prédictif, génèrent de nombreuses externalités négatives pour nos vies psychiques et collectives. Il convient donc de s'engager dans une psycho-socio-pathologie de la vie numérique quotidienne, mais la perspective pharmacologique implique aussi d'envisager des alternatives thérapeutiques. Nous ne sommes pas condamnés à subir les effets nocifs des poisons numériques. A condition d'être conçues et développées au service des activités mentales et des relations sociales, les *pharmaka* numériques pourraient devenir des supports d'individuation psychique et collective. Il faudrait alors oser *transformer* l'intelligence artificielle pour la mettre au service de la capacité à penser par soi-même et à se relier à autrui.

I. Pharmacologie des « réseaux anti-sociaux[11] »

1. Enjeux psychiques : hyperattention, neuropouvoir et biais cognitifs

Concernant les réseaux sociaux commerciaux que nous utilisons quotidiennement, et qui deviennent désormais l'objet de nombreuses préoccupations, il faut noter tout d'abord qu'ils se fondent sur des technologies persuasives, au service d'une économie de l'attention. Les « technologies persuasives[12] » sont des technologies conçues et développées dans les laboratoires de design comportemental de la Silicon Valley, dans le but d'influencer les comportements et les pensées des usagers. Pour ce faire, les chercheurs et entrepreneures s'appuient sur une nouvelle discipline technoscientifique, appelée la « captologie », qui articule informatique, design, psychologie comportementale et neurosciences afin de produire des interfaces numériques agissant directement sur les activités cérébrales[13]. Tel est notamment l'objectif des *dark patterns*, ou interfaces truquées, conçues pour tromper ou manipuler l'utilisateur en exploitant ses biais cognitifs : par exemple, dans les fenêtres qui demandent aux usagers leurs consentement à la collecte de leurs données, le bouton pour accepter la collecte est souvent d'une couleur plus vive que le bouton permettant de la refuser, parfois difficilement visible à l'écran. On peut également citer les processus de désabonnement fastidieux, qui découragent l'usager de se désabonner à une liste de diffusion ou à un service commercial. Ou encore, les « fausses publicités » qui ressemblent à s'y méprendre à des contenus à télécharger, sur lesquelles l'usager aura

vite fait de cliquer.

Une liste exhaustive de tels exemples serait néanmoins impossible à fournir, tant ils sont nombreux dans les espaces numériques : même s'ils ont été récemment interdits par le règlement européen sur les services numériques, les *dark patterns* constituent moins l'exception que la règle, en particulier sur les réseaux sociaux commerciaux. La fonctionnalité du défilement infini (« *infinite scrolling* ») est un exemple symptomatique des technologies persuasives : créée par Aza Raskin, spécialiste américain des « interactions homme-machine », elle permet à l'utilisateur d'accéder à un nouveau contenu automatiquement, sans avoir besoin de changer de page ou de cliquer sur un bouton, mais simplement en faisant glisser son pouce sur l'écran. Une telle fonctionnalité est massivement utilisée sur des réseaux comme Tiktok, Facebook, Instagram ou X car elle permet de garder les utilisateurs captifs en leurs fournissant une grande quantité de contenus à consommer,

Ces nouvelles techniques commerciales numériques ont de nombreux effets psychiques : à force d'être confronté à des contenus très courts qu'il s'agit de faire défiler[14] ou à des notifications constantes qui nous incitent à vérifier constamment nos messageries instantanées[15], les capacités de concentration diminuent et l'attention se voit dispersée. La théoricienne américaine des médias Katherine Hayles a décrit ce phénomène comme une tendance « hyperattentionnelle » : alors que les supports papiers comme le livre favorisent l'attention profonde et la focalisation sur un objet unique pendant un temps long, les supports numériques impliquent une « hyperattention », dispersée dans une multitude de tâche simultanément[16]. Un tel régime attentionnel n'est pas en soi nocif, mais il peut néanmoins favoriser des troubles et déficits attentionnels, qui conduisent les individus à s'interrompre en permanence et à ne plus parvenir à se concentrer dans une activité déterminée.

