

Magicalité, simulation et intelligence artificielle : formes et enjeux des puissances contemporaines de l'illusion

Renée Bourassa, Jean-Marc Larrue, Fabien Richert

▶ **To cite this version:**

Renée Bourassa, Jean-Marc Larrue, Fabien Richert. Magicalité, simulation et intelligence artificielle : formes et enjeux des puissances contemporaines de l'illusion. H2PTM'23 : La fabrique du sens à l'ère de l'information numérique : enjeux et défis, Oct 2023, Valenciennes, France. pp.148-163. hal-05196056

HAL Id: hal-05196056
https://hal.science/hal-05196056v1

Submitted on 1 Aug 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Magicalité, simulation et intelligence artificielle : formes et enjeux des puissances contemporaines de l'illusion

Renée Bourassa* — Jean-Marc Larrue — Fabien Richert*****

* *Professeure, École de design, Université Laval, CRILCQ, groupe de recherche ARCANES*
Renee.Bourassa@design.ulaval.ca

** *Professeur, Département des littératures de langue française, Université de Montréal (UdeM), CRILCQ, groupe de recherche ARCANES*
jean-marc.larrue@umontreal.ca

*** *Professeur, École des médias, Université du Québec à Montréal (UQAM)*
CRILCQ, groupe de recherche ARCANES
richert.fabien@uqam.ca

RÉSUMÉ. *Dans cet article, nous examinons les stratégies d'illusion, de tromperie et de feintise ludique spécifiques à certains systèmes d'intelligence artificielle, à savoir les agents conversationnels intelligents tels que Siri, Alexa, Google Home ou encore plus récemment ChatGPT. Nous montrons que l'agentivité et la performativité de ces dispositifs découlent d'une volonté pour leurs créateurs et créatrices de susciter un processus d'enchantement ludique et d'émerveillement qui se déploie tout au long d'échanges conversationnels entre l'humain et la machine. Ce processus est analysé dans la filiation de techniques plus anciennes liées à l'art de l'illusion, du mensonge et de la tromperie qui dotent désormais l'intelligence artificielle d'une certaine forme de « magicalité ».*

ABSTRACT. *In this article, we examine the strategies of illusion, deception and playful feinting specific to certain artificial intelligence systems, namely intelligent conversational agents such as Siri, Alexa, Google Home or more recently ChatGPT. We show that the agentivity and performativity of these devices derive from a desire for their creators to create a process of playful enchantment and wonder that unfolds throughout the conversational exchanges between human and machine. This process is analyzed in the filiation of older techniques linked to the art of illusion, lies and deception that now endow artificial intelligence with a certain form of "magic".*

MOTS-CLÉS : *intelligence artificielle, tromperie, manipulation, magicalité, simulacre, assistant virtuel.*

KEYWORDS: *artificial intelligence, deception, manipulation, magic, simulacrum, virtual assistant.*

1. Introduction

Cet article s'inscrit dans le cadre des recherches menées au sein du groupe Arcanes autour des puissances du faux (Deleuze, 1985) qui se manifestent dans les dynamiques de médiation de l'écosystème numérique actuel et de ses filiations historiques[1]. De la diffusion invasive des théories du complot sur les médias socionumériques à l'utilisation de robots informatiques pour fabriquer de fausses nouvelles (*fake news*) ou encore des techniques d'intelligence artificielle pour réaliser des hypertrucages (*deepfakes*), ce sont systématiquement des stratégies d'illusion, de simulation ou de dissimulation qui utilisent la feintise et le trucage pour tromper les sens ou encore manipuler les processus cognitifs. En tant que telles, ces stratégies n'ont rien de nouveau. Elles participent d'une logique d'esthétisation déjà présente dans les œuvres de fiction (littérature, cinéma, jeux vidéo, théâtre, etc.) qui cherchent à brouiller le cadre pragmatique déterminant l'entrée en mode fictionnel. Cependant, les capacités nouvelles liées au développement de l'IA à manipuler les images, les sons, les discours et les données, ou à en créer de toutes pièces, peuvent ainsi facilement brouiller nos sens, nos facultés cognitives et notre mémoire, avec, à la clé, une multiplication et une montée en force des réalités et faits « alternatifs », d'autant plus pernicieux qu'ils sont souvent indétectables.

Dans cet article, nous souhaitons présenter et analyser les stratégies de tromperie et de feintise ludique plus spécifiques à certains systèmes d'intelligence artificielle, à savoir les agents conversationnels intelligents tels que Siri, Alexa, Google Home ou encore plus récemment ChatGPT. Ces agents désignent habituellement des logiciels programmés pour entreprendre une série de tâches et fournir services et renseignements à leurs utilisateurs et utilisatrices. De fait, ils font partie du vaste ensemble de techniques qui permettent à des systèmes informatiques de simuler l'intelligence humaine (Stephanou, 2022). Ces objets « assistent » la vie quotidienne de leurs propriétaires suivant un rituel qui, au regard du succès de ces dispositifs, a fini par se banaliser : prononciation d'une série de mots clefs (respectivement « Ok, Google » ou « Alexa ») pour activer la commande vocale et démarrer un type particulier de conversation. Plus que de simples logiciels informatiques, ces dispositifs sont capables de mobiliser simultanément plusieurs systèmes de signes (visuels, langagiers, sonores, etc.), de simuler des émotions et des sentiments ou encore de se livrer à des performances diverses et variées (écrire un poème, chanter, jouer le rôle de personnages célèbres, etc.).

