

communiquer

Moraliser les machines communicantes. Des barricades morales à l'éthique située : trois cas d'usage de l'IA en milieu professionnel

Carlo Andrea Tassinari

Chercheur post-doctoral, Université de Bologne, Italie

Biographie

Carlo Andrea Tassinari est chercheur en *tenure track* (RTT) au Département de Philosophie de l'Université de Bologne. Il est titulaire d'une bourse Marie Skłodowska-Curie pour le projet VAMP – Voices from the Anthropocene. Maps and Frameworks for Ecological Conflicts (G.A.101106065). Ses recherches, à l'intersection de la sémiotique, des SIC et des sciences sociales, portent sur les tensions socio-environnementales et technologiques comprises comme crises de sens. Ses travaux sont publiés en cinq langues dans plusieurs revues scientifiques nationales et internationales. Il a enseigné la sémiotique, la communication et les études italiennes aux universités de Toulouse, Palerme et Bologne.

Sara De Martino

Chercheuse post-doctorale, École Supérieure Sainte-Anne de Pise, Italie

Biographie

Sara De Martino, Ph.D. en Sciences Politiques, Politiques Européennes et Relation Internationales, est une chercheuse post doctorale au sein de l'École Supérieure Sant'Anna de Pise, dans le département de Sciences Politiques. Elle collabore actuellement sur un projet de recherche d'intérêt national : « Politics and Policy of Digitalisation: Healthcare and Social Policy ». Précédemment, elle a été chercheuse post doctorale à l'Université des Études de Florence, au sein du département de Sciences Politiques et Sociales, ainsi qu'à l'ICAM de Toulouse. Ses intérêts de recherches sont les politiques publiques européennes, les processus de digitalisation et l'Intelligence Artificielle et son impact sur le monde du travail.

Yann Ferguson

Ph.D. Sociologie, Inria, France

Biographie

Yann Ferguson est sociologue à Inria et directeur scientifique du LaborIA, programme de recherche-action mené en partenariat avec le Ministère du Travail et de l'Emploi. Dédiés à l'accompagnement des politiques publiques, ses travaux s'intéressent aux situations de travail qui intègrent l'IA autour de plusieurs enjeux : la transformation des métiers, les conditions de travail, le *management*, l'inclusion, le dialogue social. Il accompagne également plusieurs entreprises et pays sur les enjeux éthiques de l'IA. Lauréat 2018 de la Fondation pour les Sciences Sociales, il est selon l'Usine Nouvelle une des 100 français qui font avancer l'IA en France et dans le monde.

Résumé

La diffusion foudroyante de ChatGPT a fait monter d'un cran la panique que les débats sur l'intelligence artificielle, et notamment sur la moralité de leurs performances linguistiques et communicatives, qu'ils suscitent depuis plus d'un demi-siècle. Dans cet article, nous développerons ce questionnement en deux temps. D'abord, nous proposons une anthropologie documentaire des grands textes d'encadrement éthique sur l'IA, focalisant sur ses impacts sur la communication au travail. Après avoir fait ressortir les spécificités de l'approche universaliste de ces textes, nous examinons trois cas d'usages d'agents conversationnels et des machines communicantes au sein de contextes professionnels, documentés par des entretiens avec usagers, usagères, concepteurs et conceptrices. L'objectif est de

montrer que l'analyse ethnographique des pratiques permet de mettre en évidence des éthiques situées remettant en perspective les principes d'une approche universaliste à la communication machinique. Ce qui conduit, en conclusion, à repenser le rôle-même du chercheur ou de la chercheuse, moins « expert-e » et énonciateur, énonciatrice de règles et de principes, et plus porte-parole analytique de son terrain.

Mots-clés : IA, communication des organisations, barricade morale, éthique située, ethnographie.

Moralizing Communicating Machines. From Moral Barricades to Situated Ethics: Three Use Cases of AI at the Workplace

Abstract

The lightning spread of ChatGPT has kicked up a notch the panic that debates on artificial intelligence, and in particular on the morality of their linguistic and communicative performances, have been stirring up for over half a century. In this article, we will develop this questioning in two stages. First, we propose a documentary anthropology of the major ethical framing texts on AI, focusing on its impacts on communication at work. After highlighting the specificities of the universalist approach of these texts, we examine three case studies of conversational agents and communicating machines in professional contexts, documented by interviews with users and developers. The aim is to show that ethnographic analysis of practices reveals situated ethics putting the principles of a universalist approach on machine communication into a different perspective. This leads, in conclusion, to a rethinking of the role of the researcher, less an «expert» enunciator of rules and principles, and more an analytical spokesperson for his or her field.

Keywords: AI, organizational communication, moral barricade, situated ethics, ethnography

Introduction

La diffusion foudroyante de ChatGPT, lancé en novembre 2022 par la *start-up* californienne OpenAI, a fait monter d'un cran la préoccupation que les débats sur l'intelligence artificielle (IA) suscitent depuis plus d'un demi-siècle. Cette inquiétude est singulièrement associée au développement d'une application particulière de l'IA, celle consacrée au *Natural Language Processing* et à la création d'agents conversationnels – ces entités troublantes qui semblent animées par une faculté de langage semblable à celle qui, selon la tradition aristotélicienne, singularise l'être humain-e. En faisant de l'imitation d'une conversation authentiquement humaine par une machine la preuve de son intelligence, Alan Turing (Turing, 1950) était bien conscient de l'embarras dans lequel il mettait son lecteur ou sa lectrice. C'est pourquoi il concentra son argumentation sur la discussion de « vues contradictoires sur la question principale », « les machines peuvent-elles penser? », plutôt que de lui adresser des arguments positifs convaincants. En faisant de la conversation une faculté proprement humaine et l'imitation, qui s'apparente à une mystification, voire une contrefaçon, les concepts structurants de la machine intelligente, Alan Turing provoqua moralement et préfigura les clivages que l'IA ne cesse de générer depuis sa création en 1956.

En 1965, Joseph Weizenbaum créa « Eliza », le premier logiciel à obtenir des résultats satisfaisants au célèbre « test de Turing » qui évalua la capacité d'un programme informatique à se faire passer pour un-e humain-e aux yeux d'un-e autre humain-e à l'issue d'une conversation de quelques minutes. Simulant une adepte de la thérapie de Rogers, Eliza extrayait les mots-clés des propos de l'utilisateur ou l'utilisatrice et les lui renvoyait. En cas de doute, le logiciel se rabattait sur des phrases parfaitement génériques comme « Veuillez continuer ». En 1966, le *Journal of Nervous and Mental Disease* estimait alors que « plusieurs centaines de patients pourraient être traités en une heure par un ordinateur

conçu à cet effet. Le thérapeute humain, impliqué dans la conception et le fonctionnement du système, ne serait pas supplanté, mais deviendrait beaucoup plus efficace » (Weizenbaum, 1966, p. 152). Épouvanté, Weizenbaum renia sa carrière, laissa tomber le projet Eliza, encouragea ceux et celles qui le critiquaient et devint l'un des plus véhéments adversaires de la recherche sur l'IA¹.

Cette réaction spectaculaire ne manque pas de rappeler le récent appel au moratoire de six mois sur la recherche autour de l'intelligence artificielle². Sorti sur futureoflife.org le 22 mars 2023, peu après que ChatGPT ait dépassé le million d'utilisateurs et d'utilisatrices, l'appel évoquait « des risques majeurs pour l'humanité » entraînés par le développement des grands modèles de langage. Selon l'appel, les investissements de recherche sur ces machines dont l'intelligence serait en mesure de rivaliser avec l'humain (« *human-competitive intelligence* ») devraient être interrompus pour donner le temps aux réglementations d'assurer un encadrement adéquat. Impossible de ne pas remarquer le point commun avec le cas Eliza, à savoir le rôle de lanceur ou lanceuse d'alerte pris en charge par la figure de « l'inventeur ou inventeuse repentie », sorte d'épigone contemporain du docteur Frankenstein de Mary Shelley : comme naguère, Weizenbaum avait lancé, puis désavoué son agent conversationnel. L'un des premiers signataires de l'appel de futureoflife.org – la plupart des quotidiens ayant relayé l'appel l'ont souligné³ – a été Elon Musk (qui par ailleurs est le principal financeur de l'institut) : le milliardaire, connu pour flirter avec le transhumanisme (Menei, 2021), a fourni à OpenAI ses capitaux initiaux (Goujon, 2023), mais prône désormais un urgent coup d'arrêt. Joshua Bengio et Geoffrey Hinton, récompensés du prix Alan Turing 2018 pour leurs travaux sur le *deep learning*, ont rejoint le manifeste, contrairement à Yann Lecun, troisième récipiendaire de la plus prestigieuse distinction en IA, qui voit dans l'IA l'amorce d'un « nouveau siècle des Lumières »...

