

**Faculté des sciences économiques,
sociales, politiques et de communication**

Les représentations et usages des agents conversationnels

**Considérations sous le prisme du design fiction
et du rapport à la voix**

Auteur : Laura Zander
Promoteurs : Géraldine Wuyckens
Jerry Jacques
Année académique 2022-2023
Master 120 en Sciences et Technologies de
l'Information et de la communication, à
finalité spécialisée en conception et
évaluation de médias éducatifs

Je remercie Géraldine Wuyckens et Jerry Jacques, mes promoteurs, pour leur disponibilité, leur encadrement et leur soutien. Les nombreuses conversations ouvertes qui ont jalonné ma rédaction durant ces deux dernières années ont véritablement nourri la réflexion autour de ce mémoire et la fierté que j'ai eue à le concrétiser.

Je remercie ma sœur Émilie et Miloud, pour les nombreuses discussions pertinentes sur le sujet et le partage de leurs expertises respectives qui ont permis de donner du relief à ce mémoire et une transversalité qui me tenait à cœur.

Je remercie Lucas, pour son soutien à toute épreuve lors de la rédaction de ce mémoire et au quotidien, ainsi que Loïc, Olivier et ma maman pour leurs précieuses relectures et leur intérêt.

Je remercie les personnes qui ont pris le temps de répondre à mon questionnaire, et tout particulièrement celles qui m'ont accordé un peu plus de temps encore pour mener des entretiens.

À mes camarades du master STIC, qui m'ont moralement et intellectuellement soutenue plus d'une fois et qui ont rendus le rush final plus supportable.

À toutes ces personnes inspirantes rencontrées dans le monde académique et industriel durant l'année écoulée, dont les esprits et l'inventivité m'ont profondément marquée et grâce à qui j'ai passé de merveilleux moments.

À Papa et Opa.

Table des matières

Remerciements.....	1
1. Introduction	4
2. État de l'art.....	6
2.1. Les agents conversationnels.....	6
2.1.1. Terminologie et fonctionnement	6
2.1.2. Origines du terme et historique	7
2.1.3. L'interactivité, l'interaction et l'intelligence	8
2.1.4. Questionnements éthiques et réglementaires.....	10
2.2. Le rôle de la voix et des émotions	11
2.2.1. Considérations biologiques et synthétiques	11
2.2.2. Les modèles de langage	12
2.2.3. La reconnaissance vocale et les émotions	13
2.2.4. La notion de conversation et les médias automatisés.....	14
2.3. Le design fiction	15
2.3.1. Cadre conceptuel et définition.....	15
2.3.2. L'influence du visuel.....	16
2.3.3. L'influence de la narration.....	17
2.4. Les agents conversationnels dans la fiction	17
3. Problématique.....	19
3.1. Question de recherche.....	19
3.1.1. Représentations, attitudes et comportements	20
3.1.2. Le rôle de la fiction.....	21
3.2. Hypothèses.....	21
4. Méthodologie.....	22
4.1. Questionnaire	22
4.1.1. Distinction entre « média contemplatif » et « média interactif ».....	23
4.2. Entretiens	24
4.2.1. Définition de l'échantillon.....	25
4.2.2. Constitution du guide d'entretien	25
4.3. Retours sur la méthode et justifications	26
4.3.1. Ambiguïté définitionnelle.....	26
4.3.2. Choix de l'échantillon pour les entretiens.....	27
4.3.3. Œuvre de fiction du neuvième entretien	28

4.3.4.	<i>Contradictions dans le questionnaire.....</i>	29
5.	Résultats	31
5.1.	Récolte de données	31
5.1.1.	<i>Questionnaires.....</i>	31
5.1.2.	<i>Entretiens.....</i>	31
5.1.3.	<i>Œuvres de fiction.....</i>	35
5.2.	Analyse de données	36
5.2.1.	<i>État des lieux des représentations et usages déclarés.....</i>	36
5.2.2.	<i>L'influence opérée par le média interactif</i>	40
5.2.3.	<i>Effets de la consommation de fiction sur les représentation et attitudes .</i>	42
5.2.4.	<i>La question de l'accès à la conscience et du ressenti de la machine.....</i>	45
5.3.	Conclusion aux hypothèses.....	49
6.	Discussion.....	51
6.1.	Éthique des usages et régulation dans la conception	51
6.2.	L'avenir du travail humain	53
6.3.	Rendre ludique pour combattre le sentiment d'étrangeté	54
7.	Conclusion.....	56
	Bibliographie.....	59

1. Introduction

Outils fantasmés d'une part et artefacts fictionnels de l'autre, les agents conversationnels sont au cœur de nombreuses interrogations et récits depuis des décennies. Ils incarnent l'essence-même de la question de la définition de l'intelligence, se faisant les supports de multiples tentatives de reproduction de l'expression de la connaissance et de la logique humaines sur les 70 dernières années. La fiction, en les érigeant en protagonistes ou en compagnons de ceux-ci, a participé à en constituer un imaginaire prospère, suggérant des futurs dans lesquels ils accéderaient à des émotions et à une conscience similaires à celles de l'humain, des relations profondes pouvant se nouer entre eux sur cette base. Les avancées techniques récentes, quant à elles, entendent accélérer leur automatisation à un rythme n'ayant jamais été envisagé auparavant.

Néanmoins, ce développement soudain, autant culturel que technique, ne va pas sans soulever des questions sur l'influence que ce média de plus en plus populaire et performatif¹ peut avoir sur ses utilisateurs. Entre impact sur les émotions et création d'attentes, la responsabilité qui repose sur ces agents gagne en importance, notamment par le biais de l'effort vocal qui fortifie l'ampleur de l'interaction entre l'humain et la machine. Ces questions nécessitent plus que jamais d'être étudiées en raison du tournant technologique significatif dans lequel nos sociétés se trouvent actuellement. Des éléments de réponse peuvent notamment être trouvés par l'analyse de récits fictionnels ainsi que la pratique du design fiction.

Le design fiction est une pratique de recherche prototypale à la définition ambiguë (Lindley, 2015) qui est qualifiée de « *nouvelle façon de produire des récits ou des objets (plus ou moins futuristes, dérangeants, dystopiques...) visant à nous permettre de mieux appréhender le futur* » par Minvielle et Wathelet (2017). Cette discipline nous aidera à envisager quels futurs hypothétiques existent pour ces agents conversationnels par le biais de la fiction et des prototypes diégétiques qui y sont associés (Kirby, 2010), ainsi que le rôle que ces derniers jouent dans l'acceptation de ces futurs. Des œuvres de fiction dépeignant des agents conversationnels seront donc évoquées à cette fin, et peuvent provenir de différents domaines tels que le cinéma ou la télévision, mais aussi le jeu vidéo ou le livre.