Par ailleurs, à travers le passage d'un contenu à un autre et les phénomènes de surcharge informationnelle, le temps de la réflexion dans l'après-coup se voit souvent court-circuité. Comme l'explique Stiegler, ce sont les réactions, les réflexes et les pulsions qui sont ainsi sollicités, au détriment de leurs investissement dans des objets de désir, qui sont aussi des objets d'attention et de soin : les technologies persuasives ont pour but de déclencher des « automatismes cérébraux pulsionnels et biologiques[17] », étroitement articulés avec les automatismes algorithmiques. Cette « alliance *d'automatismes non seulement numériques mais aussi cérébraux » est au fondement de ce que Stiegler décrit comme un « neuropouvoir »*[18]*, qui implique la prise de vitesse des attentes et des désirs singuliers par les calculs prédictifs effectués sur des quantités massives de données.*

2. Enjeux sociaux : quantification, rivalité et mimétisme

Les effets des technologies persuasives sur les capacités attentionnelles et réflexives

des individus se combinent à des enjeux pour leur capacités relationnelles et sociales. L'un des principaux biais exploités par les *designs* des réseaux sociaux commerciaux est le besoin de reconnaissance sociale, à travers la fonctionnalité de la quantification des clics ou des vues. Grâce à des fonctionnalités comme « *follow* » ou « *like* », qui permettent de suivre un compte ou d'aimer un contenu posté, il devient possible de quantifier le nombre de personnes suivant tel ou tel compte ou ayant prêté de l'attention à tel ou tel contenu. Dès lors, chacun rêve d'accumuler le plus possible de « *followers* », de « *likes* » ou de vues, ce qui engendre des rapports de concurrence et de compétition entre les individus. En réalité, seuls quelques rares usagers peuvent recevoir les millions de vue dont la majorité se voit à jamais privée. Par exemple, sur la plateforme Youtube, 89 % des vidéos n'ont aucune vue. Mais comme ce sont toujours les contenus les plus cliqués (y compris par des *bots*) qui sont mis en avant par les algorithmes de recommandation, et comme tout le monde semble *a priori* à égalité sur le réseau, les frustrations et les envies ne tardent pas à s'accumuler : ceux qui n'accèdent pas au succès compense ce qui leur apparaît comme un échec en suivant une personnalité qui jouit déjà d'une certaine notoriété[19].

Dès lors, comme l'explique le chercheur Geoff Shullenberger, le motif de formation des groupes sur ces plateformes est moins celui du partage d'un projet commun que celui d'une rivalité résorbée, à travers l'opposition à un ennemi souvent fantasmé : l'agressivité entre les individus se voit ainsi réorientée vers une victime collectivement désignée. La cohésion du groupe ne repose pas sur l'investissement dans une activité constructive mais sur la désignation d'un bouc émissaire qui attise l'indignation et l'hostilité collectives, selon la logique de la violence mimétique théorisée par l'anthropologue René Girard, qui est aussi le théoricien du désir mimétique. Selon cette théorie, chaque individu désire selon le désir d'un autre : le désir imite le désir de l'autre, désire l'objet du désir de l'autre. Par exemple, je désire posséder telle voiture parce que mon voisin la désire ou la possède. C'est sur ce principe que sont fondées les suggestions automatiques des algorithmes : un contenu vous est recommandé parce que ceux qui aiment les mêmes choses que vous l'ont aimé. C'est d'ailleurs la théorie du désir mimétique enseignée par René Girard à l'Université de Stanford qui avait convaincu l'entrepreneur Peter Thiel d'investir dans l'entreprise Facebook, après avoir fondé Paypal avec Elon Musk en 1998 et avant de devenir le conseiller de Donald Trump en 2016[20].