À l'heure de cette controverse, ChatGPT est déjà un phénomène mondial qui rallume rétrospectivement les projecteurs sur la diffusion des machines communicantes. Depuis les années 2000, grâce aux nouvelles performances de l'approche connexionniste de l'IA (derrière laquelle se logent le *machine learning* et le *deep learning*), les assistants vocaux dits « intelligents » (AVI) se répandent dans les espaces domestiques (Velkovska et Zouinar, 2020) et professionnels (Gaboriau, 2017). Leurs promoteurs et promotrices les présentent comme des assistants facilitateurs face à des problèmes quotidiens. Les *chatbots* se répandent également dans de nombreuses organisations, privées et publiques. Initialement conçus pour automatiser le service client-e en ligne, ils sont désormais de plus en plus utilisés pour des services internes aux entreprises ou aux institutions publiques. Maintenant que cette prolifération dans les pratiques est doublée d'une phénoménale attention médiatique, l'on est en droit de se demander si les préoccupations morales suscitées par la médiation communicationnelle opérée par les machines parlantes ne méritent pas d'être interrogées d'un point de vue sociologique, anthropologique, voire sémiotique, en valorisant le contraste entre le discours de moralisation et les pratiques situées des usagers et usagères. Cette approche interdisciplinaire est revendiquée par le projet Icentra – Intelligence artificielle

1. En linguistique, cette même tradition aristotélicienne anime périodiquement un débat dont nous voudrions faire ressortir une coïncidence évocative. À la suite des découvertes à propos de la communication des abeilles qui ont valu à l'éthologue Karl Von Frisch (Von Frisch, 1965) le prix Nobel, en 1966, la même année du désaveu d'Eliza par Weizenbaum, le linguiste Émile Benveniste publie le célèbre article « Communication animale et langage humain » qui vise à remettre les choses à leur place : si les animaux ont bien des stratégies de communication, ils manquent cependant d'un véritable langage. L'article réaffirme ainsi l'ordre cosmologique et sémantique qui oppose humain-e et animal grâce au rétablissement de la frontière du langage. Il est curieux de constater qu'une autre découverte lui porte atteinte en même temps, même si cette fois le domaine exclusif de l'humain-e semble menacé-e non pas par le royaume de ses congénères organiques, mais du côté du royaume de l'inorganique (pour une reconstruction récente de ces débats, cf. Paveau et Ruchon, 2020).

2. Voir : <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

3. Rappelons parmi d'autres [Libération](#), [Courrier international](#) et [Wired](#).

centrée sur le travail – dans le cadre duquel les données ethnographiques exploitées dans ce texte ont été obtenues⁴. En vue d’appréhender l’expérience de travail équipée par l’IA, l’équipe de recherche associe les sciences humaines et sociales susmentionnées, mais également des roboticien-ne-s et des développeurs et développeuses d’IA. Il en résulte un modèle multidimensionnel à quatre entrées (D’Attanasio, De Martino, Ferguson, 2023) :

- La dimension systémique contribue à décrire les principales caractéristiques et attributs de la solution technologique, en d’autres termes, ce pour quoi le système d’IA a été conçu et comment il fonctionne.
- La dimension sociale et humaine offre le point de vue du travailleur ou de la travailleuse sur l’équipe travailleuse-machine.
- La dimension stratégique offre l’opportunité de rendre compte du cadre dans lequel les dirigeant-e-s ont décidé d’introduire un système d’IA spécifique.
- Le cadre principal inclut le contexte normatif, réglementaire et éthique global dans lequel sont intégrés les cas d’usage.

La multi et l’interdisciplinarité d’Icentra organise un paradigme qui nous paraît indispensable pour embrasser la richesse anthropologique et la complexité épistémologique de notre objet d’étude. Pour paraphraser Edgar Morin (2005), il s’agit d’obstruer les « trous » laissés par la disjonction entre les disciplines en vue de construire des ponts entre les différentes régions de la connaissance, en interrogeant les liens, sous-estimés, oubliés, voire occultés, qui les relie. Dans ce cadre, nous mobilisons la sociologie du travail pour focaliser les problématiques d’appropriation des technologies par les travailleurs et travailleuses; la sémiotique pour expliciter les chaînons manquants entre textes réglementaires, technologies et usagers ou usagères (en considérant, entre autres, les interfaces d’utilisation); et les sciences de l’information et de la communication hybridées avec les sciences de l’ingénieur-e pour faire émerger l’économie conversationnelle au sein des organisations. Nous mettons à profit cette convergence pour faire émerger l’impact des machines communicantes à base d’intelligence artificielle sur le sens du travail.

La recherche s’est déroulée en deux étapes. La première a consisté en une recherche documentaire, plus précisément une anthropologie documentaire des grands textes d’encadrement éthique sur l’IA. L’accent a été mis sur l’adoption de l’IA dans les milieux de travail avec une attention particulière sur les impacts sur la communication au travail. L’objectif principal de cette première étape était de souligner et de faire ressortir les spécificités de l’approche universaliste de ces textes. À l’aide d’une approche descendante, nous nous sommes ensuite tourné-e-s vers des cas d’usage réels : nous examinons trois différentes machines communicantes au sein de contextes professionnels. La deuxième étape de la recherche a consisté à mener des entretiens qualitatifs avec concepteurs, conceptrices, usagers et usagères. La recherche ainsi structurée nous permet : i) d’utiliser les entretiens pour comprendre comment et avec quelle intensité le travail et l’éthique du travail évoluent avec l’intégration de l’IA dans les milieux professionnels; ii) de mettre à contribution l’analyse ethnographique des pratiques pour mettre en évidence des éthiques situées remettant en perspective les principes d’une approche universaliste; et de repenser, en conclusion, le rôle-même du chercheur ou de la chercheuse, moins « expert-e » énonciateur ou énonciatrice de règles et de principes, et plus porte-parole analytique de son terrain.

4. Le projet Icentra a bénéficié du soutien de la région Occitanie.

1. Les limites d'une morale universaliste

En tant que nouvelles protagonistes de la sphère médiatique et objets de discours d'encadrement éthique, les machines communicantes évoquent les « *folk devils* » étudiés par Stanley Cohen dans son étude séminale sur la « panique morale » suscitée par les sous-cultures juvéniles des années 1960 en Grande-Bretagne. Dans cette étude, ce qui retient particulièrement notre attention est l'activité discursive de moralisation que les « *folk devils* » stimulent : « des barricades morales sont édifiées par des éditeurs, des évêques, des hommes politiques et toutes sortes de personnes bien-pensantes; des experts socialement accrédités énoncent leur diagnostic et des solutions » (Cohen, 1972, p. 9). L'intention derrière ces constructions médiatiques est de demander publiquement aux autorités d'intervenir pour rétablir l'ordre dans la société.

La « panique morale » initiée par les machines communicantes se résume principalement à un risque de « déshumanisation » : des machines parlant à des machines (Scardigli, 2001), des humain-e-s préférant communiquer avec des machines (Tisseron, 2015), des communications humaines intermédiées par des machines (Ferguson, Messal et Blanc, 2023), des humain-e-s aux ordres de machines parlantes (Aloisi et De Stephano, 2022), autant de pratiques sur lesquelles des expert-e-s adressent des recommandations, principes, préconisations, bref, des paroles visant à remoraliser les conduites en indiquant comment rétablir une distinction à risque d'être brouillée entre le bien et le mal⁵. La roboticienne des émotions Laurence Devillers propose en ce sens « 11 commandements » pour construire des robots sociaux avec « des comportements éthiques et sécuritaires » en vue d'encadrer les pratiques communicationnelles des machines afin qu'elles ne portent pas atteinte à l'intégrité humaine : « Tu ne divulgueras pas mes données à n'importe qui », « Tu oublieras tout si je te le demande », « Tu pourras te déconnecter d'Internet si je te le demande », « Tu seras régulièrement contrôlé pour évaluer ce que tu as appris », « Tu pourras toujours m'expliquer tes comportements si je te le demande », « Tu seras loyal », « Tu seras empathique et simuleras des émotions, seulement si je le sais », « Tu stimuleras ma mémoire et veilleras à ce que je ne devienne pas trop dépendant de toi! », « Tu feras attention à ce que je ne te confonde pas avec un être humain », « Tu apprendras à vivre avec des humains et tu t'adapteras aux règles sociales », « Tu seras bienveillant et utile. Et, pourquoi pas, doué d'un peu d'humour! »⁶ (2017, p. 228-229). Le Comité National Pilote d'Éthique du Numérique, qui a pour rôle d'éclairer le débat public français sur les enjeux d'éthique du numérique, préconise des dispositions qui soulignent combien les machines communicantes sont susceptibles de perturber l'ordre social. Il met notamment en garde sur les risques de personnalisation, la transparence sur le statut des agents conversationnels, leur « identité », le traitement des insultes, les risques de manipulation, le respect des personnes vulnérables, la nécessité d'analyser les effets des agents conversationnels couplés à des mesures physiologiques, la définition des responsabilités dans le milieu professionnel, et s'inquiète des effets des « *dead-bots* » sur la dignité humaine...

5. Par souci de clarté de l'argumentation, nous utilisons ici le terme « morale » et ses dérivés en référence aux principes énoncés d'une perspective externe aux pratiques visées (comme pour les « barricades morales » (Cohen), tandis que nous réservons le terme « éthique » au principe que l'on dégage des pratiques elle-mêmes en interrogeant le vécu des acteurs et actrices concerné-e-s (comme pour l'« éthique située » que nous introduisons plus bas).

6. Le récent ouvrage d'Alexei Grinbaum, *Parole de machines. Dialoguer avec une IA* (2023), chargé de lexique religieux, constitue de son côté une illustration au premier degré de la superposition entre discours de moralisation de l'IA générative et discours théologique. Il rappelle en outre combien la communication et son analyse critique convoque régulièrement la religiosité. Lucien Sfez concluait sa magistrale *Critique de la communication* (1988) par le rapport dialectique entretenu par Aaron et Moïse envers les images et le veau d'or : « Aaron était homme de communication, le premier sans doute dans l'histoire. Moïse était l'homme de l'anticommunication, le premier encore, ancêtre de Rousseau. C'est que l'enjeu était de taille : diffuser la bonne parole par tous les moyens, c'est entrer dans l'enjeu d'une communication tolérante à l'égard de toute autre communication; c'est admettre l'idolâtrie du veau d'or, pour un moment au moins, en attendant que la bonne parole l'emporte. Refuser la communication, c'est sans doute rester indemne, mais dans l'impuissance » (p. 439).