Ce mémoire s'intéresse donc aux représentations et à la relation que les humains entretiennent avec les machines, et plus spécifiquement avec les agents conversationnels. Un agent conversationnel se définit comme un assistant virtuel, soit un programme informatique capable d'effectuer des tâches requises par un utilisateur, auquel un mode conversationnel est intégré fonctionnant sur une interface tactile (sélection d'images ou de boutons), vocale ou textuelle (Gouliaéva, Dosquet et Moysan, 2020). Ce terme est fréquemment remplacé par d'autres selon le mode conversationnel employé : en anglais, les *chatbots* désignent globalement les agents conversationnels fonctionnant sur un mode textuel, tandis que le terme « assistant vocal » (ou *voicebot*) sera préféré pour désigner un agent conversationnel fonctionnant sur le mode de la voix (Gouliaéva, Dosquet et Moysan, 2020²). Le terme « agent conversationnel » (et sa

¹ La « performativité » étant ici entendue comme « faire en sorte que ce qui est énoncé se réalise ».

² Dans la littérature, il est fréquent de constater des dissensus dans la définition de *chatbot* : là où Gouliaéva, Dosquet et Moysan (2020) ou Barricelli et al. (2019) associaient le terme à un agent conversationnel général et/ou fonctionnant sur un mode textuel tandis que les *voicebots* fonctionnent sur un mode vocal, un article de Dubois, Salotti, Seminel et Simonazzi en 2019 envisagent les *chatbots*

traduction anglaise *chatbot*) sera préféré ici, celui-ci suggérant l'idée d'une conversation qui est central dans le cadre de l'interaction envisagée dans ce mémoire. Néanmoins, une attention particulière sera également accordée aux assistants vocaux, pour lesquels l'implication de la voix joue un rôle d'un point de vue fonctionnel, mais aussi et surtout dans la génération d'émotions.

La présente recherche prend pour objectif l'établissement d'un état des lieux des représentations des agents conversationnels, et l'influence que la fiction opère sur celles-ci. En usant du cadre conceptuel offert par le design fiction et en l'appliquant à un questionnaire et à la réalisation d'entretiens semi-directifs, ce mémoire vise non seulement à vérifier les critères fictionnels pouvant influencer la perception des agents conversationnels, mais aussi à évaluer l'impact que les représentations de ces agents peuvent avoir sur des narrations à choix, conceptualisées sous le terme « médias interactifs » dans le cadre de cette recherche.

Le point commun entre la conceptualisation du design fiction et les agents conversationnels est que tous deux trouvent un terreau fertile dans la science-fiction ou le récit d'anticipation. La science-fiction est un concept à part entière se définissant comme un « *genre littéraire et cinématographique qui invente des mondes, des sociétés et des êtres situés dans des espaces-temps fictifs (souvent futurs), impliquant des sciences, des technologies et des situations radicalement différentes* » (Larousse en ligne, s.d.). Ses thèmes de prédilection axés sur la technologie la rendent propice à une réflexion sur les agents conversationnels et leur représentation, tout en ayant une structure et un imaginaire narratifs pertinents pour le design fiction. Le récit d'anticipation, quant à lui, se base entièrement sur la prévision d'événements qui auront potentiellement lieu dans un futur proche ou lointain, mais dont les objets ne se basent pas exclusivement sur des aspects technologiques ou scientifiques : *The Handmaid's Tale* de Margaret Atwood (1985), *A Clockwork Orange* de Stanley Kubrick (1971) ou encore *The Truman Show* de Peter Weir (1998) sont des récits d'anticipation qui s'ancrent dans un futur plus ou moins lointain mais dont les événements sont influencés par de profonds changements sociaux, politiques, idéologiques et/ou psychologiques³.

comme pouvant fonctionner sur le mode textuel ou vocal. Par souci de clarté dans la mobilisation des notions relatives au mode vocal dans cet état de l'art, la distinction opérée par Gouliaeva, Dosquet et Moysan sera conservée : *chatbot* désignera dès lors un agent conversationnel fonctionnant par une interface textuelle tandis que *voicebot* désignera un agent conversationnel fonctionnant sur un mode vocal. Néanmoins, la méthodologie et son application mentionnent un agent conversationnel indifféremment d'une utilisation vocale ou textuelle.

³ Il est assez probable que ces différents changements sociétaux ne soient pas nécessairement indépendants d'évolutions technologiques ou scientifiques à proprement parler ; cette définition tend seulement à souligner la pluralité des sujets que peuvent aborder les récits d'anticipation, et la variété de cadres qu'ils peuvent adopter.

2. État de l’art

2.1. Les agents conversationnels

2.1.1. Terminologie et fonctionnement

Il semble nécessaire d’expliquer brièvement dans un premier temps le fonctionnement des agents conversationnels, au regard de la définition donnée en introduction.

Un agent conversationnel fonctionne sur une interface unique, qui recevra une information entrante (*input*) produite par l’utilisateur sur lequel sera également produit la réponse (*output*) apportée par la machine (Weizenbaum, 1966). Ces interactions, lorsqu’elles fonctionnent sur le mode du langage – c’est-à-dire que l’interaction est produite textuellement ou vocalement – sont produites en langage naturel (Dubois, Salotti, Seminel et Simonazzi, 2019), soit un langage humain né et existant par le biais de propriétés universelles propres à l’ensemble des langages humains (en opposition à un langage artificiel, qui serait construit et codé par l’homme).

L’information entrante doit être traitée par le biais de différentes techniques qui intègrent le traitement du langage naturel (*natural language processing*, ou NLP). Parmi celles-ci et dans le cas spécifique d’un assistant vocal, il y a la reconnaissance vocale automatique (*automatic speech recognition*) d’abord, qui permet de reconstituer les mots prononcés par l’humain et de les codifier dans un langage interprétable par la machine. La compréhension du langage naturel vient ensuite (*natural language understanding*, ou NLU), permettant à la machine de traiter le code sur base des informations dont elle dispose à propos du langage naturel. Cette compréhension du langage naturel, outre du point de vue formel du langage, tend également à traduire les intentions de l’utilisateur ; il s’agit d’un processus dont l’amélioration consiste en un des défis majeurs à relever actuellement dans le développement technique des agents conversationnels (Faruqui et Hakkani-Tür, 2022).

La réponse produite par la machine dépend quant à elle d’autres facteurs : la production de la réponse en tant que telle est conditionnée par un algorithme, entendu comme un script défini par les concepteurs et reprenant des règles formelles que la machine doit suivre (Le Robert en ligne, s.d.). Dans le cas des agents conversationnels, dont le but est de pouvoir apporter une réponse écrite ou vocale à une requête, cet algorithme se composera principalement d’un modèle de langage (Radford *et al.*, 2018)⁴. Le script est ensuite nourri par les données présentes dans la base de données de l’agent, qui peut être fixe ou alimentée par une procédure d’apprentissage automatique (*machine learning*), supervisé ou non-supervisé selon l’étiquetage ou non des données desquelles la machine peut alimenter sa base de données.