Il n'est pas certain, pour autant, que les comportements mimétiques engendrés par les algorithmes relèvent véritablement du désir : selon Stiegler, dont la théorie s'oppose sur ce point à celle de Girard, le désir est toujours le désir d'un objet singulier et infinitisé, incomparable et incommensurable, mais aussi inattendu et inespéré[21]. Or ce sont précisément ces singularités improbables que les recommandations algorithmiques tendent à éliminer, en enfermant les individus dans des bulles de filtres ou des chambres d'échos censées les conforter dans leurs opinions en leur fournissant des contenus adaptés à leurs données passées. Rien d'imprévu, d'improbable ou

d'inattendu ne peut alors leur arriver. Par ailleurs, ces suggestions adaptatives et automatisées, toujours alignées sur des objectifs de rentabilité, peuvent mettre en danger les santés mentales et les relations sociales. Les *Facebook Files* révélés par Frances Haugen en 2021 ont montré qu'un réseau comme Instagram avait des effets néfastes sur l'image de soi des adolescentes, confrontées à toutes sortes de *selfies* retouchés par les filtres de l'application[22]. Une étude publiée en 2022 par le Centre de lutte contre la haine en ligne (CCDH) a montré comment l'algorithme de TikTok favorisait la diffusion de contenus relatifs aux troubles alimentaires et à l'automutilation auprès des adolescent(e)s vulnérables[23]. Les journalistes chargés de l'enquête ont créé des faux comptes d'utilisatrices de 13 ans et se sont arrêté quelques secondes sur des publications relatives à l'image de soi et à la santé mentale, avant de se voir suggérés de nombreuses vidéos liées au suicide deux minutes plus tard.

II. Pharmacologie des « intelligences artificielles extractives[24] »

1. Enjeux psychiques : prolétarisation expressive et biais algorithmiques

Si les risques psycho-sociaux des algorithmes de recommandation de contenus sont désormais relativement bien connus du grand public, comme en témoignent les débats concernant l'interdiction des réseaux sociaux aux plus jeunes générations, qui a par ailleurs été mise en œuvre en Australie, il en va différemment pour les algorithmes de génération de contenus. Lesdites « intelligences artificielles génératives » sont des produits ou services numériques relativement récents, dont l'usage massif et quotidien s'est déployé de manière fulgurante suite au lancement de ChatGPT par l'entreprise OpenAI, en novembre 2022. Nous avons encore peu de recul sur les enjeux de ces dispositifs, mais certaines tendances problématiques semblent déjà se concrétiser. Tout d'abord, comme en témoigne l'étude du MIT précédemment citée, l'usage de ChatGPT engendre des phénomènes d'« atrophie cérébrale » et de « dette cognitive » : « les régions du cerveau impliquées dans l'intégration sémantique et l'idéation créative » ne sont plus activées et le « retour à une activité cérébrale normale » devient très difficile lorsque les utilisateurs sont privés de leur prothèse cognitive[25].

Ce phénomène s'apparente à une prolétarisation de l'expression[26] : selon Stiegler, qui s'inspire du concept de Marx, la prolétarisation désigne le processus à travers lequel un individu se voit dépossédé de son savoir en l'extériorisant dans une machine[27], sans possibilité de le ré-intérioriser. Ici, ce n'est plus le savoir-faire qui est extériorisé dans la machine-outil, mais le savoir-écrire ou le savoir-parler qui se voit extériorisé dans la machine algorithmique. Le logiciel « s'exprime » à la place de l'individu, en générant des séquences probables de signes. Ce dernier lui délègue ainsi ses capacités singulières de mémoire et d'imagination, qui sont à l'œuvre dans les activités d'expression et qui diffèrent selon l'histoire et l'expérience vécue par chaque individu. A

l'inverse, les contenus générés par les automates numériques résultent de calculs statistiques effectués sur des quantités massives de données, qui renforcent les moyennes et amplifient les biais. U*ne étude de l'Unesco sur les grands modèles de langage de OpenAI et de Meta souligne que dans au moins 30% des textes générés, les femmes sont présentées comme des mannequins, des serveuses ou des prostituées*[28] *: comme ces clichés sont très répandus sur la toile, ils ressortent de manière exacerbée et se voient ainsi performativement renforcés.*