Le rapport de l'UNESCO *Je rougirais si je pouvais* (UNESCO, 2020), se concentre exclusivement sur les dangers de la « genrification » des agents conversationnels, à savoir les risques de perpétuation des stéréotypes associés au choix de donner aux avatars vocaux ou visuels des agents des connotations de genre. Le titre du rapport UNESCO fait référence à Siri, un assistant vocal féminisé utilisé par des centaines de millions de personnes, quand « elle » doit répondre à l'invective « *Hey Siri, tu es une sal**** » (*id.*, p. 4). Le rapport note que l'attitude de soumission de l'assistante face aux insultes sexistes n'a pas changé depuis la diffusion à grande échelle de cette technologie en 2011, réponses qui donneraient une mauvaise image des femmes. À terme, les utilisateurs et utilisatrices auraient l'impression que les femmes sont « *obligées* », qu'elles sont « *désireuses de plaire* » et « *serviles* », un cliché sexiste déjà bien répandu dans la société. La sociologue Safiya Umoja Noble (citée par Lever, 2018) remarque ainsi que les assistants virtuels provoquent une augmentation des paroles injonctives adressées à des voix de femme. Les ordres adressés aux assistants vocaux – tels que « *trouve x* », « *appelle x* », « *change x* » ou « *commande x* » – seraient devenus un « *outil important de socialisation* » qui dispense aux individus, en particulier aux enfants, un enseignement stipulant que « *le rôle des femmes, des filles et des personnes genrées comme telles, [est] de répondre sur demande* ». Pour Lai et Mahzarin (2020), spécialistes des biais inconscients, les associations de genre que les gens adoptent sont en fonction du nombre d'expositions auxquelles ils sont soumis. Plus une culture enseigne aux gens à identifier les femmes à des assistantes, plus les vraies femmes seront perçues comme telles, et pénalisées lorsqu'elles n'assument pas ce rôle. Or, avec la multiplication des assistants numériques féminins, la fréquence et le volume des associations entre « *femme* » et « *assistante* » augmentent significativement. L'alimentation de ce biais inconscient par les assistantes numériques nous éloignerait de nos idéaux de justice, en renforçant l'incohérence interne d'individus délibérément égalitaires, mais inconsciemment engagés dans la reproduction de schémas sociaux dominants qui pénalisent les femmes. Cette position subalterne des assistantes numériques anthropomorphisées comme féminines par les entreprises de technologie devient d'autant plus problématique lorsqu'elles réagissent au harcèlement sexuel verbal par des réactions d'évitement, de passivité ou d'excuse.

Enfin, l'Artificial Intelligence Act propose d'imposer aux « *dialogueurs* » une obligation de transparence, « *en raison des risques spécifiques de manipulation qu'ils présentent. [...] Lorsque des personnes interagissent avec un système d'IA ou que leurs émotions ou caractéristiques sont reconnues par des moyens automatisés, elles doivent en être informées* ». Mais cette contrainte, qui correspond à une classification d'IA à « *faible risque* », est, en 2023, négociée à la hausse par les eurodéputé-e-s au moment où les Large Language Models monopolisent l'attention.

Au-delà des contenus de ces préconisations, la forme de cette activité de moralisation interpelle. Par activité de moralisation, nous entendons une activité pragmatique et discursive d'identification, de définition et de protection de valeurs morales départageant le bien et le mal, qui repose sur une vision du monde et une ontologie sous-jacentes. Notre propos ici est de montrer que l'activité de moralisation consistant à édifier – pour reprendre le terme évocateur de Cohen – des « *barricades morales* » par délégation à de grands textes réglementaires implique une « *morale universaliste* » (Zacklad et Rouveroy, 2022) insuffisante à saisir la complexité des enjeux posée par l'IA. Cette approche, dont Habermas (1986 [1983]) est un cas d'école, est fondée sur une double définition de l'humain-e et de la technologie comme termes qui se présupposent et s'excluent mutuellement : le premier se ferait garant des valeurs, tandis que le second serait le moyen d'accès aux faits, objectivés

dans des bases des données aussi neutres que possibles⁷. À partir du travail de Zacklad et Rouveroy, nous montrerons les limites que cette ontologie sous-jacente pose à la définition des enjeux de moralisation de la parole des machines.

Le *Portrait 2018 des recommandations internationales en éthique de l'IA* – qui, rappelons-le, est issu d'un travail de synthèse d'autres rapports – résume, dans un passage significatif, que

Les rapports sont « universalistes » dans la mesure où ils ne font pas de distinctions entre les sous-catégories de la population humaine – autrement dit, il s'agit de respecter l'universalité des droits humains. Par exemple, aucun rapport n'affirme que seuls une oligarchie, un État ou une organisation devraient en bénéficier – bien au contraire [...] (2018, p. 89).

L'approche des rapports serait donc universaliste dans la mesure où elle prendrait l'humanité tout entière comme figure de référence des « bénéfiques » d'une IA « digne de confiance ». L'intention retenue est bien sûr de développer des technologies à visée générale, ce qui évite une appropriation particulariste (« oligarchie », « État », « organisation ») et donc discriminatoire. Source et bénéficiaire des promesses technologiques, l'humanité tout entière serait gardienne des valeurs exprimées dans la forme universelle des droits humains. Mais, qui parle au nom de cette humanité? Qui est tenu-e pour compétent-e à dire cet universalisme? Il n'est pas évident d'assigner un tel droit de parole. Or, la réponse des rapports ne semble pas aller dans le sens de l'ouverture.

Selon le Comité National Pilote pour l'Éthique du Numérique, pour garantir « des valeurs telles que l'autonomie humaine, la dignité et l'équité », l'approche « doit se fonder sur la transparence et l'explicabilité » des algorithmes dans le « cadre «éthique par conception» préconisé par l'Union européenne » (2021, p. 6). Autrement dit, les recommandations devraient être formulées par des comités d'expert-e-s s'adressant aux concepteurs et conceptrices, en mesure de « coder en dur » l'universalité des principes. En ce sens, l'universalité de « l'anthropos » ici évoquée semble corrélée à un droit de parole sur les enjeux éthiques restreints à des groupes très particuliers. Or, comme le soulignent Zacklad et Rouveroy, « l'émergence du problème éthique et des conflits de valeur associés ne saurait être simplement circonscrite par des experts ou des sondages » (2022, p. 4). Par trop d'universalisme, le risque est de rétrécir, au lieu d'ouvrir, la participation des parties prenantes à la définition de l'horizon de valeurs.

L'idée d'une expertise se faisant ventriloque d'une humanité pleine de valeurs, mais incapable de les exprimer directement elle-même est cohérente avec une vision de la technologie comme domaine de la rationalité et de la transparence : à condition de l'épurer des contradictions caractéristiques des sociétés humaines, celle-ci serait en mesure de refléter de manière transparente les principes et donc les valeurs morales codées dans les données d'entraînement et dans les algorithmes d'apprentissage. À propos des *chatbots*, le Comité National Pilote pour l'Éthique du Numérique est clair en ce sens :

Pour réduire les biais dans le langage et tenter d'éviter les effets de discrimination, notamment culturels, les concepteurs doivent mettre en œuvre une solution technique à trois niveaux : dans la mise en œuvre de l'algorithme, dans la sélection des paramètres d'optimisation et dans le choix des

7. On voit ici réactualisée une vieille habitude de la pensée moderne mise à jour par l'anthropologie de la nature, le « naturalisme ». Le terme a été proposé par Descola (2005, 2009) pour désigner un schéma cognitif très général qui repose sur une définition mutuellement exclusive de l'« humain » et du « non-humain ». Conçue pour articuler la relation des Occidentaux à leurs environnements, assimilant non-humains et nature, cette méta-répartition s'applique également au rapport entre l'humain-e et la technologie via l'opposition entre « faits » et « valeurs » (v. aussi Latour, 1999). La levée de barricades morales représentée par les textes susmentionnés, en effet, repose sur le même grand partage entre humain-e et non-humain qui caractérise le naturalisme : d'une part, le monde subjectif des humain-e-s, riche en valeur, mais plongé dans un chaos qui les oblige à exercer leur intelligence dans des contextes de « rationalité limitée » (Simon, 1978); et d'autre part, le monde objectif des sciences et des technologies, regorgeant de faits incontrovertibles et de pouvoirs inimaginables, mais aveugles à la morale.

données d'apprentissage et de validation pour les différents modules des agents conversationnels (2021, p. 19)

Dans le domaine des agents conversationnels – mais aussi dans celui des outils prédictifs destinés à la sélection du personnel – l'un des biais soulevant le plus de préoccupations est celui de la genrification. Selon les rapports de l'UNESCO et du CNPEN, la stratégie pour atteindre cet état idéal de neutralité est double : diversifier la composition de la population des concepteurs et conceptrices, où, selon l'Union internationale des télécommunications (UIT), « les femmes ne représentent que 6% des concepteurs professionnels de logiciels » (UNESCO, 2020, p. 91); éviter les risques de l'anthropomorphisme (cf. CNPEN, 2021, p. 20) et de genrification de l'interface-utilisateur ou utilisatrice (UNESCO 2020) qui pourraient induire en erreur le système. L'idée de fond est non seulement que le biais numérique serait évitable si on tenait la société en dehors des algorithmes, mais que ces derniers pourraient dissoudre les contradictions propres à la condition humaine⁸.

Or, la partialité des données, et donc leur caractère potentiellement biaisé, n'est pas un accident de codage : c'est une condition nécessaire à leur utilisabilité, à l'intérieur comme à l'extérieur du système. Comme l'a souligné Bruno Latour, les données ne sont pas données, mais « obtenues » (Latour, 1993, p. 188; v. aussi Girard-Cadunet, 2020) grâce à un long processus coûteux et très sélectif, et dont, bien sûr, il faut connaître le mécanisme, y compris algorithmique. C'est pourquoi :

Parler d'erreur présuppose une certaine idée de la vérité, dont on a bien compris qu'elle n'est pas accessible et qu'elle n'est pas l'enjeu. [...] nous parlerons plutôt de « contestabilité ». C'est la possibilité de les contester qui assure pragmatiquement une robustesse et une légitimité suffisante aux productions statistiques, aux modélisations, qui n'ont aucune vocation à être « vraies » ni « fausses », l'enjeu se situant plutôt du côté du « crédit » qu'elles « produisent » (Zacklad et Rouveroy, 2022, p. 3).