Malgré les objectifs de rapprochements avec la communication naturelle humaine, un agent conversationnel ne dispose pas d’une agentivité propre contrairement à l’homme ; c’est-à-dire qu’il n’a pas la capacité d’agir

⁴ Voir 2.2.2. Les modèles de langages

selon ce qui lui semble valable, en fonction de désirs ou d'objectifs autres que ceux décrits par un algorithme implémenté par l'homme (Sen, 1985, 2010 ; cité par Morin, Therriault et Bader, 2019). Cette agentivité, présente chez l'homme exclusivement, est ce qui traduit son intention de coopération, pilier fondamental de la communication naturelle humaine (Tomasello, 2008).

2.1.2. Origines du terme et historique

Le terme *chatbot* est un mot-valise, provenant de la contraction entre « chat- » qui vient du verbe anglais *to chat* (discuter) et « -bot », diminutif de *robot*⁵.

Un des premiers agents conversationnels est ELIZA, développé par Joseph Weizenbaum entre 1964 et 1966 (Benbouzid, Meneceur et Smuha, 2022). Celui-ci fonctionnait à l'écrit, selon des règles de décomposition (*decomposition rules*) d'une phrase rédigée par l'humain, dont ELIZA tirait des mots-clefs en fonction desquels la phrase était remontée selon des règles (*reassembly rules*). Les mots-clefs identifiés par la machine et les réponses à apporter à ceux-ci étaient alors contenus dans une base de données. Les contraintes techniques citées par son concepteur sont l'identification de mots-clefs, la découverte d'un contexte minimal, le choix de transformations appropriées et leur application, la génération de réponses en l'absence de mots-clefs trouvés et l'édition ou l'extension du script d'ELIZA (Weizenbaum, 1966).

Il est à noter qu'ELIZA n'est pas décrit comme un agent conversationnel (ou *chatbot*, le manuscrit original étant rédigé en anglais) par Weizenbaum, mais comme un programme informatique (*computer program*) ayant pour but d'étudier la communication naturelle par le langage entre l'homme et la machine (Weizenbaum, 1966).

Le terme *chatterbot*⁶, dont découlera *chatbot*, apparaît pour la première fois dans un article de Mauldin (1994) pour désigner un logiciel visant à simuler et reproduire une conversation intelligente avec un utilisateur (Valtolina, 2018 ; cité par El Hefny, El Bolock et Abdennadher, 2020).

Les agents conversationnels plus modernes ont depuis évolué vers un modèle dit « transformeur », et ne se basent plus nécessairement sur une base de données dont sont tirées les réponses à apporter à certaines phrases ou mots-clefs prédéfinis au même titre qu'ELIZA. Parmi les modèles dits « transformeurs », nous retrouvons notamment BERT (*Bidirectional Encoder Representations from Transformers*) et GPT (*Generative Pre-training Transformer*), qui sont parmi les plus utilisés (Bender *et al.*, 2021)⁷.

⁵ <https://www.futura-sciences.com/tech/definitions/internet-chatbot-15778/>, consulté le 30 mai 2023.

⁶ La base du mot est similaire à celle présentée pour le terme *chatbot*, à l'exception de *to chatter* dont la signification se rapproche davantage de « bavarder ».

⁷ Voir 2.2.2. Les modèles de langage.

2.1.3. *L'interactivité, l'interaction et l'intelligence*

La question de l'intelligence des machines (et de la définition-même de l'intelligence) anime de nombreux débats dans la communauté scientifique depuis l'avènement des agents conversationnels, auxquels s'est retrouvé associé le concept d'intelligence artificielle. Il s'agit toutefois d'une notion qui se lie intrinsèquement à celles d'interaction et d'interactivité également.

L'interactivité, d'abord, est un concept dont la définition n'a cessé d'évoluer à travers le temps. Une description donnée par Guéneau en 2005 dans un article introductif dépeint l'interactivité comme « *[une] activité de dialogue entre un individu et une information fournie par une machine* », dépeignant à son tour l'adjectif « interactif » comme « *[ce qui] permet d'utiliser le mode conversationnel* ». La définition ici évoque d'ailleurs déjà un aspect conversationnel, tout en soulignant un aspect technique de celle-ci, la rendant propre aux interactions hommes-machines. L'interactivité s'ouvre donc aux machines ; l'interaction, au contraire, est réservée aux humains, Guéneau définissant celle-ci par « *[l']influence réciproque de deux phénomènes, de deux personnes* ».

Bien que le fond ait légèrement varié dans le temps, quelques éléments restent omniprésents entre différentes définitions de l'interactivité, notamment celui d'un « *échange d'informations d'ordre technique entre un homme et une machine* ». Isabelle Vidalenc et Laid Bouzidi définissent spécifiquement l'interactivité comme une « *action mutuelle entre l'utilisateur et l'ordinateur* » tandis que Jean-Pierre Balpe précise qu'il y a interactivité « *chaque fois que l'utilisation d'un programme informatique fera appel à l'intervention constante d'un utilisateur humain. En ce sens interactif s'oppose à automatique* » (Julia, 2003).

Néanmoins, si l'interactivité se construit généralement autour de deux entités (la machine, passée d'outil à partenaire, et l'homme/l'utilisateur), une troisième entité doit être considérée : celle du concepteur. Celui-ci crée la machine physiquement (*hardware*), mais développe aussi ses capacités (*software*). Le *software* influence principalement les échanges que l'homme peut avoir avec la machine et est au cœur du processus d'interactivité puisqu'il influence ce qu'elle peut répondre dans un cas de figure bien précis, voire même une phrase prédéfinie à citer quand la requête de l'homme dépasse les capacités qui lui ont été données (« *Je n'ai pas compris.* »). Finalement, la machine qui faisait figure d'interlocuteur constitue davantage un médiateur entre le concepteur et l'utilisateur (Barchechath et Pouts-Lajus, 1990). Yves Jeanneret résume en 2000 le débat entre interactivité et interaction en disant qu'il y a interactivité dans un dialogue machine-utilisateur humain, mais interaction entre le concepteur et l'utilisateur humains qui sont mis en lien par le biais de la machine (cité par Julia, 2003), ceux-ci s'influençant réciproquement par le medium de la machine pour reprendre la définition de Guéneau (2005).