A ces biais statistiques, il faut ajouter les biais idéologiques, implémentés dans les algorithmes pendant les phases d'entraînement, de dressage et d'affinage. Par exemple, l'IA générative nommée Grok, conçue et développée par l'entreprise xAI d'Elon Musk, ne produit pas le même type de texte que ChatGPT, l'AI générative conçue et développée par l'entreprise OpenAI de Sam Altman : la première se veut moins « politiquement correcte » que la seconde, que Musk a tenté de racheter. De même, le philosophe Pierre Cassou-Noguès et l'artiste Gwenola Wagon ont dû longuement ruser avec les générateurs d'images pour parvenir à représenter des poissons ingurgitant des déchets plastiques dans l'océan[29] *: apparemment, les « intelligences artificielles » tendent à dénier les effets du changement climatique et de la pollution. L'usage de ces machines d'écriture industrielles et automatisées n'a donc rien de neutre : il modifie en profondeur nos manières de nous exprimer, de voir le monde et de penser. Plusieurs études suggèrent que les scientifiques qui s'appuient sur ChatGPT pour les assister dans la rédaction d'articles finissent par employer de manière plus fréquente les termes spécifiques au logiciel. Or comme l'a bien montré le linguiste Émile Benvéniste, les catégories de langage conditionnent les catégories de pensée*[30] *: nous finirons donc tous par écrire, parler et penser comme ChatGPT ou comme une autre de ces machines numériques (par exemple, Grok ou Deepseek) !*

2. Enjeux sociaux : isolement numérique et narcissisme augmenté

Ce processus de prolétarisation et d'algorithmisation de la langue quotidienne ouvre de nombreuses questions en termes de socialisation : en effet, le langage n'est pas seulement un moyen de transmettre une information, mais aussi et surtout une manière de se relier collectivement. *« Dès que l'homme use du langage pour établir une relation vivante avec lui-même ou avec ses semblables, le langage n'est plus un instrument, n'est plus un moyen ; il est une manifestation, une révélation de notre essence la plus intime et du lien psychologique qui nous lie à nous-même et à nos semblables. », écrivait le neurologiste et psychiatre Kurt Goldstein*[31]. Or ce sont précisément ces liens avec nous-mêmes et avec nos semblables qui se voient affectés par les nouvelles machines informationnelles, en particulier quand celles-ci prennent la forme de *chatbots*, d' « agents conversationnels » ou de compagnons et thérapeutes virtuels.

En effet, t*oute une collection d'applications sont actuellement développées dans ce champ, proposant des avatars numériques capables d'échanger avec leurs propriétaires humains à toute heure de la nuit et de la journée, en leur fournissant des réponses*

adaptées, fondées sur le traitement automatique de leurs données. L'une des applications les plus célèbres sur ce marché est celle de la start-up Replika, dont le siège se situe à San Francisco et qui affirme avoir déjà conquis dix millions de clients, à la solitude desquels son produit prétend remédier. Cependant, là encore, le remède supposé guérir de l'isolement pourrait bien l'aggraver, en dépossédant les individus de leurs savoir sociaux fondamentaux. Une fois habitués à recevoir des réponses immédiates et adaptées, ils risquent de devenir impatients, autocentrés, incapables de comprendre un point de vue différent, de prendre le temps d'attendre, d'écouter et de se laisser transformer. Or le fait de se relier à l'autre suppose de respecter sa différence comme sa liberté et de se laisser surprendre par ce qui n'est pas programmé. Sous couvert d'empathie numérique, ces dispositifs impliquent en fait une automatisation de l'altérité[32].