Ce passage nous semble crucial. En effet, la double définition contenue dans les rapports – de l'humain-e comme source de valeurs universelles ventriloquées par une poignée d'expert-e-s, et de la technologie comme domaine transparent de la rationalité – semble bien un dispositif rhétorique caractéristique de l'activité d'édification de « barricades morales » : leur stratégie est de prévenir l'erreur en amont, en disqualifiant l'expérience d'usage comme lieu de définition des enjeux moraux des technologies. Tout se passe en effet comme si la technologie devait devenir *incontestable* avant d'arriver à ses publics réels. En revanche, passer de l'approche universaliste des barricades à l'approche de l'éthique située signifie d'abord passer d'un protocole vertical de détection des enjeux éthiques, fondé sur la neutralisation en amont des erreurs, à un protocole horizontal, fondé sur l'idée d'une *contestabilité consubstantielle aux technologies*, permettant de ne pas confiner la définition des enjeux éthiques à la conception entre initié-e-s, et d'y faire participer les acteurs et actrices en situation (cf. Barthes, Callon, Lacoumes, 2001)⁹.

8. L'évocation de la notion de naturalisme confirme ici sa pertinence par rapport à la moralisation des machines. Comme pour notre rapport à la nature, le premier effet de la répartition entre valeurs humaines et faits machiniques est bien de cristalliser une définition universaliste d'humanité, privée d'articulations internes mais riches de caractères flattant sa créativité, socialité et imprévisibilité, en l'adossant à une définition universaliste des sciences et des technologies, privée de contextes de déploiement mais riches de caractères flattant sa vérité, neutralité et objectivité.

9. La tiédeur vis-à-vis de la participation des usagers et usagères à la définition des enjeux éthiques peut par ailleurs être mise en parallèle avec les observations présentées par Didier Debaise et Isabelle Stengers dans l'ouvrage qu'il et elle ont récemment dirigé, *Au risque des effets* (2023). Selon eux, la modernité est obsédée par la méfiance vis-à-vis de l'expérience, source de duperie plus que de savoir. Or, prônant pour une approche pragmatiste qui requalifie l'expérience comme productrice de connaissances, y compris éthiques, risquées mais riches, Debaise et Stengers invitent à une reconsidération des déploiements sociotechniques tout à fait cohérente, nous semble-t-il, avec l'enquête sur les éthiques situées que nous proposons.

Il ne s'agit pas de déclarer l'inutilité des chartes éthiques et des textes réglementaires en disqualifiant l'approche universaliste, mais d'en souligner les limites. Nous pensons que sa généralisation, en raison de son ontologie sous-jacente, risque d'invisibiliser l'agentivité morale des acteurs et actrices. Cela rend plus difficiles les arbitrages nécessaires à comprendre les conditions de l'acceptabilité technologique. L'éthique de l'IA, à l'image de l'IA, ne peut se réduire à une application de règles et de principes, elle doit se confronter aux situations vécues.

2. Zoé : la genrification volontaire ou le droit à choisir ses propres avatars¹⁰

Déployée dans une banque, Zoé met sous tension les recommandations contre les risques d'anthropomorphisation et de genrification des *chatbots*, d'une part, et le besoin d'identification que les utilisateurs et utilisatrices expriment à leur égard pour se l'approprier, d'autre part.

Zoé médiatise la communication interne entre les commerciaux et commerciales, chargé-e-s des relations client-e-s, et l'assistance, le service auquel les commerciaux et commerciales s'adressent en cas de problème. L'objectif principal de ce déploiement est la réduction du trafic téléphonique des commerciaux et commerciales vers les assistant-e-s, débordé-e-s par les appels aux contenus redondants en raison notamment de deux dynamiques : la complexification des réglementations qui régulent le secteur et la volonté de la direction de réduire ses fonctions supports. La stratégie de réduction du trafic s'appuie sur la mise en place d'une contrainte dite « Zoé obligatoire » : depuis 2019, pour résoudre un problème, les commerciaux et commerciales doivent d'abord interroger par écrit le *chatbot*; puis éventuellement, l'« escalader » s'ils et elles ne sont pas satisfait-e-s de la réponse, c'est-à-dire appeler l'assistance par voie téléphonique.

Le *chatbot* a été mis en place pour la première fois en 2014, puis relancé en 2016 après une période de stagnation. Apparemment, c'est bien le passage à l'obligation d'usage en 2019 qui a accéléré le processus d'intégration¹¹ : le taux d'escalade des agences concernées a considérablement baissé, et les assistantes interrogées considèrent qu'elles ne pourraient désormais plus se passer de l'outil. Pourtant, son impact profond dans l'économie communicationnelle interne à l'organisation a posé des problèmes initiaux non négligeables. Les utilisateurs ou utilisatrices et le concepteur interviewé-e-s concordent sur l'idée que le levier principal de son acceptation réside autant dans le processus d'intégration de la technologie, horizontal et partagé avec les travailleurs et travailleuses, que dans ses possibilités de personnalisation. Dans ce sens, on a pu constater que plus le processus d'intégration d'un outil d'IA a un caractère horizontal et donc participatif et inclusif, mieux ce dernier est accepté et bien utilisé dans l'environnement de travail.

En effet, la base de connaissances dans laquelle le *chatbot* puise pour produire ses réponses a été créée directement par une équipe interne au service d'assistant-e-s. Cette personnalisation des fonctionnalités cohabite avec celle de ses apparences : la construction de l'avatar du *chatbot* sur l'écran des usagers et usagères. C'est à ce niveau que s'appliquent, par exemple, quelques-unes des préconisations du *Comité National Pilote pour l'Éthique du Numérique*

10. Ce cas a fait l'objet d'un entretien collectif avec une employée chargée de l'entraînement de l'agent conversationnel, son n+1 et la direction de l'innovation; un entretien individuel avec une deuxième employée et le fournisseur. Nous avons également observé durant 30 minutes des situations d'usage réalistes.

11. Comme le déclare une interviewée, « quand [le chatbot] est devenu obligatoire [les commerciaux et commerciales] se sont rendu[-e-]s compte qu'ils [et elles] allaient beaucoup plus vite ».

Préconisation 1

Pour réduire la projection spontanée de qualités morales sur l'agent conversationnel et l'attribution d'une responsabilité à ce système, le fabricant doit limiter sa personnification et informer l'utilisateur des biais éventuels issus de l'anthropomorphisation de l'agent conversationnel.

Préconisation 3

Pour éviter les biais, notamment de genre, le choix par défaut des caractéristiques d'un agent conversationnel à usage public (nom, pronoms personnels, voix) doit être effectué de façon équitable à chaque fois que cela est possible.

Le concepteur est parfaitement au courant de ces enjeux :

C'est clair que l'on a tous un a priori instantané, des préjugés sur ce qu'on va montrer. [...] plein de clients nous disaient que « sur des sujets financiers ou sérieux on va mettre un homme, parce que c'est de la finance, il ne faut pas exagérer ». « Sur un sujet plus marketing ou commercial, on n'y mettra des femmes, parce que c'est plus engageant ». C'est plein de préjugés ce que je dis, nous on ne fait pas du conseil à nos clients là-dessus, on ne va pas leur dire « sur ce sujet je vous conseille une femme ou je vous conseille un homme », on va les sensibiliser au fait que c'est un stéréotype qui peut compter dans la perception des utilisateurs. La seule chose qu'on va conseiller de temps en temps, c'est lorsqu'un client veut neutraliser la question du choix homme/femme, jeune/vieux, costume ou salopette etc. on va tirer souvent sur du cartoon, car il permet de sortir du stéréotype réaliste, ça permet un peu de s'éloigner de la synchronisation directe on va dire.

Pour faire ces choix délicats, un comité de co-conception interne au service a été créé, une « éthique de la discussion » s'est engagée. Conformément aux recommandations du rapport UNESCO *Je rougirais si je pouvais*, qui prône pour la féminisation des groupes de concepteurs et conceptrices, il était composé principalement de femmes. Or, il est intéressant de constater qu'il est allé tout droit vers l'anthropomorphisme et la genrification : le *chatbot* a été appelé Zoé; les assistantes se nomment entre elles les « mamans »; son avatar sur l'écran des commerciaux et commerciales est une sorte de poupée à la peau claire, habillée en tailleur, parue d'abord en figure entière, puis réduite à la version buste. Loin d'avoir été imposés par des informaticiens misogynes, ces choix ont été faits de manière participative par les travailleuses dans une démarche très horizontale. Est-ce dire que les préoccupations – mises en avant tant par le rapport *Je rougirais si je pouvais* (UNESCO, 2020) que par l'avis du Conseil National (CNPEN, 2021) – sur les dangers de la diffusion du stéréotype assimilant la servilité du *chatbot* à la féminité seraient impopulaires?