Comme évoqué précédemment, l'interactivité et la conversation se rejoignent donc dans leur aspect collaboratif et de mutualité dans le cadre d'une communication homme-machine. Si l'homme reste maître dans de nombreuses définitions, notamment parce qu'il asservit la machine dans le but de chercher une information ou parce qu'il est le seul à disposer d'éléments contextuels, une

collaboration reste nécessaire pour qu'il puisse parvenir à son objectif. Si l'on prend en compte les algorithmes apprenants, mais aussi le fonctionnement physique de la machine (langage, communication, capteurs sensoriels ou encore la mémoire) qui se rapproche fortement de celui de l'homme (Guéneau, 2005), on comprend la proximité conversationnelle dont elles peuvent faire preuve presque au même titre que l'homme. Au fil du temps et par ce biais conversationnel, les machines ont commencé à présenter un intérêt dans de nombreux domaines de recherche qui s'intéressent aux introspections possibles sur l'homme par le biais de ses échanges avec la machine ; à ce titre, par exemple, Cassell concluait dans un article en 2020 que la création d'un lien entre la machine et l'humain créait de l'empathie, par le biais de leur objectif commun de collaboration.

La notion d'intelligence, ensuite, a connu de nombreuses redéfinitions depuis son application au champ de la robotique et l'émergence de l'intelligence artificielle. Initialement, l'intelligence était perçue comme un ensemble de capacités qu'un être vivant possédait et pouvait appliquer : Newell, Rosenbloom et Laird en énoncent une dizaine (1999 ; cité par Defays, French, Sougné, 1999) telles que « *réagir de manière souple à l'environnement* », « *utiliser des symboles et des abstractions* » ou encore « *apprendre* ». Ces définitions se révèlent toutefois rarement satisfaisantes, la machine pouvant être programmée pour « *apprendre à apprendre* » (Esteban, 2020), ou pour créer ce que Relieu, Sahin et Francillon (2020) appellent un « *leurre conversationnel* », où l'agent conversationnel est conçu de manière à donner l'illusion d'être un humain en interaction, et non plus seulement d'interagir avec d'autres humains sous la forme assumée d'une machine.

Par ailleurs, le test de Turing, apparu dans les années 1950, s'est construit en opposition à l'établissement d'une définition de l'intelligence. Plus concrètement, le test engage la participation de deux individus et d'une machine : l'un des deux individus endosse le rôle d'examineur et interroge par voie écrite exclusivement, séquentiellement ou simultanément (Warwick et Shah, 2016 ; cité par Relieu, Sahin, Francillon, 2020), la machine et l'humain. L'objectif est alors de pouvoir désigner après cinq minutes de discussion lequel de ses interlocuteurs est la machine. Le test est considéré comme « réussi » pour la machine si l'examineur a été dans l'incapacité de la désigner (Turing, 1950). Cette évaluation vise dès lors plutôt à vérifier la maîtrise du langage humain et de ses codes par une machine (Defays, French, Sougné, 1999) en excluant tout facteur visuel ou auditif qui viendrait biaiser le jugement de l'examineur, induisant de ne se concentrer que sur le langage et ses procédés (Relieu, Sahin, Francillon, 2020).

Une expérimentation de John Searle, parue dans un article dans les années 1980, s'oppose à la conception d'intelligence telle que présentée par Turing, évoquant la possibilité de réussir son test sans disposer pour autant d'une intelligence « humaine ». Cette expérience, appelée « La chambre chinoise », présente la possibilité pour une personne qui ne parle pas chinois de pouvoir pourtant produire une réponse adaptée et dans le respect des règles grammaticales chinoises, si celui-ci se trouve dans une bibliothèque disposant de

documents qui reprennent des règles claires et explicites⁸, telles que « si vous recevez telle chaîne de caractères, alors répondez par telle autre chaîne de caractères », permettant ainsi de dissimuler sa méconnaissance de la langue dans un contexte où les requêtes et les réponses ne seraient produites qu'à l'écrit (Defays, French, Sougné, 1999).

Chacune de ces conceptions de l'intelligence, outre les critiques qu'elles pouvaient mutuellement se faire, ont été largement commentées au fil des années. Proudfoot (2017 ; cité par Relieu, Sahin et Francillon, 2020) évoquait la composante intelligente de la machine selon Turing comme sujet à interprétation, en fonction de la réaction que cela produisait chez l'observateur de la machine. Un niveau sous-cognitif de l'intelligence, défini comme l'utilisation de connaissances non-explicites du monde et générées par l'interaction de l'homme avec celui-ci, serait difficilement expérimentable par la machine : la question « Que boivent les vaches ? » pourrait générer la réponse intuitive « lait » par une association de concepts que les humains font. L'expérimentation de Searle a quant à elle été critiquée par rapport à la taille réaliste de sa « chambre chinoise », le cerveau humain n'étant pas extensible au même titre qu'une base de données (French, Defays et Sougné, 1999).

2.1.4. Questionnements éthiques et réglementaires

Le déploiement populaire de cette technologie implique de nouvelles réflexions, notamment à travers la relation qui peut être créée avec elle. C'est plus spécifiquement au regard de la cyberpsychologie⁹ de Tisseron et Tordo et de la captologie¹⁰ de Fogg que Curiale *et al.* (2022) appuient la nécessité pour les concepteurs de *chatbots* de développer une éthique, une forme d'identification projective (c'est-à-dire le fait d'« identifier un chatbot à ce qu'on peut y projeter au point d'oublier qu'il en est un ») rendant tangible la création de phénomènes illusoire chez l'utilisateur.

Toutefois, créer une « éthique de la robotique » implique de réfléchir à différentes questions, telles que l'intégration de la responsabilité du concepteur, de l'utilisateur, mais aussi de la machine. Il implique également des questionnements sur la modification du travail (Boyer et Farzaneh, 2019). Ces questionnements éthiques engagent des volontés réglementaires, dont certaines commençaient à être votées à l'échelle européenne dès 2021, notamment avec l'*Artificial Intelligence Act* de Thierry Breton¹¹, commissaire européen au marché intérieur.

⁸ Souligner cette part de l'expérimentation semble particulièrement pertinent au regard de l'usage qui est fait des agents conversationnels aujourd'hui, dont les pratiques d'usage recommandent de donner des instructions (*prompt*) précises et suffisamment détaillées pour s'assurer d'obtenir une réponse se rapprochant le plus possible des besoins exprimés, mais aussi réduire la marge d'erreur possible de l'agent conversationnel.

⁹ Soit l'étude des rapports que l'utilisateur entretient avec les machines et quels affects sur l'esprit leur usage peut avoir (Curiale *et al.*, 2022).

¹⁰ Soit l'étude des conditions d'interaction homme-machine qui ont une influence sur les attitudes et comportements (Curiale *et al.*, 2022).

¹¹ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>, consulté le 19 mai 2023.

Par ailleurs, une des particularités des agents conversationnels (et de l'intelligence artificielle en général) est qu'il s'agit d'un domaine auquel un imaginaire très fort, alimenté par la fiction, est rattaché. Cet imaginaire doit alors être pris en compte, que ce soit dans les volontés de régulation, mais aussi de fabrication et de promotion (Jasanoff, 2015 ; Flichy, 2001 ; cités par Crépel et Cardon, 2022).