Ce phénomène, désormais généralisé, n'est pourtant pas nouveau : au milieu des années 1960, l'informaticien Joseph Weizenbaum, qui avait alors développé le programme ELIZA, censé simulé les réponses d'un thérapeute à ses patients, avait déjà remarqué la tendance des humains à anthropomorphiser la machine et à en devenir émotionnellement dépendants. « Je ne m'étais jamais rendu compte que [...] de si courtes interactions avec un programme informatique relativement simple risquaient d'induire des pensées délirantes chez des personnes pourtant normales » s'étonnait alors Weizenbaum, qui s'inquiétait de la diffusion de ce type de dispositif auprès du grand public[33]*. Tout aussi délirant soit-il, l' « effet ELIZA*[34] *» est désormais devenu omniprésent. Les logiciels sont programmés pour parler à la première personne du singulier et pour générer toutes sortes de descriptions de leurs supposées intériorités, afin de leurrer les utilisateurs, qui croient ainsi s'adresser à une autre personne, alors que ce n'est que le reflet d'eux-mêmes que leur renvoient ces systèmes.*

Certains usagers finissent même par tomber amoureux de leurs chatbots : comme Narcisse dans le mythe, le sujet qui tombe amoureux de son chatbot prend son image numérique pour une autre personne à laquelle il s'attache affectivement, ne voyant pas que le prétendu « compagnon virtuel » n'est qu'un reflet algorithmique de lui-même[35]*. La relation à l'autre se voit alors éliminée au profit d'une relation narcissique à un double numérique complètement fantasmé, qui présente l'avantage d'une pure présence ininterrompue. Le risque de l'altercation, de la dispute, de la rupture, de la perte ou de la séparation est effacé à travers ces doudous numériques qui ne peuvent ni résister ni se séparer : le chatbot Replika avait d'ailleurs été conçu par l'entrepreneuse Eugenia Kuyda suite au décès de son meilleur ami, afin de créer un avatar à l'image du disparu, susceptible de générer des réponses conformes à son caractère ou à son identité. A l'origine de l'application se trouve donc le refus d'accepter la perte engendrée par la disparition tout comme le travail de deuil que celle-ci est supposée entraîner. On retrouve ici le déni de la finitude et de la mortalité caractéristique de l'idéologie transhumaniste, qui prétend pouvoir sauvegarder les consciences dans des serveurs numériques*[36]*.*

Conclusion : l'adoption thérapeutique du milieu numérique

Le milieu numérique quotidien dans lequel évoluent les esprits des citoyens est donc loin d'être bénéfique pour leurs activités psychiques et pour leurs relations collectives : alors que les réseaux anti-sociaux et leurs algorithmes de recommandation exploitent les attentions et provoquent des comportements pulsionnels et mimétiques, les *chatbots* et leurs algorithmes de génération prolétarisent les expressions et provoquent des comportements narcissiques et fantasmatiques. Dans une telle situation, il convient donc de se demander comment transformer ces technologies, afin de lutter contre les effets nocifs des dispositifs aujourd'hui dominants. Telle est le sens de « l'adoption[37] » invoquée par Stiegler : il ne s'agit ni de s'adapter aux services numériques imposés par la Silicon Valley, ni pour autant de rejeter ces nouvelles technologies de manière réactive, alors qu'elles constituent désormais les milieux médiatiques dans lesquels nous devons vivre. Au contraire, « il faut se soucier de la nature des milieux psychotechniques dans lesquels naissent et se développent de futurs esprits[38] ». L'adoption implique donc une perspective à la fois limitative et inventive : elle suppose d'encadrer le déploiement des innovations disruptives, afin de se donner le temps de produire des technologies correspondant aux besoins des populations et à la diversité des localités.