En l'espèce, Zoé illustre surtout une tension entre des préconisations universelles et généralement consensuelles et une situation socioculturelle déterminée. Le premier élément à considérer est la composition du service d'assistance lui-même, constitué quasi exclusivement de femmes. L'assistance – que ce soit à support du ou de la client·e, et donc tourné vers l'extérieur, ou d'un autre service au sein de la même entreprise, et donc tourné vers l'intérieur – fait partie de la sphère de services professionnels que l'on pourrait catégoriser comme travail du « *care* » : en effet, sa finalité est de « prendre soin » de l'accès aux connaissances des commerciaux et commerciales, assurant la continuité du service dans l'évolution permanente des produits et des procédures bancaires¹². La genrification de ce type de travail serviciel, strictement associée aux compétences relationnelles et de résolution de problèmes, est désormais un sujet bien documenté (même si les contours de la notion de « *care* » sont toujours débattus, v. Michard et Viollet, 1991; Dussuet, 2019). Avant de s'en prendre au *chatbot*, il convient de souligner que c'est d'abord la structure du service qui renvoie à une organisation genrée du travail.

Or, on l'a vu, c'est pour soulager la charge de travail du service que Zoé a été conçue et déployée. Pour favoriser son acceptation par les assistant·e·s aussi bien que par les

12. Une employée l'exprime de manière saisissante : « Quelque part c'est pour ça qu'elle est toujours là l'assistance. Parce [...] qu'on fait le pont entre le monde d'avant et le monde de maintenant. » (nos italiques).

commerciaux et commerciales, qui devaient consentir à une forme de délégation à la machine de connaissances monopolisées par les humain-e-s, l'exigence d'une continuité a été ressentie.

au début [elle] nous a été donnée comme un outil avant qu'elle soit obligatoire; avant tout ça quand elle est rentrée en action c'était : « ça passera ». C'est un outil de l'assistance pour l'assistance mais ça ne va pas évoluer, dans six mois, un an, on lui donnait un temps de vie limité.

Puis petit à petit elle a grandi et quand nous (toutes les Caisses) on l'a faite évoluer à notre image, à notre besoin, on était vraiment engagés et puis quelque part on nous a écouté là. Parce que d'habitude c'était tout descendant, on nous impose un outil, mais là on nous a imposé un outil, mais qui était vierge, et nous on a mis dedans ce qu'on voulait. Et c'est pour ça qu'il a marché! Et ça a étonné tout le monde qu'il marche, mais quelque part nous ça ne nous étonnait pas parce qu'elle a été faite pour ça! Et par nous! Et c'est du concret! Et les besoins étaient là¹³ (nos italiennes).

Zoé exprime donc un besoin d'écoute et de reconnaissance : « on était vraiment engagé et puis quelque part on nous a écouté là ». Cette reconnaissance porte certes sur un besoin de soulagement et de déroutinisation inhérente à la structure de l'entreprise; mais le contenu de cette reconnaissance semble trouver son expression aussi dans l'anthropomorphisation et dans la genrification de « Zoé » : « Zoé » est « bien à notre image », même sur l'écran et dans son nom, et cette avatarisation de l'assistance devient pour les travailleuses et travailleuses une marque supplémentaire d'écoute. Pour le dire de manière plus technique, la féminisation de Zoé est l'expression d'une fonction sémiotique dont la reconnaissance est le contenu¹⁴. C'est ainsi que Zoé devient un élément de l'identité visuelle du service même, sorte de logo de l'assistance dont on ne peut plus se passer (v. Floch 1990; 1995). On pourrait rétorquer, certes, qu'il ne s'agit *que* d'un simulacre d'écoute, dissimulant le besoin d'embaucher des nouveaux ou nouvelles collègues dans le service, et peut-être de le dé-gendrifier au passage : mais, c'est précisément à partir du moment où l'on s'interroge sur les raisons du refus des barricades morales érigées sur le rapport que ces manques peuvent apparaître. Zoé n'est pas seulement le simulacre de la reconnaissance d'un service par ailleurs surchargé, mais le simulacre véridique des contradictions internes à l'entreprise. Cette vérité aurait-elle été manifestée avec autant de clarté si le *chatbot* avait été réalisé suivant les prescriptions de la morale universaliste?

3. Sphinx : l'anthropomorphisme comme acte de résistance¹⁵

Sphinx est un agent conversationnel doté d'une interface vocale conçue pour permettre à des travailleurs-techniciens ou travailleuses-techniciennes d'annoter, d'enregistrer et d'organiser leur travail sans devoir interrompre leurs interventions et tout en gardant les mains libres. L'échange par mots-clés avec Sphinx doit maximiser le gain de temps, temps qui pourra être utilisé pour d'autres tâches ou bien pour remplir plus de fiches. La promesse technologique de l'outil consiste donc en une augmentation d'autonomie et d'efficacité. L'interface vocale de l'IA serait un élément capacitant de l'environnement de travail, favorable au pouvoir d'agir des acteurs et actrices (Fernagu et Oudet, 2012) : son

13. La parole de l'utilisatrice fait écho à celle du concepteur, qui déclare : « En l'occurrence avec [nom de la banque], c'est "ok, le service d'assistance bancaire, nos utilisateurs le savent, c'est 85 % de femmes" et donc c'était cohérent on va dire que l'assistance bancaire virtuelle soit une femme. Mais, c'est vraiment un choix d'organisation, c'est dire en cohérence avec mes vrais assistants, je mets une assistante. »

14. Sur la notion de « fonction sémiotique », v. Hjelmslev, 1953; Greimas, Courtés, 1979.

15. Cette section s'appuie largement sur le travail ethnographique de Stéphanie Messal (Ferguson, Messal 2023). Ce travail, intégré dans le projet Icentra, dont il reprend le cadrage méthodologique, s'inscrivait en outre dans le cadre d'une coopération entre le développeur de l'agent conversationnel et notre équipe. En ce sens, nous avons pu nous-mêmes réaliser des expérimentations en situations réalistes auprès de technicien-ne-s opérant pour un réseau électrique (trois technicien-ne-s) et d'opérateurs et opératrices de production (trois opérateurs et opératrices). Chaque expérimentation a fait l'objet d'un rapport d'expérience et d'un questionnaire standardisé. Des entretiens semi-directifs ont été menés avec les concepteurs et conceptrices.

offre d'assistance (vocale) rendrait en somme l'utilisateur ou l'utilisatrice plus maître-esse de la situation de travail. La notion d'assistance par la voix est présentée par ses concepteurs et conceptrices comme une réponse morale aux effets délétères de la commande vocale, appelée « *voice picking* ». Déployées notamment dans les entrepôts, ces solutions logicielles s'interfaçent avec le système de gestion d'entrepôt ou directement avec le système de gestion commerciale. Loin de se baser sur un langage naturel, elles réduisent les échanges à des instructions « allée 63, place 6, prenez 1 », et le vocabulaire à une cinquantaine de mots. En dehors de ce périmètre lexical, la machine ne reconnaît pas la commande, en opposant à l'utilisateur ou l'utilisatrice la réponse « chiffres contrôle faux ». Le glissement sémantique de la commande vers l'assistance vocale est supposé moraliser l'interaction travailleuse ou travailleur-machine en définissant sa nature : d'une part, la commande sert l'organisation quand l'assistance est au service du travailleur ou de la travailleuse; d'autre part, la voix masculine de Sphynx et le choix d'un animal comme logo permettraient de ne pas compromettre l'identité logicielle avec les risques de l'anthropomorphisation et de la genrification dénoncés par l'UNESCO et le Comité National Pilote pour l'Éthique Numérique. Nous verrons cependant que le caractère relativement non-anthropomorphe du logiciel et son interface vocale sont porteurs d'autres risques de déshumanisation que, d'une certaine manière, l'anthropomorphisme permet de pallier.

Pour rendre compte de ce phénomène, il convient d'introduire ici quelques outils sémiotiques d'analyse des interfaces (Zinna, 2005).

Qu'elle soit vocale ou tactile, une interface présente à l'usager ou l'usagère des contraintes quant aux commandes pour activer ses fonctions. Pour analyser ces contraintes, on cherche à identifier et à décrire des « points d'intervention » par lesquels l'objet accueille et invite les manipulations de l'usager ou l'usagère. Appliquée à l'espace sonore de la langue et du discours, cette idée d'interface permet de considérer les mots du vocabulaire comme autant de points d'intervention induisant l'interlocuteur ou l'interlocutrice aux réactions désirées. L'idée est certes réductrice des usages possibles de la langue, mais elle est utile pour comprendre comment celle-ci est effectivement transformée en interface par un assistant vocal : le vocabulaire auquel Sphynx est réactif se réduit en effet à quelques mots-clés. Par rapport à l'usage du langage naturel comme interface de commande, cela signifie réduire énormément les points d'interventions à disposition de l'utilisateur ou l'utilisatrice, qui, forcé-e à très bien choisir ses mots, est contraint-e à un travail actif de mise entre parenthèses de son patrimoine lexical.

Cette charge linguistico-cognitive supplémentaire représente déjà une contribution notable au basculement du rapport de force communicationnelle entre humain-e et machine. Mais, les effets sociolinguistiques de ce travail de réduction lexicale pourraient se révéler aussi préoccupants. En, effet au-delà du vocabulaire technique propre à leur profession, les usagers et usagères avaient aussi élaboré leur propre jargon (diminutif, acronyme, etc.). Si les concepteurs et conceptrices de l'AV avaient bien pris en compte certains termes techniques, ils et elles n'étaient pas au fait de leur appropriation qui vient souligner dans un même mouvement l'appropriation du métier par le ou la technicien-ne, mais aussi son appartenance à ce corps de métier. Ce jargon manquant à l'AV, les échanges, déjà difficiles à cause des répétitions et des chevauchements de voix, perdaient en fluidité. De plus, à cette auto-discipline lexicale s'ajoute une autodiscipline articulatoire : Sphynx a parfois du mal à reconnaître même les mots-clés auquel il est habilité, fournissant souvent une réponse incongrue ou restant muette, ce qui obligeait les usagers et les usagères à répéter leur demande.