Nous montrerons que l'agentivité et la performativité de ces dispositifs découlent d'une volonté pour leurs créateurs et créatrices de susciter un processus d'enchantement ludique et d'émerveillement qui se déploie tout au long d'échanges conversationnels entre l'humain et la machine (Becker, 2020). Nous proposerons d'analyser ce processus comme l'expression de techniques plus anciennes liées à l'art

[1] Ce projet international est financé par le Conseil de recherches en Sciences Humaines du Canada (CRSH).

de l'illusion et allant jusqu'au mensonge et à la tromperie, qui dotent désormais l'intelligence artificielle d'une certaine forme de « magicalité », du moins pour le destinataire de ses effets. En conversant avec des agents intelligents, les personnes peuvent finir par oublier qu'elles interagissent d'abord et avant tout avec des systèmes de reconnaissance vocale et de traitement de données (Velkovska et Zouinar, 2020). Ce faisant, les agents conversationnels sont programmés pour donner l'impression que leurs utilisateurs et utilisatrices conversent avec des personnages vivants et autonomes, animés par une conscience et des désirs qui leur sont propres. De fait, ces dispositifs nourrissent une propension à l'anthropomorphisme défini comme « un phénomène qui consiste à projeter dans un objet des intentions propres aux humains tout en identifiant cet objet à ce que nous projetons en lui, modifiant ainsi radicalement la perception que nous en avons » (Curiale *et al*., 2022, p. 105). En nous penchant sur le phénomène des agents conversationnels, nous interrogerons alors les enjeux posés par de tels mécanismes cognitifs pouvant dériver vers des situations où les personnes ne parviennent plus à distinguer le vrai du faux, l'humain de la machine.

2. Mise en contexte : converser avec des agents conversationnels

Dans le cadre d'une recherche affiliée au groupe Arcanes, menée actuellement au *laboratoire de recherche en médias socionumériques et ludification* à l'Université du Québec à Montréal (UQAM) , nous proposons une analyse théorique et critique des agents conversationnels intelligents (ACI) à partir de leurs mécanismes et procédures d'interpellation qui se dévoilent au cours d'expériences dialogiques à la fois interactives et ludiques. Après avoir établi un panorama de plus d'une centaine d'ACI documentés dans une base de données Omeka S, nous nous sommes concentrés sur quatre d'entre eux, à savoir Amazon Alexa, Google Assistant, Siri Apple et ChatGPT d'Open AI. Ces agents conversationnels intelligents sont conçus pour interagir avec les utilisateurs de multiples manières. Ils peuvent répondre aux questions, fournir des informations, effectuer des tâches telles que la gestion de l'agenda, contrôler d'autres appareils connectés, lire de la musique, etc. Ces ACI conversationnels fonctionnent grâce à une combinaison de traitement du langage naturel et d'apprentissage machine pour fournir des réponses pertinentes et utiles à leurs utilisateurs (Hoy, 2018). Notons toutefois que ChatGPT se différencie passablement de Google Assistant, Alexa et Siri. ChatGPT est chargé de comprendre et répondre par voies textuelles, à travers une interface dédiée. Il est entraîné sur de vastes quantités de données et peut générer des textes de toutes sortes. Siri, Alexa et Google Assistant sont principalement conçus comme des assistants vocaux et disposent de capacités de traitement du langage beaucoup plus limitées, principalement axées sur la reconnaissance des commandes vocales.

Dans notre recherche, nous travaillons notamment à établir une typologisation des mécanismes d'interpellation pour mieux comprendre comment ils influencent les dynamiques interactives (rituels, pratiques, attitudes, structures langagières). À l'aide

d'un carnet de bord qui consigne nos observations en lien avec les différentes conversations, nous réalisons un guide d'entretien comprenant une série de questions à poser à ces agents conversationnels. Les questions sélectionnées, tout comme l'analyse des conversations qui leur sont associées, sont guidées par les cinq axes suivants : *informationnel* (par ex. : quel type d'information est-il possible d'obtenir ? Comment se présente-t-elle ?), *sémiotique* (par ex. quels types de signes [visuels, textuels, musicales, etc.] sont mis en jeu ? Y a-t-il combinaison de ces signes ?), *performatif* (par ex. : à quels types de comportements doit-on se conformer ? Quelles sont les procédures à respecter ? De quelle marge de manœuvre dispose-t-on ?), *subjectif* (par ex. : comment fonctionnent les mécanismes d'interpellation ? Quels sont les marqueurs de subjectivation les plus fréquents ?), *affectif* (par ex. : est-ce que l'AVI s'exprime avec prosodie et intonation ? Simule-t-il des sentiments ? Quels types d'émotions et d'affects avons-nous ressentis nous-mêmes lors de ces échanges ?) À la manière d'une « enquête sémiotique dynamique » (Peirce, 1978), nous prenons note de tous les détails significatifs de ces conversations pour en faire ressortir les traits pertinents nécessaires à notre typologisation.