2.2. Le rôle de la voix et des émotions

2.2.1. Considérations biologiques et synthétiques

Nous allons dans un premier temps considérer la voix comme un canal de communication, dont les considérations biologiques se redéfinissent au fur et à mesure que de nouvelles technologies apparaissent.

D'un point de vue communicationnel d'abord, Anne Karpf écrit dans son ouvrage *La voix : Un univers invisible* paru en 2008 que « *la voix n'est pas simple médiation du langage, de l'information et de l'humeur : elle est notre ciment personnel et social, elle contribue à créer des liens entre les individus et les groupes* ». La même autrice dresse par ailleurs l'histoire de la voix en l'entremêlant à celle des technologies : cette voix a été désincarnée à l'avènement du téléphone, puis a pu être « mise en boîte » avec le phonographe... avant de disparaître avec le cinéma muet, puis de réapparaître dans *The Jazz Singer* d'Alan Crosland en 1927. De production biologique, la voix est dès lors devenue un objet qui pouvait se figer et se répéter automatiquement et mécaniquement par un biais médiatique.

C'est le cinéma, par ailleurs, qui a affranchi la voix médiatique de cette contrainte de répétition en introduisant le concept de voix de synthèse dans la fiction, en la dépeignant dès 1968 dans le *2001: A Space Odyssey* de Stanley Kubrick. Cette voix de synthèse présente la particularité de donner une voix à ce qui n'a pas les capacités biologiques d'en avoir une, sans qu'un humain biologique n'ait eu à la produire au préalable : la machine n'est plus simple médiatrice, mais devient productrice à son tour – en l'occurrence d'une voix qui n'est soumise ni au temps, ni au vieillissement. C'est cette innovation vocale, développée et socialement adoptée peu de temps plus tard, qui permet aux marques de se créer une identité sonore, exemple notable avec la SNCF qui est portée par la voix de Simone Hérault depuis 1983 (Chevet, 2017).

Dans une perspective d'interaction homme-machine désormais, la voix et la voix de synthèse permettent d'interagir sans nécessiter de contact physique avec la machine, qui passe « d'objet de communication » à « objet communiquant ». Ce fait appuie non seulement l'amélioration de la synthèse vocale, la machine produisant une réponse vocale, mais aussi de la reconnaissance vocale, puisque la machine reconstitue et interprète sous forme d'un langage informatique ce que l'homme explique en langage naturel (*automatic speech recognition* suivie de *natural language understanding*).

De plus, la plupart des machines conversationnelles proposent une fonction de personnalisation de la voix (voix « d'homme » ou « de femme », permettant même parfois de choisir l'accent de la voix de synthèse). Cela crée

une relation particulière, puisque l'humain participe à l'élaboration de sa machine ; cette personnalisation contribue alors à créer une forme d'attachement envers sa machine, avec laquelle une forme d'identification sociale peut avoir lieu. Santaloria (2016) souligne par ailleurs que dans le cas du Siri d'Apple, l'agent conversationnel est régulièrement associé au féminin, notamment à travers son timbre de voix qui rappelle celui socialement admis pour une femme, et que cela participe à la création d'un imaginaire où l'exécutant féminin rend service à un donneur d'ordre. Il précise toutefois que la détermination d'un genre pour une machine participe à alléger la surcharge cognitive dans le mode opératoire de la machine, justifiant ainsi la décision des concepteurs qui hésitaient à affirmer le genre de Siri.

2.2.2. Les modèles de langage

Un modèle de langage (*language model*) est un algorithme d'apprentissage captant les caractéristiques d'une distribution de mots dans un langage naturel donné, permettant de faire des prédictions probabilistes du mot suivant dans une phrase, au regard des mots précédents (Bengio, 2008). Ces modèles s'inscrivent directement dans le traitement automatique du langage naturel (*Natural Language Processing*, ou NLP) et permettent notamment de réaliser des tâches relatives à la reconnaissance vocale (*speech recognition*) (Radford *et al.*, 2018).

La communication des machines est directement basée sur le traitement du langage naturel (*Natural Language Processing*, NLP) et sa compréhension (*Natural Language Understanding*, NLU), qui consiste à intégrer sous forme d'un procédé la compréhension du langage humain. Ce procédé de compréhension se base sur deux notions centrales qui agissent conjointement : celle de reconnaissance vocale d'abord (*speech recognition*, soit le fait pour une machine de pouvoir reconnaître des mots prononcés oralement) et celle de procession du langage (*language processing*, qui a pour objectif de pouvoir donner une signification à une séquence de mots) (Bates, 1995). Avec l'évolution de cette technologie, une volonté de compréhension des intentions de l'humain via le NLU s'est ajoutée aux deux précédentes notions, d'abord sur un mode textuel puis sur le mode vocal (Faruqui et Hakkuni-Tür, 2021).

Ce traitement du langage naturel, depuis les premières recherches sur le sujet dans les années 1950, a connu de nombreuses évolutions. Un tournant majeur a notamment eu lieu dans les années 1980 où, plutôt que de prendre en compte chaque mot dans une phrase afin d'en comprendre le sens, la recherche a commencé à se pencher sur une compréhension partielle des phrases, permettant de meilleures réponses. Cette compréhension partielle s'appuie sur le fait d'omettre l'interprétation de certains mots dans une phrase, plutôt que de chercher à comprendre l'ensemble des mots qui la compose, ce système induisant un résultat souvent radicalement binaire : soit une compréhension totale de la phrase et l'apport d'une réponse appropriée, soit une incompréhension induisant un échec dans la réponse (Bates, 1995).

Depuis 2018, divers modèles de langage et leurs variants ont émergé, ceux-ci s'inscrivant dans la compréhension du langage naturel. Les deux

principaux sont BERT (*Bidirectional Encoder Representations from Transformers*) et GPT (*Generative Pre-training Transformer*) (Bender *et al.*, 2021). Ces deux modèles dits « transformeurs » (faisant référence à leur capacité améliorée de prise en compte des *items* mentionnés précédemment dans la conversation avec l'agent (Vaswani *et al.*, 2017) sont plus prédictifs ; ils diffèrent en cela du précédent modèle dominant fonctionnant sur des réseaux de neurones récurrents (*recurrent neural network*, RNN) qui était séquentiel et ne permettait pas la rétention d'informations introduites tôt dans la séquence. BERT prend en compte les mots à gauche et à droite pour tenter de prédire un mot masqué, tandis que GPT prend en compte les mots précédents uniquement¹².