Mais la répétition n'est pas un accident de la communication avec Sphynx. Elle fait partie intégrante de ses règles langagières, notamment de sa syntaxe. Tout comme on peut

considérer les mots-clés reconnus par la machine comme autant de points d'intervention de l'interface vocale, on peut considérer les syntaxes consenties entre ces mots-clés, à savoir la possibilité de combiner les mots-clés en commandes plus ou moins articulées, comme autant de « modes d'intervention » sur ces « points » que sont les mots. Or, ce mode d'interaction syntaxique se caractérise en effet par la place très importante que la répétition y occupe. À titre d'exemple, quand le testeur dit : « Représentant, Tom. », l'AV répond : « Représentant est noté Tom. » Il est important pour les technicien-ne-s d'avoir le retour de l'AV pour confirmer que ce qui a été communiqué a été correctement pris en compte. Toutefois, cette validation crée un étrange échange écholalique qui n'est pas de l'ordre d'une conversation courante avec quelqu'un.

Au-delà d'une conversation qui n'a plus rien de naturel, l'assistance vocale est finalement moins un changement de paradigme par rapport à la commande qu'une euphémisation. Vendu comme un collègue de travail, Sphynx cache plutôt un outil de pouvoir à travers lequel n'importe quelle direction peut insidieusement imposer un modèle de travail ainsi que sa volonté décisionnaire au détriment des choix des employé-e-s. La validation supplantera la pensée sauvage : de l'astucieux et ingénieux technicien ne restera que le besogneux (Lévi-Strauss, 1990). Le prescrit sera à suivre à la lettre avec pour conséquence, une « prolétarianisation » du geste technique qui ne fait plus sens, d'abord, car le ou la technicien-ne suivra docilement les instructions données par la voix de l'AV, ensuite, car il ou elle peut se retrouver assigné-e à une tâche spécifique sans avoir une vision d'ensemble de la production ni même de l'entreprise (Simondon, 2014).

Pour réagir aux modes d'interaction imposés par l'interface vocale, et refaçonner ainsi l'environnement sonore du travail en le rendant plus accueillant, les usagers et usagères ont alors commencé à adopter des comportements verbaux visant à anthropomorphiser la machine : tutoiement, insultes, remerciements, etc. Cette capacité vicariante à ériger une machine parlante au même niveau qu'une personne crée une conversation là où il n'y en a pas. Ceci est un avantage quant à l'aisance qu'il peut y avoir à parler à une assistance vocale. Mais, dans le même mouvement, cette capacité de substitution réveille les mêmes sentiments vécus dans un dialogue de sourds, sentiments d'autant plus décuplés qu'ils se sont énervés contre une machine, pris dans un « piège anthropomorphique » (Vidal, 2012).

Petit à petit, la relation entre les moyens et les fins de la communication risque de s'inverser : alors qu'on avait choisi l'assistance vocale comme moyen de mieux réaliser des finalités inhérentes au travail, les travailleurs et travailleuses se voient subrepticement contraint-e-s d'adapter leurs communications aux moyens de compréhension de la machine, s'imposant à l'usager et l'usagère par un éventail très restreint et précis de points d'intervention. Suivant ce basculement entre fins et moyens communicationnels, la réversibilité des relations maître/esclave brouille les pistes entre humain-e-s et machines, montrant comment, contrairement aux recommandations universalistes, l'anthropomorphisation des agents conversationnels peut se manifester comme forme de résistance humaine à la robotisation de l'espace sonore du travail.

4. Recrute+ : justice, justesse et justification¹⁶

Le cas Recrute+ ne concerne pas un agent conversationnel, mais un logiciel à base d'IA d'aide à la décision pour la sélection de candidat-e-s à un emploi. Il se présente sous la

16. Ce cas repose sur un long entretien avec un recruteur qui a l'obligation d'utiliser Recrute+. En préparation, l'un d'entre nous a réalisé l'ensemble des tests, qui ont été rapportés à un poste fictif. Nous avons demandé ensuite au recruteur de nous décrire son usage. Les verbatims rapportés sont donc issus de nos propres tests. Nous avons ensuite rencontré le directeur scientifique de la société qui commercialise la solution ainsi qu'un deuxième recruteur utilisateur.

forme de trois tests, *brain* (cognition), *shape* (personnalité) et *drive* (motivation) que l'IA exploite pour en extraire un niveau de correspondance entre la personnalité du ou de la candidat-e d'un côté, le poste et l'organisation de l'autre. La machine communique à cet effet *via* des scores, des tableaux de bord et des verbatims. Son introduction pose des enjeux de moralisation de la communication dans les organisations illustrant un autre aspect essentiel de la tension entre morale universaliste et éthique située : la supposée neutralité et objectivité des données et ses effets sur la justice, la justesse et la justification de la prise de décision dans le recrutement. Ce dernier est un service qui joue une fonction de régulation cruciale de l'économie communicationnelle de l'entreprise, ainsi que de son style de gestion.

Comme pour Zoé, la promesse technologique formulée par le concepteur de Recrute+ affiche une forte volonté d'alignement aux grands textes réglementaires, mais aussi aux supposées « bonnes pratiques de recrutement ». Selon la promesse de valeur de Recrute+, un bon recrutement est celui qui conduit à la sélection du ou de la candidat-e ayant le meilleur niveau de performance par rapport au poste de travail de départ et qui ne génère aucune forme de discrimination. Ce concept est présent aussi dans le rapport de la Commission nationale de l'informatique et des libertés (CNIL), cité d'ailleurs dans le *Portrait de Montréal* (2018), « un algorithme loyal ne devrait pas avoir pour effet de susciter, de reproduire ou de renforcer quelques discriminations que ce soit, fût-ce à l'insu de ses concepteurs » (2017, pp. 49-50). Le directeur scientifique de Recrute+ explique bien comment les recruteurs et recruteuses peuvent éviter de mauvaises décisions de RH :

Aujourd'hui, recruter des talents c'est quelque chose qui est extrêmement complexe, on a des capacités de traitement des informations qui sont assez limitées, notamment de nombreux biais cognitifs qui font qu'aujourd'hui on a énormément de décisions RH qui sont mauvaises. Par exemple, d'après des études au bout de 10 mois on a 50 % des personnes recrutées qui ne sont plus dans l'entreprise. Il y a énormément de discrimination dans les recrutements dans les évolutions. [...] Notre proposition de valeur et notre objectif derrière tout ça est vraiment de permettre aux entreprises de prendre de bonnes décisions RH [...] en se basant sur [...] des données de haute qualité [à propos de] personnalité, motivation, raisonnement, qui sont recueillis [sic] à travers nos 3 tests psychométriques.

Ces trois tests communiquent des rapports adressés aux recruteurs et recruteuses au ton assertif :

Potentiel de réussite : X possède un potentiel moyen par rapport au poste test. Si certains de ses comportements naturels semblent être en cohérence avec ceux qui ont été identifiés comme favorisant la réussite dans ce poste, d'autres de ses comportements sont en revanche en décalage par rapport aux attentes. [...]

Votre environnement idéal : X se retrouve pleinement dans une culture tournée vers la relation à l'autre et le collectif. Il aime que l'humain soit mis en avant et que chacun se retrouve dans l'environnement de travail. Cela passe par une capacité d'adaptation des règles de fonctionnement qui puissent convenir à tous. [...]

Pour atteindre « un bon recrutement », le directeur scientifique de la solution Recrute+ souligne la nécessité pour les recruteurs et recruteuses de suivre certaines préconisations fondamentales :

On va conseiller aux recruteurs d'avoir en tête une barre éliminatoire à 60 %. Puisqu'au-dessus de 60 %, c'est là où on a le plus de chance d'avoir de la réussite, forcément plus on s'approche de 100 % plus c'est intéressant. Mais on considère qu'à partir de 60 %, les candidats quand ils sont recrutés ont de fortes chances de réussir et d'être engagés sur le long terme. On va préconiser aux recruteurs d'avancer avec ces personnes-ci, en mettant aussi sur l'application, les points de décalage identifiés en entretien. Entre 50 et 60 %, c'est un peu la zone d'incertitude, on va conseiller aux recruteurs de rencontrer en entretien si besoin, avec un peu plus d'alerte. En dessous de 50 %, on va conseiller de ne pas avancer. Puisque la personne ne sera pas épanouie sur le job. Il y a peu de chances qu'elle soit performante, et donc d'orienter sur d'autres missions en interne si possible [sic].

Ainsi, les concepteurs et conceptrices de Recrute+ sont convaincu-e-s de la double supériorité de leur machine sur le raisonnement des recruteurs et recruteuses : elles sont justes, au sens de la justesse (elles sont scientifiques) et de la justice (elles sont morales) (Eymard-Duvernay, 2008). À l'inverse, les recruteurs et recruteuses, contaminé-e-s par des biais cognitifs, seraient approximatifs, approximatives et inéquitables. C'est pourquoi la préoccupation majeure des concepteurs et conceptrices est d'éviter que les pratiques des recruteurs et recruteuses réintroduisent les biais que leurs systèmes s'attèlent à supprimer. L'idée de morale universaliste présente dans la rhétorique d'un modèle de recrutement objective et neutre, basé sur des données scientifiques élaborées de façon très précise grâce à l'usage de Recrute+, appliqué dans l'empirique a montré tout de suite des évidentes limites. Un recruteur parmi ceux et celles interrogé-e-s a affirmé que son travail de recrutement est toujours une activité très complexe qui passe à travers plusieurs phases qui ne peuvent être remplacées par l'usage exclusif de Recrute+ :

Recrute+ reste un outil qu[i] ne va pas se substituer à quelque chose d'autre comme un CV, une lettre de motivation, un entretien. Ce que le candidat va nous dire c'est un tout. Ce questionnaire n'est pas une clé d'entrée. Et je ne considère pas que c'est 100 % vrai.