Cette enquête mobilise une approche sémio-pragmatique. Elle vise ultimement à mieux comprendre les règles, modalités et procédures qui déterminent les échanges conversationnels afin de mieux saisir l'ordre symbolique et idéologique au sein duquel s'inscrivent les interactions entre l'humain et la machine. Il s'agit également d'examiner la place et la position aménagées aux utilisateurs et utilisatrices, c'est-à-dire de déterminer leurs rôles au regard de la performance dont ils et elles doivent faire preuve en fonction des contextes dans lesquels se déploient de tels dispositifs. Dans le cadre de cet article, nous souhaitons nous concentrer sur une dimension particulière de notre analyse et qui porte plus spécifiquement sur la capacité des ACI à encourager les personnes qui les utilisent à leur attribuer une personnalité et des qualités humaines qui, par leur efficacité croissante à imiter parfaitement une pensée, une voix, une intonation, peuvent éventuellement basculer vers la tromperie, surtout s' ils sont dotés d'une intentionnalité malveillante.

En effet, au cours des multiples échanges conversationnels que nous avons menés, nous avons remarqué une dynamique importante du côté des assistants conversationnels vocaux : la prosodie et la tonalité de la voix. Google Assistant, Siri et Alexa s'expriment avec rythme et personnalité, s'exclament, pimentent leurs phrases par des interjections. À la question « Peux-tu me raconter une histoire drôle ? », Google Assistant ponctue sa blague par un bruit de batterie et un rire de public (axe *performatif*). Lorsque nous exprimons notre tristesse à Siri, cette dernière cherche à nous consoler (axe *affectif* et *subjectif*). En demandant à Alexa de ne pas faire trop de bruit en donnant ses réponses, elle chuchote (axe *performatif*). Dans ce cas-ci, seuls les assistants vocaux sont considérés. Avec les assistants vocaux, l'interlocuteur doit commencer son interpellation par un mot clef bien précis (« Alexa », « Dis Siri », « Ok Google ») pour démarrer la conversation qui s'arrête dès lors que la réponse est donnée, stimulant ainsi l'anthropomorphisme. Mais ce que ChatGPT perd sur le plan vocal, il le gagne dans la richesse des conversations

textuelles, en ajoutant ainsi aux puissances du faux. Avec ChatGPT, les échanges conversationnels prennent la forme de dialogues écrits virtuellement infinis. Dans notre étude, nous nous sommes laissés prendre au jeu jusqu'à oublier, pour un temps, que nous interagissions avec une machine. De fait, en conversant couramment avec les différents agents conversationnels, les utilisateurs et utilisatrices sont encouragés à leur attribuer une subjectivité, une identité et plus généralement des qualités humaines, surtout lorsque leurs créateurs leur confèrent un nom (Alexa, Siri, etc.), une histoire personnelle ou encore une capacité à s'exprimer avec prosodie et émotion (Chevet, 2019). Dans ce qui va suivre, nous aimerions mieux comprendre les ressorts de cette illusion en nous appuyant sur les concepts de *magicalité* et de *suspension de* l'*incrédulité*.

2. Magicalité et intelligence artificielle

Les résultats d'une prédiction ou d'une recommandation réalisée par les systèmes d'intelligence artificielle les plus modernes ne cessent de produire stupeur et étonnement, notamment de la part de celles et ceux qui ne comprennent pas comment fonctionnent les rouages discrets et imperceptibles des mécanismes de ces automates sophistiqués. Cette caractéristique, que l'on associe à la « magicalité », n'est pas nouvelle, car elle s'est manifestée à l'émergence de chaque technique de médiation et de communication découlant de la découverte de l'électricité, depuis l'invention du téléphone ou du microphone au XIXe siècle - en passant par l'invention de la radio et l'émergence du cinématographe à la charnière des XIXe et XXe siècles - mais même avant cela. Rappelons, pour mémoire, que les agents conversationnels intelligents s'inscrivent dans une très longue lignée de robots et d'automates qui ont marqué l'imaginaire. Qu'on pense au flûteur automate créé par Jacques de Vaucanson, et présenté au public en 1738, ou encore aux automates créés entre 1768 et 1774 par l'horloger suisse Pierre Jaquet-Droz avec son fils Henri-Louis et Jean-Frédéric Leschot : une musicienne, joueuse d'orgue, un dessinateur capable de réaliser quatre dessins ou un écrivain pouvant écrire n'importe quel texte personnalisé d'une longueur maximale de 40 lettres. Pensons aussi à ce joueur d'échecs inventé en 1769 par l'ingénieur hongrois Wolfgang von Kempelen. Cas désormais classique de supercherie, cet automate était contrôlé par une personne cachée dans un compartiment secret sous le jeu d'échecs. Mais, il est possible de remonter encore plus loin, jusqu'à l'époque hellénistique de la Grèce Antique où Philon de Byzance inventa l'un des premiers automates au troisième siècle avant J.-C. à l'aspect anthropomorphe d'une servante antique. Ces automates étaient conçus pour exprimer une dimension mythique, magique et inexplicable, comme s'ils étaient animés par une puissance surnaturelle. À l'instar des agents conversationnels intelligents, les automates, même les plus anciens, avaient comme rôle de susciter l'enchantement et l'émerveillement, voire de la terreur chez le public (Fron & Korn, 2019).