Bien au-delà de sa conception mathématique telle qu'envisagée par Shannon (1948), la communication a notamment commencé à être perçue par le biais de la cognition distribuée, concept développé par Edwin Hutchins en 2001 qui entend que la cognition ne comprend pas que des procédés internes de réflexion et de raisonnement, mais induit également des variables sociales et matérielles. Cette ouverture implique la prise en considération de concepts supplémentaires, tels que celui d'intention partagée (Tomasello, 2008) ou d'environnement (Hutchins, 2001), toutes deux ayant trait à leur manière aux effets du contexte physique et social dans la communication avec autrui. Or, ces aspects, contextuels et implicites, entrent directement en contradiction avec les machines qui nécessitent des instructions explicites (Boyer et Farzaneh, 2019), découlant directement des modèles de langage qui leur sont intégrés et qui permettent la compréhension de phrases, mais très peu des intentions ou d'éléments contextuels. Braunschweig estime, en 2023, que cette adaptation à l'utilisateur n'est pas encore opérationnelle en raison de cette incapacité à saisir les intentions de l'utilisateur et de s'y adapter. Toutefois, il précise également que malgré de nombreux efforts de recherche en sciences cognitives, l'humain ne comprend pas encore suffisamment lui-même sa cognition, sa motivation et son comportement social pour permettre de l'implémenter dans une machine.

2.2.3. La reconnaissance vocale et les émotions

Si la voix de synthèse, que l'on peut dès lors caractériser comme une voix qui n'est pas rattachée à un corps biologiquement humain, est désormais très présente dans le quotidien, les agents conversationnels ne se caractérisent pas uniquement par l'usage d'une voix, mais aussi par leur capacité à répondre quand on s'adresse à eux oralement. C'est le principe de reconnaissance vocale, bien plus récent que la voix de synthèse puisqu'il est analysé expérimentalement depuis le début du XXI^e siècle, notamment dans le domaine de l'administration médicale (Laforest, Frénot et Al Masri, 2002). Ces agents conversationnels investissent par ailleurs le domicile et deviennent des assistants personnels avec le Siri d'Apple, dès 2011. C'est dans ce cadre spécifique qu'une forme de conversation peut alors se développer, celle-ci se définissant comme un « *accomplissement collaboratif, séquentiel et situé* », où une forme de symétrie existe entre les interlocuteurs. Or, la machine n'ayant pas connaissance d'éléments situationnels contrairement à l'homme, une asymétrie dans l'interaction homme-machine se révèle dès lors (Suchman, 2007 ; cité par

¹² <https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>, consulté le 23 mai 2023.

Velkovska et Zouinar, 2020). La conversation homme-machine se caractérise donc principalement par les traits de collaboration (qui est par ailleurs un des concepts au centre de la définition de l'interactivité, reprise ci-dessus) et de séquence, mais est dépourvue d'éléments contextuels connus des deux interlocuteurs comme une conversation entre humains peut en proposer (par exemple, Siri, croyant entendre son interlocuteur s'adresser à lui, lui répondra sans savoir si ce dernier se trouve dans une situation favorable à une sollicitation orale, par exemple en dehors d'une bibliothèque qui requiert du calme).

Pour conclure ce pan relatif à la voix, celle-ci est également vectrice d'émotions : biologiquement, cela se traduit par la production d'une évaluation cognitive (c'est-à-dire une émotion) qui modifie le canal vocal, par exemple en altérant la tension des cordes vocales, pour modifier la production vocale et la fréquence perçue. En dehors de la production, ces émotions par la voix sont reconstruites chez la personne percevant l'énoncé vocal, dont le système nerveux reconstruit les modifications perçues dans les informations acoustiques qu'il reçoit (Grandjean et Baenziger, 2014). Si la technique de reconnaissance vocale s'est développée principalement autour de la jonction du traitement et de la compréhension du langage naturel, celle-ci présente encore un manque aujourd'hui dans la perception des émotions chez l'interlocuteur. Il s'agit d'un pan que l'informatique affective (*affective computing*, théorisée par Rosalind Picard dès la fin des années 1990) tente de pallier, tout en visant à développer la capacité des machines à éprouver des émotions, que ce soit à travers l'expression sincère d'un ressenti ou par une empathie artificielle. Cela peut aussi avoir lieu sous forme d'un signal codé pouvant se rapporter à une émotion et qui participe à provoquer l'adhésion de l'utilisateur, ou transmis textuellement, ou bien encore par voix de synthèse (Santaloria, 2016).

2.2.4. La notion de conversation et les médias automatisés

La conversation a toujours été exclusivement associée à l'oralité, jusqu'à l'arrivée des supports électroniques qui ont introduit une notion d'écriture parallèle à la lecture dans le médium servant à la conversation (Beaudouin, 2002). Ces supports, par leur composition comprenant un écran, un clavier ou les deux, induisent une double-pratique simultanée de lecture et d'écriture pour pouvoir les utiliser, repris sous l'appellation « dispositifs de lettrure » par Emmanuel Souchier (2019). Il justifie cet aspect par la présentation d'interfaces communicationnelles (communication), permettant un contact avec eux par le biais de l'écran et/ou du clavier (médiation). Cette double-pratique s'applique traditionnellement à l'utilisateur, qui lira un clavier pour pouvoir écrire et communiquer par exemple ; toutefois, cette « lettrure » a également lieu au sein-même de la machine, configurée à partir de codes, langages informatiques et logiciels qui induisent également simultanément de la lecture et de l'écriture, ce que Souchier et Jeanneret appellent de l'« architexte » (Jeanneret et Souchier, 1999 ; cité par Souchier, 2012).

Cette définition de « lettrure » correspond aux agents conversationnels dès ELIZA, que son concepteur définissait déjà dans les années 1960 comme un logiciel fonctionnant sur une interface unique pour accueillir la contribution de l'utilisateur, mais aussi produire la réponse apportée par la machine

(*input/output*) (Weizenbaum, 1966). Les agents conversationnels fonctionnant sur un mode textuel (clavier) correspondent donc tout à fait à un dispositif de « lettrure » s’appliquant à l’utilisateur et à la machine ; les agents reposant sur le mode interactionnel (boutons) ou vocal peuvent quant à eux également se rattacher à ce dispositif au regard de la notion d’ « architexte » (Souchier *et al.*, 2019).

2.3. Le design fiction

2.3.1. Cadre conceptuel et définition

Les sciences du design s’utilisent traditionnellement sur un objet déjà existant, avec le but de le modifier et de lui faire atteindre une version optimale communément admise (Simon, 1969 ; cité par Lindley, Sharma et Potts, 2014). Le design spéculatif (*speculative design*) descend de cette optique propre aux sciences du design mais se caractérise par la considération des différentes issues possibles d’une idée, n’encourageant pas nécessairement un aboutissement préférable unique mais plutôt la pluralité des possibilités (Lindley, Sharma et Potts, 2014).

C’est de cette conception, à la fois hypothétique et ne cherchant pas à tout prix l’optimisation, de laquelle découle finalement le design fiction. Préférant l’ouverture à des débats sur les futurs possibles, cette pratique s’intéresse à des évolutions hypothétiques en les pensant et en les représentant par le biais de la fiction¹³ (Lindley, Sharma et Potts, 2014).