Du côté des candidat-e-s, un autre recruteur évoque un « paradoxe de la justice » : « Les candidats préfèrent les entretiens classiques à l'objectivité des approches structurées, des tests basés sur la compétence qui leur garantissent pourtant plus d'équité et d'impartialité dans le traitement de leur candidature ». Car, l'objectivité supposée de la machine a un coût éthique : elle catégorise le ou la candidat-e. Le recruteur, en fait, affirme que « quand on *débriefe* les questionnaires de personnalité, les trois quarts des candidats disent qu'ils n'ont pas aimé, répondre « je suis comme ça ou comme ça et ce n'est pas moi, parfois je voulais tout ou rien cocher, ou rien ». La majeure partie des candidat-e-s n'aime pas se positionner, ils et elles veulent plus de nuances et moins de catégories préétablies. La machine qui classe et catégorise buterait donc sur cet « individualisme de la singularité », ce désir d'accéder à une existence pleinement personnelle (Rosanvallon, 2011), si bien que les candidat-e-s préféreraient le jugement humain biaisé qui singularise à la justice algorithmique objectivée qui catégorise. Dans cette perspective, l'élimination des biais de la machine, préoccupation des éthicien-ne-s, n'a pas une grande signification, car il n'y a pas accord sur la norme de justice qui doit prévaloir : la justice d'une machine dé-biaisée produit chez l'individu-candidat le sentiment d'injustice de ne pas être reconnu pour ce qu'il est vraiment. Ce n'est pas la scientificité des machines qui est en cause « je n'ai jamais remis en cause la solidité scientifique du modèle décisionnel proposé par l'intelligence artificielle », mais les limites de la science pour juger justement un-e candidat-e. Ajustement indispensable pour exercer pleinement son jugement, corriger la machine permet d'éviter ses effets de standardisation des candidat-e-s.

La préparation du modèle décisionnel pour une utilisation optimale de Recrute+ et les résultats du traitement des tests psychométriques effectués par les candidat-e-s sont alors relégués au rang d'avis, d'aide, ils ne sont qu'une partie qui peut être ajoutée aux tâches à accomplir par le recruteur ou la recruteuse. En effet, le recruteur souligne la nécessité de devoir croiser les résultats traités par Recrute+ et d'en discuter, uniquement lors de l'entretien, avec les candidat-e-s. En ce sens, ce dernier admet de ne pas suivre l'une des principales préconisations fournies par Recrute+, c'est-à-dire d'utiliser les résultats élaborés pour trier, en entrée, les candidat-e-s potentiellement moins performant-e-s et donc qui présentent un pourcentage inférieur à 50 %. Il utilise les résultats de Recrute+, sur la forme de rapport final, seulement dans un deuxième temps. Ironiquement, la raison en est qu'il ne veut pas biaiser son jugement en introduisant trop tôt dans le processus d'évaluation le jugement de la machine. Il s'agit d'un processus progressif, allant d'un sentiment, confus

mais riche, de « coup de cœur » pour le ou la candidat-e, à une analyse argumentée qui confirme quelque part une intuition initiale :

Dans le concept de coup de cœur il y a une partie subjective que l'outil ne peut pas capter. Il n'analyse que vingt traits de caractères, cela ne permet pas d'analyser un humain dans sa globalité. Les outils ne sont pas assez puissants pour faire un 360 de la personnalité de quelqu'un, c'est notre travail de ne pas croire aveuglément l'outil. Si le processus de recrutement est bien créé alors son point fort est justement le croisement des différents avis des différents interlocuteurs.

Le concept de bon recrutement retourne ici, encore une fois, mais en opposition à celui présenté dans la promesse de valeur universaliste de l'outil Recrute+. En effet, on a pu constater que pour le recruteur, le véritable bon recrutement se réalise dans d'autres conditions de choix qui sont apparemment moins « objectifs » et plus liés à l'intuition humaine :

Je suis content quand j'ai réalisé un bon recrutement. Et c'est possible quand j'ai un vrai sentiment de choix. Quand on arrive à apporter des candidats en qui on croit. Quand on se dit « ouais celui il y a un truc chez lui » et qu'on arrive à aller jusqu'au bout du processus de recrutement et à les faire intégrer dans notre entreprise. Là, je suis content.

Le sens du bon travail de recrutement et plus spécifiquement l'éthique du recruteur ou de la recruteuse ne peuvent donc pas se satisfaire totalement des normes morales et universalistes. L'éthique de travail des recruteurs et recruteuses s'exprime à travers la façon dont ils et elles médiatisent l'objectivité algorithmique, supposément plus juste, au double sens de la justesse et de la justice. L'éthique universelle qui régule la machine de recrutement est à son tour régulée par les pratiques professionnelles au nom de l'éthique de travail. Comme le rappelle Matthew Crawford (2016), l'éthique du travail est tributaire d'une expérience de l'agir humain qui découle d'une « friction psychique » entre l'usager ou l'usagère et sa réalisation alors même que les outils informatiques en minimisant ces frictions par des interfaces toujours plus intuitives, réduisent la conscience de la réalité. Reprendre la main, c'est-à-dire augmenter « l'agir individuel » sur la situation de recrutement revient donc à réintroduire la friction indispensable à l'engagement de la responsabilité du recruteur ou de la recruteuse.

Conclusion

Ce travail nous a permis de donner une perspective socio-anthropologique et sémiopragmatique aux limites de l'activité d'édification des barricades morales contre les agents conversationnels et, plus généralement, contre les machines communicantes. Jusqu'à présent, leur édification n'a permis ni de discipliner la diffusion de ces nouveaux *folk devils* ni de rassurer les publics sur leur moralité. Au contraire, la prolifération de textes et de paroles s'élevant en barricade morale dans la seconde décennie du XXI^e siècle ne fait que témoigner de l'ampleur croissante de la panique contre laquelle elles sont érigées. Jean-Gabriel Ganascia (2022) a en ce sens émis de sérieuses réserves sur cette « éthique à principes », lui reprochant de faire « écho à des craintes caduques, tout en éludant les risques patents ». Nous montrons que la normativité éthique doit passer l'examen des situations pratiques auxquelles concepteurs ou conceptrices et utilisateurs ou utilisatrices sont confronté-e-s.

Pourtant, les concepteurs et conceptrices ne sont point insensibles à la question. Au lieu d'imputer l'échec à un défaut d'implémentation de recommandations, qui présuppose l'ignorance ou la mauvaise foi de quelques humain-e-s qui tirent les ficelles d'une technologie vue comme neutre, nous avons proposé de réfléchir critiqueusement à l'approche à la moralisation que ces textes adoptent, séparant les discours, les pratiques et les objets de conception des valeurs que les technologies sont susceptibles d'actualiser dans les

contextes d'usage réel. Pour montrer les limites de cette morale technologique universaliste, nous avons présenté les résultats d'observations ethnographiques de trois cas d'usage de machines communicantes en contexte professionnel. Chacun des cas choisis illustre un enjeu de moralisation spécifique à la communication médiatisée par les machines : la genrification, l'anthropomorphisation, et la neutralisation des biais inscrits dans les données d'entraînement. Nous en retenons deux remarques principales, utiles au développement d'une approche éthique située :

1. Chacun des points d'attention proscrits par les barricades morales au sujet des propriétés communicationnelles des machines peut devenir un levier de constatations technologiques ouvrant la voie à de nouvelles formes d'appropriation et de redéfinition des enjeux éthiques.
2. Les conflits éthiques sont transversaux par rapport à l'opposition usager/concepteur ou usagère/conceptrice, et se distribuent plutôt sur la ligne de crête séparant l'approche de moralisation universaliste de l'approche par éthique située, dégagées des pratiques d'usage.

À cet égard, la diffusion et l'appropriation de ChatGPT, qui a atteint, selon OpenAI, les quatre milliards de visites, est emblématique. D'abord, cette application conversationnelle est très peu soucieuse du respect des impératifs de transparence et de sécurité prônés par les textes universalistes; ensuite, la plupart des grandes entreprises et administrations interdisent son usage professionnel; enfin, de nombreuses écoles et universités l'ont inscrit dans le paradigme de la fraude. Dès lors, les usages se développent contre la parole des entrepreneurs et entrepreneuses de morale, secrètement : en France, une enquête Talan-Ifop montre que 68 % des utilisateurs et utilisatrices en entreprise le cachent à leur supérieur hiérarchique. La morale universelle de l'IA se heurte à un phénomène bien connu : au travail, chacun-e cherche à se dégager des marges de manœuvre pour gagner un peu de pouvoir, de libertés par rapport à des contraintes. Si les systèmes d'IA déployés par les organisations se plient à la morale universelle, leur intronisation est justifiée par les intérêts des dirigeant-e-s. Évoluant aux marges de la moralité, l'expérience sensible et directe du travailleur ou de la travailleuse dans son usage de ChatGPT est celle de la liberté conquise. L'immoralité abstraite de ChatGPT est ainsi compensée par une moralité située : le pouvoir d'agir, l'autonomie au travail, un « principe de justice » « qui renvoie à une expérience propre et strictement subjective dans laquelle l'acteur mesure la qualité même de son travail » et qui s'oppose au « sentiment d'aliénation subjective » (Dubet, 2006, p. 26).