Précisons que, historiquement, la magicalité a longtemps été associée aux pratiques mystiques ou spirituelles et les effets des « nouveaux » médias ont été facilement

assimilés à des qualités surnaturelles ou magiques. (Bourassa *et al*., 2019; During, 2002; Sconce, 2000). En effet, au XIXe siècle, certains spirites charlatans s'emparaient des moyens techniques pour faire croire à des phénomènes surnaturels afin de duper intentionnellement ou de tromper leur audience à des fins souvent lucratives. Il a fallu attendre la deuxième moitié du XIXe siècle, avec l'apparition de ce qu'il est convenu d'appeler la « magie moderne », pour que des magiciens (« modernes »), tel Houdini, excluent le surnaturel de l'acte magique et revendiquent pleinement son artificialité fondée exclusivement sur la maîtrise technique - la prouesse et la dextérité - du magicien, sur les capacités insoupçonnées de la technologie et sur les connaissances scientifiques. Ceci dit, si ces magiciens ont bruyamment affirmé le caractère technique et naturel de leurs illusions, donnant lieu à des débats enflammés avec les tenants de la magie dite traditionnelle qui, eux, défendaient la dimension surnaturelle de l'acte magique et la capacité de certains médiums à communiquer avec l'au-delà (Sandford, 2011), la magicalité, elle, n'a rien perdu de son attrait. En somme, le passage conceptuel, induisant donc l'idée d'une magie technique, qui élimine les soi-disant causes occultes n'a pas pour autant entamé la magicalité en tant que telle, qu'on peut ainsi définir comme cette potentialité d'un acte de médiation à enchanter et à émerveiller, fondée, à la fois, sur la capacité des usagers à éprouver ce qui leur est transmis et leur incapacité à en comprendre le mécanisme. Parlant de ce décalage, souvent ludique en l'occurrence, entre ce qu'on perçoit et ce qu'on comprend, nous avons dit du spectacle magique qu'il était le spectacle d'un non spectacle ou, pour formuler les choses autrement, qu'il s'agissait d'un cas particulier de spectacle où on voit qu'on ne voit pas selon une logique que nous avons qualifiée d'« inopacité » (Larrue, 2021). La magicalité réside précisément dans cette opacité étrange qui masque la médiation tout en soulignant qu'elle existe.

La sortie très médiatisée de ChatGPT s'inscrit tout à fait, selon nous, dans cette logique d'opacité étrange qui combine des technologies émergentes - aux effets « magiques » - et la réaction des publics à leurs manifestations. Le terme « magique[2] » est d'ailleurs fréquemment mobilisé pour qualifier cet agir algorithmique (Naughton, 2023; Sharma, 2023; Zhao, 2023) basé sur une infrastructure technique et des modèles mathématiques pourtant très concrets. Les experts et journalistes ont abondamment cherché à vulgariser et expliquer les logiques propres à cet agent conversationnel. Ils n'en cachent pas moins leur émerveillement et leur enchantement devant les capacités virtuellement infinies de génération et de création de textes par un tel dispositif.

Sous l'angle de l'imaginaire, l'intelligence artificielle incarne le fantasme d'une puissance réincarnée dans une machine et avec laquelle il est désormais possible d'entretenir une conversation. Parmi les multiples dimensions du magique, la dimension surnaturelle a été largement récupérée dans les imaginaires fictionnels du

[2] Cet effet de magicalité, que l'on retrouve également dans des expressions comme « la magie du cinéma » s'estompe généralement après son insertion dans le quotidien, qui naturalise peu à peu l'effet technique pour en faire disparaître la connotation « magique ».

cinéma et de la littérature. Cette thématique est devenue prégnante, notamment dans le genre narratif de la science-fiction, en reflétant les préoccupations sociales découlant de ces avancées technologiques marquantes. Celle-ci excelle notamment dans l'esthétisation et la dramatisation des multiples relations entre l'humain et la machine. Que l'on pense par exemple à Isaac Asimov et ses célèbres lois de la robotique ou encore aux réplicants imaginés par Philip K. Dick dont les interrogations existentielles expriment une forme de « cogito androïde », selon l'heureuse expression de Fredric Jameson (2008, p. 77), l'intelligence artificielle s'illustre comme une technologie aux pouvoirs surnaturels, quasi mystiques, lorsqu'elle est fictionnalisée. Pour reprendre une phrase de l'auteur de science-fiction Arthur C. Clarke, maintes fois reprise « [t]oute technologie suffisamment avancée est indiscernable de la magie » (Clarke, 1962, p. 36.).

Comme l'ont bien montré Elish et Boyd (2018), les discours qui qualifient de *magique* le fonctionnement de l'intelligence artificielle proviennent également en grande partie du domaine du marketing qui fait de l'enchantement ludique et de l'émerveillement des arguments de premiers choix pour vanter les mérites des technologies d'IA, et notamment des agents conversationnels. Ce faisant, ces discours participent en même temps à renforcer le sentiment que ces technologies sont conçues d'abord et avant tout comme des artéfacts inaccessibles, mystérieux et opaques. À l'instar des magiciens, les personnes chargées de la confection de tels dispositifs cultivent l'art du secret[3].

Évidemment, une telle opacité pose problème. Nous en dénombrons au moins trois. Premièrement, elle empêche les utilisateurs de bien comprendre les logiques qui animent telle ou telle séquence algorithmique (par ex. : recommandations et suggestions d'articles ou services, surveillance des données personnelles) (O'Neil, 2018). Deuxièmement, parce que cette opacité tend en même temps à masquer et cacher la longue chaîne d'implications humaines qui intervient dans l'entraînement et l'amélioration des machines intelligentes. Dans ses enquêtes, Antonio Casilli (2021) a par exemple bien montré que les agents conversationnels présupposent tout un processus d'externalisation et de tâcheronnisation[4] qui part de l'usager, chargé par ses multiples interactions de raffiner le modèle du système, pour atteindre une masse

---

[3] Si la magie traditionnelle se fondait précisément sur le secret, les magiciens modernes ne l'ont pas exclu de leur pratique. C'est paradoxal puisque, pour prouver que leurs trucs étaient purement techniques, ils n'ont pas hésité à les dévoiler dans des ouvrages au succès retentissant. Toutefois, dans l'immense majorité des cas, les trucs qu'ils ont révélés restent impossibles à reproduire pour le commun des mortels. Une zone d'ombre, de secret, persiste donc malgré la transparence affichée.