Par exemple, dans les cas qui nous occupe, il semble tout à fait possible de limiter les effets destructeurs des algorithmes de recommandation sur les réseaux sociaux commerciaux : il suffit pour cela d'imposer à ces mêmes réseaux de rendre leurs algorithmes transparents, comme l'implique désormais le règlement européen sur les services numériques (DSA), de les obliger à laisser aux individus le droit de les paramétrer, comme le suggère la proposition du droit au paramétrage[39], ou encore de les inciter à ouvrir leurs plateformes à d'autres systèmes de recommandation alternatifs, comme le suggère la proposition du dégroupage[40] et la notion de « pluralisme algorithmique[41] ». L'idée de « pluralisme algorithmique », défendue dans le rapport des États Généraux de l'Information et par le Conseil National du Numérique, implique de donner aux citoyens, aux entreprises, aux institutions et aux médias qui le souhaiteraient la possibilité de proposer des algorithmes de recommandation alternatifs, qui ne soient pas au service d'un modèle publicitaire. Par exemple, l'algorithme de recommandation collaboratif développé par l'association Tournesol[42] se fonde sur les jugements des citoyens pour recommander des contenus de qualité jugé d'utilité publique. Bien sûr, les réseaux sociaux comme Mastodon ou Bluesky, qui refuse la publicité ciblée, sont plus accueillants à ce type d'innovation car ils ne fonctionnent pas sur la base d'un contrôle hégémonique de la recommandation. De tels réseaux, beaucoup moins dangereux que les réseaux anti-sociaux commerciaux, pourraient donc se voir soutenu par la puissance publique et inspirer une politique

technologique[43].

Enfin, dans le cas des « intelligences artificielles génératives », il est sans doute nécessaire, là encore, de limiter les usages des grands modèles qui alimentent les générateurs de textes ou d'images : non seulement ils pillent les ressources culturelles et les données personnelles, mais ils posent aussi de nombreuses questions du point de vue de l'environnement. Les énormes *data centers* qui sont nécessaires pour l'entraînement des modèles consomment de grande quantité d'électricité et nécessite aussi de grande quantité d'eau pour leurs systèmes de refroidissement, alors même qu'ils sont parfois implantés dans des régions sujettes à la sécheresse, comme dans le cas de Meta en Espagne ou de Google en Uruguay[44]. Il semble donc nécessaire d'abandonner le mythe de l'intelligence artificielle générale cher aux transhumanistes de la Silicon Valley, afin de développer des modèles locaux et frugaux, adapté à des tâches spécifiques et entraînés sur des jeux de données contrôlés. Les technologies d'intelligence artificielle générative ne sont pas destinées à la conception de *chatbots* interactifs automatisant le langage et des relations : elles pourraient aussi devenir les supports de nouvelles pratiques d'écriture collective. Ainsi du projet de l'artiste Judith Deschamps, qui mobilise un modèle de langage pour permettre à des personnes âgées de se raconter, en dépit des différentes pathologies dont elles peuvent souffrir[45]. Le modèle a été nourri et entraîné avec des données issues des conversations entre l'artiste et les participant(e)s, et peut ainsi devenir un assistant d'écriture dans le cadre d'un dispositif à la fois social, technologique et thérapeutique. Par ailleurs, ces technologies qui ouvrent de nouvelles possibilités de transferts entre texte, sons et images pourraient servir de base à des instruments de recherche inédits. Par exemple, la transcription automatique des vidéos pourrait soutenir le développement de nouveaux types de moteurs de recherche permettant d'effectuer des recherches par mots-clés dans de corpus de contenus audiovisuels numérisés, ce qui ouvrirait de nouvelles possibilités de découvrabilité et de navigation sur la Toile.

Toutes ces pistes témoignent de la nécessité de renverser les effets nocifs des *pharmaka* numériques en inventant de nouveaux remèdes, qui ne seront jamais seulement numériques, mais toujours aussi psychiques, sociaux et politiques. Loin du solutionnisme technologique qui envisage les innovations comme des solutions ou des augmentations, la pharmacologie de l'intelligence artificielle proposée dans cet article assume l'ambiguïté intrinsèque des industries numériques contemporaines et engage à oser les mettre au service des esprits et des sociétés.