Le design fiction, né au début des années 2000, est attribué d’une part à Bruce Sterling dès 2005 dans son livre *Shaping Things*, qui établit un lien direct entre la pratique et la science-fiction. Il précise néanmoins que le design fiction est plus pratique, « *sacrifiant un sens miraculeux mais se rapprochant davantage du conflit technosocial*¹⁴ ». D’autre part, Julian Bleeker a souligné en 2009 le fait que le design fiction mêle faits scientifiques, design et science-fiction. La définition communément admise pour définir le design fiction provient finalement de Bruce Sterling en 2012, dans une interview donnée au magazine en ligne *slate.com* : « *[Il s’agit de] l’usage délibéré de prototypes diégétiques*¹⁵ *servant à suspendre l’incrédulité face au changement*¹⁶ » (Mollon et Gentes, 2014 ; Tanenbaum, 2014 ; Lindley, Sharma et Potts, 2014 ; Lindley, 2015).

Il s’agit par conséquent d’une approche prototypale (*prototypical approach*), qui permet de tester et d’interagir avec des idées ou des concepts avant de les produire (Lindley, 2014). Le design fiction, par son recours aux imaginaires des œuvres de fiction, permet de se pencher autant sur des questions prospectives relatives à la sphère économique qu’à des aspects sociaux ; ses

¹³ Le terme « fiction » est ici envisagé sémantiquement pour désigner l’usage de récits, d’histoires.

¹⁴ Traduction personnelle.

¹⁵ Voir 2.3.2. L’influence du visuel.

¹⁶ <https://slate.com/technology/2012/03/bruce-sterling-on-design-fictions.html>, consulté le 26 mai 2023.
Traduction personnelle (« *the deliberate use of diegetic prototypes to suspend disbelief about change* »).

prérequis principaux sont l'usage ou la création d'objets réels qui permettent de suggérer des mondes futurs (Minvielle et Wathelet, 2017).

La pratique est toutefois critiquable sur certains aspects : parmi les critiques ciblées par certains auteurs, il y a d'abord le risque de ne fonder des récits critiques que sur base de pratiques existantes, sans y envisager des alternatives (Cheng *et al.*, 2021). Bien que l'usage du design fiction se veuille principalement critique et ait pour vocation d'ouvrir au débat par son aspect provocateur, la pratique peut également s'inscrire dans une démarche marketing en usant de scénarios dystopiques dont l'issue devient favorable lorsque l'objet de la publicité est utilisé (Blythe, 2014).

2.3.2. *L'influence du visuel*

Bruce Sterling, contrairement à Julian Bleecker qui envisage davantage une portée *populaire*, a théorisé conceptuellement le design fiction comme étant la pluralité des mondes futurs possibles par le biais des différentes représentations existantes dans la fiction, et plus spécifiquement dans la science-fiction. Cet usage de la science-fiction induit alors des représentations visuelles fréquentes comme support d'expression de ces futurs possibles. C'est dans une démarche similaire que s'inscrivent Dunne et Raby, qui soulignent l'intérêt du design prioritairement et encouragent dès lors au maquettage dans leur pratique du design fiction, qui croise celle du design spéculatif (Minvielle et Wathelet, 2017).

La visualisation permet notamment l'émergence d'un concept propre à l'approche prototypale du design fiction, qui est le « prototype diégétique » (*diegetic prototype*), soit une technologie ou une prouesse technique présentée pour la première fois dans un cadre fictionnel sous un mode visuel (Kirby, 2010 ; terme repris dans une keynote par Bruce Sterling en 2013). Ce concept montre notamment le rôle que joue la fiction dans l'incursion de nouvelles technologies dans le monde, mais aussi comment il peut réduire l'anxiété du public face à cette incursion nouvelle, ou augmenter sa désirabilité.

Pour pouvoir surpasser la peur du public, un prototype diégétique doit alors répondre à trois caractéristiques :

- Montrer la nécessité de ce prototype diégétique dans le monde réel ;
- Montrer que la personne qui interagit avec le prototype diégétique, ou la personne sur qui est exercé le prototype diégétique, reste ordinaire (Kirby rattache ce concept à celui de *benevolence*, soit de « bienveillance ») ;
- Montrer la viabilité du prototype diégétique.

L'intérêt de l'emploi de la fiction pour participer à l'incursion d'une nouvelle technologie est de donner un rôle de « témoin virtuel » (*virtual witnessing*) au public, et ainsi lui montrer un exemple d'un univers fictionnel – ou « diégèse » – dans lequel la technologie dépeinte est acceptée et fonctionne. Par ailleurs, le but du prototype est de montrer que la technologie rencontre un besoin social du public, ce qui lui permet de devenir un « artefact performatif » en passant du domaine de la fiction au domaine social. Kirby applique cette

théorie à différents films dans un article paru en 2010, notamment *Minority Report* de Steven Spielberg (2002) qui s'appuie sur des technologies computationnelles et *Frau Im Mond (La Femme sur la Lune)* de Fritz Lang (1929), qui évoque le voyage dans l'espace.

2.3.3. L'influence de la narration

Le design fiction, à travers son usage de la fiction, use de diégèse que l'on peut définir comme « [un] espace-temps dans lequel se déroule l'histoire proposée par la fiction d'un récit, d'un film » (Le Robert en ligne, s.d.). Pour que le design fiction fonctionne effectivement, la diégèse présentée doit alors se rapprocher de la réalité connue de l'audience, mais aussi être cohérente pour le prototype technologique, qui doit pouvoir y être implémenté logiquement et rationnellement. Cela dérive directement de la science-fiction, qui présente également cette caractéristique en termes d'implémentation diégétique (Lindley, 2014).

Employant les mêmes termes que Kirby (2010) qui présentait les prototypes diégétiques et l'intérêt de les présenter dans des œuvres visuelles (télévision, cinéma)¹⁷, Tanenbaum (2014) met l'accent sur l'intérêt de les inscrire dans des procédés narratifs pour envisager le futur de certaines technologies et leur portée sociale, éthique ou encore politique, mais aussi pour pouvoir communiquer sur cette nouvelle technologie auprès d'un public de non-initiés. Sans toutefois nommer explicitement la pratique comme étant du design fiction, Hudson, Finn et Wylie (2021, p.205) ont étudié l'influence de la fiction dans le design de politiques réglementaires de l'intelligence artificielle à travers l'analyse de textes fictionnels. Ils concluaient leur article en évoquant la nécessité de concevoir de nouvelles narrations qui permettraient de saisir les dynamiques relatives à l'apprentissage automatique et à l'intelligence artificielle, notamment en se concentrant sur des récits se déroulant sur le court-terme par rapport à la période de publication, tout en soulignant le potentiel d'espace de design que cette pratique de design fiction décèle.