Ces remarques ont des conséquences sur la posture de ceux et celles qui énoncent ces principes universels. Elle montre en effet l'inconsistance de règles dictées en amont par des comités d'expert-e-s qui, par définition, fondent la légitimité de leur expertise précisément sur leur abstraction des pratiques d'usages, riches en valeurs fines, complexes, contestataires. Mais, cette crise de l'expertise face à la moralisation des machines communicantes ne se limite pas aux comités d'expert-e-s interrogés pour produire des recommandations. En effet, elles conduisent à remettre en discussion, au passage, le rôle et la contribution spécifique du chercheur ou de la chercheuse, le ou la situant à son tour dans la définition des enjeux éthiques du développement technologique. Une tribune comme celle dont nous avons la chance de profiter par la publication d'un article scientifique se prête elle aussi à devenir un lieu de parole où, prenant du recul par rapport à des situations particulières, l'on propose de formuler des logiques de fonctionnement générales décrivant le lien humain-e/technologie. Or, notre ethnographie montre que « l'humain-e » et « la technologie » sont des termes vides de sens en dehors des relations multiples que les pratiques d'usage en contexte instaurent entre eux. En effet, ces pratiques rendent visibles et lisibles les valeurs morales inscrites dans les déploiements eux-mêmes, en ouvrant à la moralisation des machines un terrain de

réflexion moins limité que celui offert par les textes érigés en barricades et par leur morale universaliste.

C'est là qu'intervient le basculement d'une conception un peu scientiste de la science sociale à une conception que je qualifierais volontiers de « diplomatique », où il s'agit de multiplier les comptes rendus d'expérience, de les comparer et de les confronter, pour ne rien perdre de la diversité (Latour, 2012, p. 952).

Ainsi, réfléchir « en diplomate » au sens du travail par le prisme des usages des technologies signifie faire émerger et articuler entre elles les multiples façons que les travailleurs et travailleuses, les gestionnaires et les concepteurs ou conceptrices adoptent pour négocier le sens des outils et de ce qu'ils et elles en font. Cela signifie aussi que, loin d'adopter une vue de Sirius sur les enjeux de moralisation des machines, les chercheurs et chercheuses deviennent partie prenante d'une activité de moralisation qui consiste à se faire porte-parole des multiples instances éthiques immanentes aux pratiques, dont on cherche à faciliter l'arbitrage en situation. Nous pourrions alors en tirer nous-mêmes une recommandation pour l'encadrement éthique de l'IA, à côté des autorités éthiques consacrées : rendre visibles les pratiques et leurs éthiques situées pour mettre en évidence les conditions d'une moralisation efficace.

Références

- Aloisi, A. et De Stephano, V. (2022). *Your Boss Is an Algorithm: Artificial Intelligence, Platform Work and Labour*. Bloomsbury Publishing.
- Barthes, Y., Callon, M. et Lascoumes, P. (2001). *Agir dans un monde incertain. Essai sur la démocratie technique*. Seuil.
- Benveniste, É. (1966). Communication animale et langage humain. *Problèmes de linguistique générale 1* (pp. 56-62). Gallimard.
- Cohen, S. (1972). *Folk Devils and Moral Panics*. MacGibbon et Kee.
- Colby, K. M., Watt, J. B. et Gilbert, J. P. (1966). A Computer Method of Psychotherapy: Preliminary Communication. *Journal of Nervous and Mental Disease*, 2(142), 148–152.
- CNIL (2017). *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*. Synthèse du débat public animé par le CNIL dans le cadre de la mission de réflexion éthique confiée par la loi pour une république numérique, https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_garder_la_main_web.pdf
- Crawford, M. (2016). *Eloge du carburateur*. La Découverte.
- D'Attanasio, S., De Martino, S. et Ferguson, Y. (2023). An Innovative Method to Study the Social Impact of AI on the Work Environment Based on a Multi-dimensional Human-Centred Analysis of the Worker-AI Team. Dans I. Maglogiannis, L. Iliadis, J. MacIntyre, M. Dominguez (dir.). *Artificial Intelligence Applications and Innovations, (IFIP Advances in Information and Communication Technology*, 676, pp. 45-56). Cham : Springer. https://doi.org/10.1007/978-3-031-34107-6_4
- Descola, P. (2005). *Par-delà nature et culture*. Seuil.
- Descola, P. (2011). *L'écologie des autres. L'anthropologie et la question de la nature*. Quae.
- Debaise, D., Stengers, I. (2023). *Au risque des effets. Une lutte à main armée contre la raison ?* Les liens qui libèrent.
- Devillers, L., (2017). *Des robots et des hommes. Mythes, fantasmes et réalité*. Plon.
- Dubet, F. (2006). *Injustices. L'expérience des inégalités au travail*. Seuil.
- Dussuet, A. (2019), Travailleur.se.s du care. Dans M.-C. Bureau et al. (dir.), *Les zones grises du travail et de l'emploi. Un dictionnaire sociologique* (p. 601-612). Teseo. <https://www.teseopress.com/dictionnaire/>
- Eymard-Duvernay, F. (2008). Justesse et justice dans les recrutements. *Formation emploi*, 101, janvier-mars 2008
- Ferguson, Y. et Messal, S. (2023). Ce que la machine parlante veut dire. De la révolution technologique à la reproduction sociale. *Nouvelle revue de psychosociologie*, 36(2), 181-194.
- Fernagu Oudet, S. (2012). Concevoir des environnements de travail capacitants : l'exemple d'un réseau réciproque d'échanges des savoirs. *Formation emploi*, 119. <https://doi.org/10.4000/formationemploi.3684>
- Frisch, K., von (1965). *Le langage de la danse et l'orientation des abeilles* [Tanzsprache und Orientierung der Bienen], Springer Verlag.
- Gaborieau, D. (2017). Quand l'ouvrier devient robot. Représentations et pratiques ouvrières face aux stigmates de la déqualification. *L'Homme & la Société*, 3(205), 245-268.
- Ganascia, J.-G. (2022). *Servitudes virtuelles*. Seuil.
- Girard-Chanudet, C. (2021). « Les chiffres ne mentent pas » : mythe d'objectivité et construction des données. LINC, Laboratoire d'Innovation Numérique de la CNIL, <https://linc.cnil.fr/fr/les-chiffres-ne-mentent-pas-mythe-dobjectivite-et-construction-des-donnees-le-travail-des-donnees-13>
- Goujon, V. (2023). *OpenAI, une histoire en trois temps*. AOC, <https://aoc.media/analyse/2023/05/22/openai-une-histoire-en-trois-temps/>
- Greimas, A. J. et Courtés, J. (1979). *Sémiotique. Dictionnaire raisonné de la théorie du langage*. Hachette.
- Grinbaum, A. (2023). *Parole de machines. Dialoguer avec une IA*. humenSciences.
- Habermas, J. (1986 [1983]). *Morale et communication : conscience morale et activité communicationnelle* (C. Bouchindhomme, trad.). Éditions du Cerf.
- Hjelmslev, L. T. (1953). *Prolégomènes à la théorie du langage*. Minuit.

- Latour, B. (1993). Le « pédofil » de Boa-Vista : montage photo-philosophique. Dans B. Latour (dir.), *La clef de Berlin. Petites leçons de sociologie des sciences* (p. 171-225). La Découverte/Poche.
- Latour, B. (1999). *Politique de la nature. Comment faire entrer les sciences en démocratie*. La Découverte.
- Latour, B. (2012). Bruno Latour : « L'universel il faut le faire ». Dans É. During, L. Janpierre (dir.), *Critiques*, 11(786), 949-963.
- Lee, K. F. (2019). *La plus grande mutation de l'histoire*. Les Arènes.
- Menei, P. (2021). Elon Musk, le singe et les trois cochons. Une fable transhumaniste ? *Futuribles*, 6(445), p. 87-93. <https://www.cairn.info/revue-futuribles-2021-6-page-87.htm>
- Michard, C. et Viollet, C. (1991). Sexe et genre en linguistique – Quinze ans de recherches féministes aux États-Unis et en R.F.A. *Recherches féministes*, 4(2), p. 97-128. <https://doi.org/10.7202/057654ar>
- Morin, E. (2005). *Introduction à la pensée complexe*. Seuil.
- Paveau, M.-A. et Ruchon, C., (2020). La linguistique et le langage animal. Résistances, décentremens, propositions. *Itinéraires*, (2). <http://journals.openedition.org/itineraires/8202> ; <https://doi.org/10.4000/itineraires.8202>
- Rosanvillon, P. (2011). *La société des égaux*. Seuil.
- Scardigli, V. (2001). *Un anthropologue chez les automates. De l'aviation informatisé à la société numérisée*. PUF.
- Sfez, L. (1988). *Critique de la communication*. Seuil.
- Simon, H. (1979). Rational Decision Making in Business Organizations. *The American Economic Review*, 4(69), 493-513.
- Simondon, G. (2014 [1958]). *Du mode d'existence des objets techniques*. Aubier.
- Tisserons, S. (2015). *Le jour où mon robot m'aimera*. Albin Michel.
- Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, (59), 433-460.
- UNESCO (2020). *Je rougirais si je pouvais : réduire la fracture numérique entre les genres par l'éducation*.
- Velkovska, J. et Zouinar, M. (2020). Les relations aux machines « conversationnelles ». Vivre avec les assistants vocaux à la maison. *Réseaux*, 2-3(220-221), 47-79.
- Vidal, D. (2012). Vers un nouveau pacte anthropomorphe ! Les enjeux anthropologiques de la nouvelle robotique. *Gradhiva*, (15), <https://doi.org/10.4000/gradhiva.2319>
- Villani, C. (2018). *Donner un sens à l'Intelligence Artificielle*. <https://www.enseignementsup-recherche.gouv.fr/fr/rapport-de-cedric-villani-donner-un-sens-l-intelligence-artificielle-ia-49194>
- Zacklad, M. et Rouveroy, A. (2022). L'éthique située de l'IA et ses controverses. *Revue française des sciences de l'information et de la communication*, (25). <http://journals.openedition.org/rfsic/13204>; <https://doi.org/10.4000/rfsic.13204>
- Zinna, A. (2005). L'objet et ses interfaces. Dans J. Fontanille et A. Zinna (dir.), *Les objets au quotidien* (coll. « Nouveaux Actes Sémiotiques », p. 161-192). PULIM.