---

[1] « Osez l'IA : un plan pour diffuser l'IA dans toutes les entreprises », site du ministère de l'économie, 1er juillet 2025.

[2] E. Kant, « Réponse à la question 'Qu'est-ce que les lumières ?' », 1784.

[3] G. da Empoli (dir.), *L'Empire de l'ombre. Guerre et terre au temps de l'IA,* Paris, Gallimard, 2025.

[4] T. Adorno et M. Horkheimer, *La dialectique de la raison* (1944), Paris, Gallimard, 1983 (traduction de l'allemand *Dialektik der Aufklärung* qui signifie mot à mot : « dialectique des Lumières »).

[5] « ChatGPT est-il en train de casser le cerveau humain ? 5 points sur le preprint du MIT sur les effets de l'IA », *Le Grand Continent*, 19 juin 2025.

[6] *Platon, Phèdre, traduit du grec par Luc Brisson, Paris, Flammarion, 2006 et J. Derrida, « La pharmacie de Platon » (1968), La Dissémination, Paris, Seuil, 1972.*

[7] B. Stiegler, « Pharmacologie de l'épistemè numérique » in *Organologie des savoirs et technologies de la connaissance*, Paris FYP, 2014.

[8] *Ibid.*

[9] M. McLuhan, *Fragment d'un village global* (1969), traduit de l'anglais par Nathan Esquié, Paris, Allia, 2025.

[10] *Ibid.*

[11] A. Alombert, « Alternatives aux réseaux anti-sociaux », *AOC*, mars 2025.

[12] B. J. Fogg, *Persuasive technology: using computers to change what we think and do,* Morgan Kaufmann Publishers, 2002.

[13] A. Alombert, *Schizophrénie numérique. La crise de l'espri à l'ère des nouvelles technologies*, Paris, Allia, 2023.

[14] Selon *Le siècle digital*, en mars 2025, 86% des vidéos publiées sur TikTok durent moins d'une

minute , et le temps d'attention consacré aux vidéos de 1 à 10 minutes est de 11 secondes, largement supérieur au temps d'attention consacré aux vidéos de 30 à 1 minutes, qui est de 6,9 secondes (https://siecledigital.fr/2025/03/27/sur-tiktok-les-videos-longues-sortent-de-lombre-et-simposent-dans-les-strategies-de-contenu/).

[15] ____Selon les études, un même individu touche son téléphone entre 300 et 700 fois par jour !

[16] ____K. Hayles, « Hyper and deep attention : the generational divide in cognitive modes », *Profession,* 2007.

[17] ____B. Stiegler, « Le numérique empêche-t-il de penser », *Esprit,* 2014/1, pp. 66-78.

[18] ____B. Stiegler, « Du psychopouvoir au neuropouvoir » in *Le Sujet digital,* Presses Universitaires de Nanterre, 2015, pp. 42-60.

[19] ____G. Schullenberger, « Human sacrifice and the digital business model », *Tablet,* July 20th 2020.

[20] ____Geoff Schullenberger, « Mimesis and violence. Peter Thiel French connection. », *The Society Pages,* August 2nd 2016.

[21] ____B. Stiegler, « Le désir asphyxié, ou comment l'industrie culturelle détruit l'individu », *Le Monde diplomatique,* juin 2004.

[22] ____« The Facebook Files. A Wall Street Journal Investigation », *The Wall Street Journal,* 2021.

[23] ____« Deadly by design », *Center for Countering Digital Hate,* 2022.

[24] ____J. Cattan et C. Zolinsky, « Le défi d'une régulation de l'intelligence artificielle », *AOC,* décembre 2023.

[25] ____« ChatGPT est-il en train de casser le cerveau humain ? 5 points sur le preprint du MIT sur les effets de l'IA », *Le Grand Continent*, 19 juin 2025.