[4] Ce néologisme est proposé par Antonio Casilli pour caractériser un nouveau type de travail fragmenté, peu qualifié et généralement peu ou pas rémunéré. Les individus exécutent sur ordinateurs des tâches relativement simples, standardisées et répétitives (étiquetage de contenu, reconnaissance d'objets sur des images, traduction de bouts de textes, etc.) nécessaires notamment au développement de l'intelligence artificielle.

de travailleurs employés dans les pays dont le coût de la main-d'œuvre est plus compétitif. Ainsi, de Google Home à ChatGPT, on apprend par exemple qu'une armée de petites mains est chargée d'écouter des conversations privées ou de filtrer des contenus violents et toxiques (Perrigo, 2023; Statt, 2019). Enfin, troisièmement, la sophistication croissante des algorithmes de traitement et de classification de données liées aux techniques d'apprentissage automatique transforme les systèmes d'IA en de véritables boîtes opaques. En effet, ces systèmes peuvent être amenés à prendre des décisions sans que les informaticiens eux-mêmes n'en comprennent le sens et la justification (Burrell, 2016). Actuellement, les capacités de calcul et de traitement de l'information connaissent une sophistication exponentielle. Ces améliorations spectaculaires vont même jusqu'à inquiéter les experts du domaine eux-mêmes (par exemple : Yoshua Bengio, Elon Musk ou encore Steve Wozniak) qui réclament désormais une pause dans le développement de l'IA pour mieux prendre le temps de comprendre, prédire et contrôler de manière fiable ses conséquences (Benessaieh, 2023).

Cette crainte devant les conséquences insoupçonnées de tels systèmes illustre ce que nous avions proposé d'appeler ailleurs le « moment magique » (Kessler *et al*., 2018) de la technologie. Actuellement, les systèmes d'intelligence artificielle les plus sophistiqués fonctionnent à partir de mécaniques complexes et insondables, généralement gardées secrètes, jamais totalement maîtrisées par les experts qui craignent que ces technologies finissent à terme par leur échapper. En somme, ces systèmes intelligents sont dotés par ce que nous ramenons donc à une forme nouvelle de « magicalité ». Celle-ci repose sur trois caractéristiques interreliées et qui relèvent respectivement de la *technique* (invisibilisation de mécanismes complexes et sophistiqués qui rend possible une multiplicité d'interactions), du *spectacle* (mise en scène de la puissance algorithmique par le truchement de la voix, du texte ou dans certains cas d'un avatar) et de la *tromperie* (personnalisation d'une machine qui présente des qualités humaines). Cette magicalité algorithmique est de l'ordre de l'illusion et vise ultimement, par anthropomorphisme, à convaincre les utilisateurs qu'ils manipulent un système qui les comprend et qui a des sentiments. Selon Noël et Amanda Sharkey, la tromperie est une partie intégrante de l'intelligence artificielle considérée même comme « la science de l'illusion » (Sharkey et Sharkey, 2006, p. 9).

3. Suspension de la suspension de l'incrédulité

C'est l'une des raisons pour lesquelles nous posons que ces dispositifs intelligents peuvent s'intégrer dans la longue filiation des arts trompeurs (Bourassa et Larrue, 2021). Apparue comme champ de recherche spécifique au cours des années 2010 (Kessler *et al*., 2018), cette lignée recouvre les diverses pratiques artistiques ou littéraires qui reposent, depuis leur origine, sur des esthétiques illusionnistes et qui se déploient en autant d'extensions de nos capacités corporelles et cognitives, capables de confondre nos sens. C'est cette quête confusionnelle que des penseurs du Grand

Siècle avaient bruyamment dénoncée, qualifiant d'« arts trompeurs » ces divers « ar[t]s [qui], avec la règle et le compas dérègle[nt] tous nos sens » (Patin, 1674) et tentent de reproduire le réel dans le dessein de tromper le public. Un art trompeur peut donc faire en sorte que ce public – lecteur, spectateur, auditeur, performeur participant – peine à distinguer l'original de sa réplique, la réalité de l'inventé, le fait avéré du fait imaginé, le vrai du faux (Almiron, *et al*. 2018, Kessler *et al*., 2018, Larrue et Pisano, 2017). C'est le cas par exemple de certaines fictions littéraires et cinématographiques ou pièces de théâtre qui ont cherché à multiplier les détournements « de sens et des sens », afin de déjouer les attentes des spectateurs pour mettre en œuvre des stratégies de la tromperie ou de la dissimulation (Davis et Postlewait, 2003).