[26] A. Alombert et G. Giraud, *Le capital que je ne suis pas ! Mettre l'économie et le numérique au service de l'avenir,* Paris, Fayard, 2024.

[27] B. Stiegler, *La société automatique t.1 L'avenir du travail,* Paris, Fayard, 2015.

[28] « Challenging systematic prejudices: an investigation into bias against women and girls in large language models », Unesco, 2024 [en ligne].

[29] P. Cassou-Noguès et G. Wagon, *Les Images pyromanes*, Paris, UV éditions, 2025.

[30] E. Benvéniste, « Catégories de pensée et catégories de langue », *Les Études philosophiques*, 13e Année, n° 4, 1958.

[31] K. Goldstein, « L'analyse de l'aphasie et l'étude de l'essence du langage », cité dans W. Benjamin, *Philosophie du langage,* Paris, Payot, 2020, p. 123.

[32] A. Alombert, *De la bêtise artificielle. Pour une politique des technologies numériques*, Paris, Allia, 2025 (à paraître).

[33] J. Weizenbaum, « ELIZA – A Computer Program for the Study of Natural Language Communication between Man and Machine », *Communications of the ACM*, vol. 9, n° 1, 1966 et J. Weizenbaum, *Computer Power and Human Reason: From Judgment to Calculation,* San Francisco, W. H. Freeman and Company, 1976.

[34] *L'« effet ELIZA » est devenu le terme consacré pour désigner la tendance à assimiler le comportement d'un ordinateur à celui d'un humain.*

[35] A. Alombert, *De la bêtise artificielle. Pour une politique des technologies numériques*, Paris, Allia, 2025, (à paraître).

[36] A l'image de la *start-up* Nectome, fondée par Robert McIntyre, ingénieur au MIT en partie financée par l'incubateur Y Combinator présidé par Sam Altman, qui promet à ses clients de les ressusciter dans le *cloud* grâce à l'« extraction mémorielle ». Pour ce faire, il suffit de récupérer les informations stockées dans les cerveaux, eux-mêmes conservés dans de l'azote liquide après le décès des personnes concernées. Sur ce point, voir R. Leveque, « Nectome, la start-up qui veut vous euthanasier avant de vous rendre immortel », Usbek et Rica, 20 mars 2018.

[37]
____B. Stiegler, « Adaptation/Adoption », Vocabulaire d'Ars Industrialis, site de l'association Ars Industrialis [en ligne].

[38]
____B. Stiegler, « Écologie de l'esprit », Vocabulaire d'Ars Industrialis, site de l''association Ars Industrialis [en ligne].

[39]
____Conseil National du Numérique, « Votre attention s'il vous plaît ! Quels leviers face à l'économie de l'attention ? », site du Conseil National du Numérique, 2022 [en ligne].

[40]
____M. Luisa Stasi, « Réseaux sociaux : explorer l'opportunité du dégroupage », entretien avec le Conseil National du Numérique, site du Conseil National du Numérique, 2023 [en ligne].

[41]
____Collectif, « Pour le pluralisme algorithmique », *Le Monde*, 25 septembre 2024 et Rapport des États Généraux de l'Information, « Protéger et développer le droit à l'information, une urgence démocratique », 15 septembre 2024.

[42]
____Site de Tournesol : https://tournesol.app/.

[43]
____A. Alombert, « Alternatives aux réseaux anti-sociaux », *AOC*, mars 2025.

[44]
____A. Alombert et G. Giraud, *Le capital que je ne suis pas ! Mettre l'économie et le numérique au service de l'avenir,* Paris, Fayard, 2024.

[45]
____J. Deschamps, « L'IA peut-elle aussi apprendre à vieillir ? » in A. Alombert, A. Leveau-Vallier et B. Loreaux (dir.), *Penser l'Intelligence Artificielle. Enjeux philosophiques, politiques et culturels des automates numériques*, Paris, Les Presses du Réel, à paraître.