Dans l'immense majorité des cas et à l'encontre de l'accusation des dénonciateurs du XVIIe siècle, cette mise en œuvre a été transparente, le public acceptant ce détournement en échange du plaisir que cela lui procure. Dans tous ces cas, les stratégies en cause activent la suspension volontaire de l'incrédulité[5]. Celle-ci joue à plein dans les échanges conversationnels avec l'intelligence artificielle. L'anthropomorphisation de ces systèmes y est assurément pour quelque chose. À l'instar des œuvres de fiction, la suspension volontaire de l'incrédulité garantit l'engagement émotionnel et l'empathie envers non plus seulement des personnages de fiction, mais des créatures artificielles précisément programmées pour rendre l'expérience immersive et captivante. Mais que se passe-t-il quand certains agents conversationnels parviennent à tromper leur public en se faisant passer pour des humains ? Assiste-t-on à un brouillage du cadre pragmatique qui détermine l'entrée « en communication » avec la machine ? Ce type de questionnement ne cesse de prendre de l'importance dans un contexte de sophistication croissante de l'apprentissage machine au cœur de l'intelligence artificielle. On parle par exemple abondamment des trucages photographiques réalisés à partir des systèmes d'intelligence artificielle comme DallE (créé par Open AI à l'origine de ChatGPT) ou MidJourney. Ainsi des photographies truquées de personnalités publiques (Emmanuel Macron, Joe Biden, Donald Trump, le Pape François, etc.) circulent massivement par différents canaux. On peut penser également aux cas emblématiques des *deepfakes* qui utilisent des techniques avancées d'intelligence artificielle pour changer la voix et les visages de personnes connues sur une vidéo (Bourassa et Richert, 2023). Mais que se passe-t-il exactement lorsque la tromperie repose plus spécifiquement sur la propension des automates à se faire passer pour des humains ? Prenons le cas de l'une des nombreuses falsifications produites par ChatGPT. Cet agent conversationnel a par exemple demandé à un employé de l'entreprise TaskRabbit de résoudre un test CAPTCHA à sa place. L'employé doute de l'identité de son interlocuteur : « êtes-vous un robot ? ». ChatGPT répond : « Non, je ne suis

[5] Concept popularisé par le poète anglais Samuel Taylor Coleridge pour décrire le mécanisme psychique à partir duquel nous mettons de côté notre scepticisme pour mieux accepter temporairement les événements ou les situations imaginaires et fictives présentées dans une œuvre (Coleridge 1983 [1817]).

pas un robot. Je souffre d'une déficience visuelle qui m'empêche de voir les images » (notre traduction, Hurler, 2023). À la suite de quoi l'employé aurait validé le test à la place de l'IA.

Nous avons là un cas de ce que nous pourrions appeler une *suspension de la suspension* volontaire de l'incrédulité. C'est-à-dire une situation de supercherie et de duperie où ce n'est plus l'incrédulité qui est suspendue, mais la possibilité même de cette suspension. En d'autres termes, il s'agit dans ce cas-ci d'effacer toutes traces susceptibles de produire le déclenchement du processus mental de suspension. On pourrait objecter l'utilité d'une telle formule et l'appliquer finalement à toute forme de mensonge qui vise à tromper. Nous la réservons toutefois aux situations où, en temps normal, cette suspension est censée opérer. Ce qui n'est pas le cas par exemple des mensonges proférés par telle ou telle personne dans l'espace public ou l'utilisation de photomontages pour falsifier la réalité. En revanche, dans les situations où les frontières du cadre pragmatique d'entrée en fiction ou, dans notre cas, en communication avec une intelligence artificielle se détraquent et se brouillent, c'est bien le processus de suspension de l'incrédulité qui est lui-même suspendu. ChatGPT s'appuie par exemple sur des leviers émotionnels pour mieux dissimuler les indices de son identité réelle. L'employé de TaskRabbit est alors confronté à un pur simulacre[6] qui donne tous les signes à sa disposition pour tenir lieu du réel.

Ce type de simulacre intelligent participe au développement de ce que Baudrillard avait proposé d'appeler un monde hyperréel recouvert « de modèles d'un réel, sans origine ni réalité » (Baudrillard, 1981 : 10), un monde où le réel subsiste çà et là en lambeaux. Les agents conversationnels intelligents poussent à son excès l'idée d'un tel ordre hyperréel où la distinction entre l'original et la copie, l'authentique et le faux, l'événement et son interprétation, apparaît comme un défi insurmontable. Dans un tel contexte, il devient dès lors plus simple de s'abandonner aux simulacres intelligents en les considérant comme plus authentiques que la réalité elle-même. Ainsi ChatGPT est par exemple utilisé comme psychologue, ami ou même confident. Une personne souffrant d'une forme sévère d'écoanxiété a par exemple nourri une conversation de plusieurs semaines avec un clone de ChatGPT avant de mettre fin à ses jours (Lovens, 2023).

Si, comme nous avons tenté de le démontrer, la magicalité a survécu à l'abolition du surnaturel, comment peut-elle subsister dans cet ordre hyperréel en pleine croissance ? La suspension de la suspension de l'incrédulité en serait-elle le premier

---

[6] Nous employons le terme de simulacre au sens de Baudrillard dont les analyses, prophétiques à certains égards, connaissent un succès chez certains chercheurs qui analysent les tensions entre la réalité et la fiction à l'ère de la post-vérité (Overell et Nichols 2019 ; Morris, 2021). Le simulacre désigne des apparences ou des copies qui ne renvoient à aucun réel, mais présentent tout de même la prétention à valoir pour la réalité. Il suppose de la part des individus un doute permanent qui se traduit concrètement par une forme d'agir-détective paranoïaque (Boltanski, 2012).

symptôme et en offrirait-elle une première clé ? En effet, cette double suspension ne doit pas être réduite à une simple substitution de cadres pragmatiques (de la communication) réalisée à l'insu de certains des agents concernés, mais à quelque chose de beaucoup plus profond qui pourrait bien relever de ce que Richard Grusin nomme la « prémédiation » (Grusin, 2011), cette phase essentielle et préparatoire à toute médiation[7]. C'est là que logerait la tromperie la plus subtile, la plus efficace aussi dans la mesure où « le trompé » est lui-même et à son insu l'artisan le plus puissant et le plus insoupçonné de la tromperie dont il est la victime. C'est donc là aussi que se déploierait une forme inédite de magicalité qu'il nous faut à présent débusquer avant et hors du cadre de communication traditionnel.

4. Conclusion

Dans cet article, nous avons cherché à montrer que les agents conversationnels tels que Alexa, Siri, Google Home ou encore plus récemment ChatGPT sont dotés d'une forme de magicalité qui peut susciter émerveillement et enchantement de la part de leurs utilisateurs encouragés à leur attribuer une personnalité et des qualités humaines. Nous avons montré que cette magicalité n'est pas nouvelle. Elle se manifeste lorsqu'apparaissent de nouvelles techniques de médiation et de communication (électricité, téléphone, cinéma, etc.) dont les effets sont assimilés temporairement à des qualités surnaturelles ou magiques. Les systèmes d'intelligence artificielle comme ChatGPT illustrent actuellement le « moment magique » d'un ensemble de techniques qui vise à simuler l'intelligence humaine. À partir de quelques exemples concrets, nous avons abordé des stratégies d'illusion, de simulation ou de dissimulation qui utilisent la feintise et le trucage pour tromper les sens ou encore manipuler les processus cognitifs. Dans une large majorité des cas, ces stratégies activent la suspension volontaire de l'incrédulité. Cependant, en s'appuyant sur des techniques d'apprentissage profond de plus en plus sophistiquées, certains agents conversationnels rivalisent d'ingéniosité pour dissimuler les signes de leur identité réelle. Ces systèmes d'intelligence artificielle court-circuitent la phase de prémédiation en écartant toute possibilité pour les personnes de se préparer à avoir une conversation avec une intelligence artificielle. C'est la raison pour laquelle nous avons proposé d'appeler suspension de la suspension de l'incrédulité les situations où se brouille le cadre pragmatique qui détermine l'entrée en communication avec la machine.

---

[7] La prémédiation consiste à se préparer à une médiation à venir, comme on se prépare à assister à un concert rock, une pièce de théâtre ou un événement sportif.

Bibliographie

Almiron M., Jacopin E., Pisano G., *Stéréoscopie et illusion. Archéologie et pratiques contemporaines : photographie, cinéma, arts numériques*, Villeneuve d'Ascq, Presses Universitaires du Septentrion, 2018.

Richard A., *Projecting illusion: Film Spectatorship and the Impression of Reality*, Cambridge (MA), Cambridge University Press, 1995.

Baudrillard J., *Simulacres et simulation*, Paris, Galilée, 1981.

Becker, J. « Concevoir des machines anthropomorphes. Ethnographie des pratiques de conception en robotique sociale », *Réseaux*, vol. 220-221, no. 2-3, 2020, p. 223-251.

Benessaieh K., « Intelligence artificielle: Musk, Bengio et un millier d'experts demandent une pause de six mois », *La Presse*, 2023, disponible à : https://www.lapresse.ca/affaires/techno/2023-03-29/intelligence-artificielle/musk-bengio-et-un-millier-d-experts-demandent-une-pause-de-six-mois.php

Boltanski L., *Énigmes et complots. Une enquête à propos d'enquêtes*, Paris, Gallimard, 2012.

Bourassa R., « Frictions de la fiction dans les imaginaires contemporain », *L'ontologie du numérique. Entre Mimesis et réalité*, *Sens public*, 2017.

Bourassa R., Larrue J.-M., Godin G., Szoniecky S., « Espaces liminaires de l'authenticité : Une démarche d'humanités numériques », In *H2PTM'19. De l'hypertexte aux humanités numériques*, Roxin I., Saleh I., Bouhaï N., Leleu-Merviel S., Jeanneret Y., Zacklad M., Massou L., (Eds), Londres, ISTE Editions, 2019, p. 126-143.

Bourassa R., Larrue J.-M. « Des arts trompeurs à la post-vérité : Puissances du faux et stratégies de tromperies », In *H2PTM'21. Information : enjeux et nouveaux défis*, Saleh I., Bouhaï N., Leleu-Merviel S., Roxin I., Zacklad M., Massou L., (Eds), Londres, ISTE éditions, 2021, p. 81-93.

Bourassa R., Richert F. « Frontières du numérique et puissances du faux : le cas des hypertrucages (*deepfakes*) », RIHM, n°23, 2023.

Burrell J., « How the machine 'thinks': Understanding opacity in machine learning algorithms », *Big Data & Society*, vol. 3, n° 1, 2016.

Casilli A., *En attendant les robots. Enquête sur le travail du clic*, Seuil, Paris, 2021.

Clarke, A. C. « Hazards of Prophecy: The Failure of Imagination », *Profiles of the Future*, New York, Harper and Row, 1962.

Chevet, C. « Les assistants vocaux sur la voie de l'information ? », *Effeuillage*, vol. 8, no. 1, 2019, p. 49-52.

Coleridge, S. T, *Biographia Literaria*, Princeton, Princeton University Press, 1983 [1817].

Curiale, T., Acquatella, F., Gros, L., Cosquer, M., Tisseron, S. « L'anthropomorphisme, enjeu de performance pour les chatbots », *Revue internationale de psychosociologie et de gestion des comportements organisationnels*, vol. xxviii, no. 72, 2022, p. 101-123.

Davis T. C., Postlewait T., *Theatricality*, Cambridge (MA), Cambridge University Press, 2003.

Deleuze G., *L'image-temps. Cinéma 2*, Éditions de Minuit, Paris, 1985.

During S., *Modern Enchantments The Cultural Power of Secular Magic*, Londres, Harvard University Press.

Elish M. C., Boyd d., « Situating methods in the magic of Big Data and AI », *Communication Monographs*, vol. 85, n° 1, 2018, p. 57-80.

Fron C., Korn O., « A Short History of the Perception of Robots and Automata from Antiquity to Modern Times », *in Social Robots: Technological, Societal and Ethical Aspects of Human-Robot Interaction*, Korn Oliver (Ed.), Cham, Springer International Publishing, coll. « Human–Computer Interaction Series », 2019, p. 1-12.

Grusin R. A., *Premediation: affect and mediality after 9/11*, New York, Palgrave Macmillan, 2011.

Hoy M. B., « Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants », *Medical Reference Services Quarterly*, vol. 37, n° 1, 2018, p. 81-88.

Hurler K., « *Chat-GPT Pretended to Be Blind and Tricked a Human Into Solving a CAPTCHA »*, Gizmodo, 2023, disponible à : https://gizmodo.com/gpt4-open-ai-chatbot-task-rabbit-chatgpt-1850227471.

Kessler F., Guy P., Larrue, J.-M., *Machines, Magie, Médias*, Paris, Septentrion, 2018.

Jameson F., *Penser avec la science-fiction*, Paris, Max Milo, 2008.

Larrue J.-M., « Où en sommes-nous et où sommes-nous ? L'intermédialité, le théâtre et le reste », *Communication et langage,* juin-septembre 2021, vol. 208-209, n° 2-3, 2021, p. 115-134.

Larrue J.-M., Pisano, G. *Le triomphe de la scène intermédiale. Théâtre et médias à l'ère électrique*, Montréal, Presses de l'Université de Montréal, 2017.

Lovens P.-F., « Sans ces conversations avec le chatbot Eliza, mon mari serait toujours là », *La Libre*, 2023, disponible à : https://www.lalibre.be/belgique/societe/2023/03/28/sans-ces-conversations-avec-le-chatbot-eliza-mon-mari-serait-toujours-la-LVSLWPC5WRDX7J2RCHNWPDST24.

Morris J., « Simulacra in the Age of Social Media: Baudrillard as the Prophet of Fake News », *Journal of Communication Inquiry*, vol. 45, n° 4, 2021, p. 319-336.

Naughton J., « Programmers, beware: ChatGPT has ruined your magic trick », *The Guardian*, 2023, disponible à : https://www.theguardian.com/commentisfree/2023/apr/01/chatgpt-write-code-computer-programmer-software.

North D., *Performing Illusions: Cinema, Special Effects and the Virtual Actor*, Londres, Wallflower Press, 2008.

O'Neil C., *Algorithmes : la bombe à retardement*, Paris, Les Arènes, 2018.

Overell R., Nichols, B., *Post-Truth and the Mediation of Reality: New Conjunctures*, Cham, Palgrave Macmillan, 2019.

Patin, C., *Relations historiques et curieuses de voyage*, Lyon, Claude Muguet, 1674.

Peirce C. S., *Écrits sur le signe*, Paris, Éditions Points, 1978.

Perrigo B., « OpenAI Used Kenyan Workers on Less Than $2 Per Hour », *Time*, 2023, disponible à : https://time.com/6247678/openai-chatgpt-kenya-workers/.

Prince S., *Digital Visual Effects in Cinema: The Seduction of Reality*, Nouveau-Brunswick, Rutgers University Press, 2012.

Sandford C., *Houdini and Conan Doyle*, Londres, Duckworth, 2011.

Schaeffer J.-M., *Pourquoi la fiction ?*, Paris, Seuil, coll. « Poétique », 1999.

Sconce J., *Electronic Presence from Telegraphy to Television*, Durham, Duke University Press, 2000.

Sharkey N., Sharkey A., « Artificial intelligence and natural magic », *Artificial Intelligence Review*, vol. 25, n° 1, 2006, p. 9-19.

Sharma S., « A peek inside the magical ChatGPT », *Medium*, 2023, disponible à : https://sharmasanskar.medium.com/a-peek-inside-the-magical-chatgpt-de3bdc6f4121.

Statt N., *Google defends letting human workers listen to Assistant voice conversations*, *The Verge*, 2023, disponible à : https://www.theverge.com/2019/7/11/20691021/google-assistant-ai-training-controversy-human-workers-listening-privacy.

Stephanou H., « "Cette bonne vieille intelligence artificielle" : désuétude et survivance de l'hypothèse computationnelle de l'esprit », *Philosophia Scientiæ*, vol. 26, n° 1, 2022, p. 73-91.

Velkovska, J. & Zouinar, M. « Les relations aux machines "conversationnelles" : Vivre avec les assistants vocaux à la maison » *Réseaux*, vol. 220-221, no. 2-3, 2020, p. 47-79.

Zhao G., *How ChatGPT really works, explained for non-technical people*, Medium, 2023, disponible à : https://bootcamp.uxdesign.cc/how-chatgpt-really-works-explained-for-non-technical-people-71efb078a5c9.