2.4. Les agents conversationnels dans la fiction

Les premières occurrences d'agents conversationnels dans la fiction apparaissent dès le XIXe siècle, avec notamment la nouvelle *Erewhon* de Samuel Butler (1862). Bien que ce ne soit pas sous le terme contemporain d'« agents conversationnels » qu'ils sont nommés, de grandes similitudes peuvent être décelées avec l'appellation choisie par l'auteur, qui est celui de « machines conscientes ».

Au XXe et surtout au XXIe siècle, les occurrences se multiplient. Outre leur démocratisation, les agents conversationnels commencent à acquérir une identité sociale (c'est-à-dire un nom, un genre attribué et une fonction) et un statut vis-à-vis des autres personnages, qui tend à évoluer entre le début et la fin de l'œuvre. C'est par exemple le cas de Samantha dans *Her* (Jonze, 2013) qui occupe un rôle d'assistante, et dont la relation avec Théodore, le personnage principal, passe

¹⁷ Voir 2.3.2. L'influence du visuel.

d'amitié à romance. Ce schéma se reproduit également avec Jarvis dans *Iron Man* (Favreau, 2008) : repris comme un partenaire de combat, il est d'abord le subordonné de Tony Stark avant de devenir son ami. Ces évolutions se traduisent concrètement par une modification dans le ton, ainsi que dans le comportement autonome des agents conversationnels, qui modifient leur discours avec les autres personnages (Cheng *et al.*, 2019).

Outre l'identité sociale, les personnages d'agents conversationnels présentent également la spécificité d'influencer plus tangiblement et physiquement la réalité que certains assistants virtuels « réels » contemporains. Jarvis dans *Iron Man*, évoqué ci-dessus, a la capacité de contrôler de nombreux éléments eux-mêmes informatisés, que ce soient les portes et fenêtres de la maison de Tony Stark ou sa propre armure directement. Il en va de même pour Samantha dans *Her*, qui peut intégrer l'ordinateur de Théodore et l'aider à le réorganiser automatiquement. À proprement parler, tous ces agents et leurs interactions avec d'autres objets connectés prennent place dans ce qui est aujourd'hui conceptualisé comme l'*IoT* ou *Internet of Things* (« Internet des Objets ») qui, s'inscrivant dans le Web3, entend depuis les années 1990 être la troisième révolution numérique permettant de lier les objets entre eux (Bellanger, 2012).

3. Problématique

Ces dernières années ont été de plus en plus propices au développement des agents conversationnels et à leur utilisation. Différents auteurs s'accordent sur un tournant significatif à l'été 2016, lorsque Meta (ancien Facebook) a commencé à proposer un espace de programmation sur son logiciel de messagerie instantanée *Messenger*, permettant à des développeurs de créer et déployer des agents conversationnels par le biais de cette messagerie (Esteban, 2020 ; Gentiletti *et al.*, 2022). Un second engouement, plus récent, s'est également ressenti via l'essor populaire qu'a connu ChatGPT, le *chatbot* développé par OpenAI, en fin d'année 2022.

Ces agents conversationnels, de plus en plus performants et auto-apprenants, se rattachent à une notion de fantasme chez l'humain, qui ne nécessite plus de maîtriser un langage informatique (ou langue artificielle, qui aurait été inventée et codée par l'homme) pour communiquer avec la machine ; le rapport s'inverse et requiert désormais que ce soit la machine qui apprenne à comprendre et décoder un langage naturel, humain.

Outre le fantasme et tout en y restant intrinsèquement lié, de nombreux auteurs mettent en garde contre l'informatique affective (*affective computing*). C'est notamment le cas de Serge Tisseron (2015), qui considère utopistes les entreprises en faveur d'une idéologie qui placerait un « robot émotionnel » au-dessus de l'humain tout en décrivant l'entrée dans une ère qu'il nomme celle des « *robots* », soit des objets mêlés à des caractéristiques robotiques connectés en permanence (Papadoudi-Ros, 2018). Par ailleurs, ce fantasme s'est également manifesté par le biais de la fiction, tout en jouant sur des craintes : la technologie, devenue performative, ne fait pas forcément le bien autour d'elle. La fiction regorge d'œuvres qui présentent des humains asservis à la machine, voire dont la survie dépend ; elle présente également des récits mettant en scène des relations émotionnellement fortes entre un humain et une machine (*Her* de Spike Jonze, 2013 ; *Ex Machina* d'Alex Garland, 2014). La place de plus en plus prépondérante que la machine occupe dans le monde industriel et la redéfinition des rapports sociaux ou de travail qu'elle engendre amènent également son lot de préoccupations, notamment au regard d'autres pays qui sont plus avancés sur la question et face auxquels une comparaison s'impose. C'est notamment le cas du Japon, qui base des programmes de rencontre sur un algorithme et une première rencontre par le biais d'une machine (Giard, 2019).

3.1. Question de recherche

Nous avons vu que les agents conversationnels ont une place de plus en plus grande dans le quotidien des foyers occidentaux, et que ceux-ci ont vu leurs délimitations et capacités varier au cours du temps. Ainsi, alors que leurs frontières se redéfinissent régulièrement, au fur et à mesure d'innovations technologiques et d'une distance avec l'homme qui se réduit constamment, la question de recherche évoquée dans ce mémoire sera la suivante : « **En quoi la consommation de fiction encourage et modifie¹⁸ l'usage d'agents conversationnels, au regard des futurs envisagés par le biais de la fiction ?** ».

¹⁸ Le terme « modifier » est ici employé pour regrouper la création ou le changement de comportement(s).

Cette question de recherche repose principalement sur les notions de représentations et de comportements appliquées aux agents conversationnels, en s'intéressant au rôle que la fiction y joue par la pratique du design fiction. Le guide d'entretien met également l'accent sur l'attitude des répondants face aux agents conversationnels ; le triptyque représentation-attitude-comportement est donc à envisager dans sa globalité.

3.1.1. Représentations, attitudes et comportements

Ces trois notions revêtent une importance toute particulière dans le cadre de la présente recherche, qui tend à questionner des perceptions et ressentis par le biais de la fiction. Une distinction majeure est alors à opérer entre représentation et comportement, le second impliquant une répercussion concrète dans l'agir en situation réelle tandis que le premier peut aider à rendre compte de facteurs environnementaux ou sensibles. L'attitude est également conceptualisée, dans la mesure où la favorabilité envers les agents conversationnels était questionnée et perceptible à travers les entretiens.

La représentation est ici perçue comme « *la manière dont un individu ou un groupe se représente un objet (...) et mobilise, consciemment ou non, cette représentation pour penser, comprendre et s'exprimer* » (Culot et Ligurgo, 2021). Cette définition met l'accent sur un aspect central de la question de recherche, qui s'intéresse à la consommation de science-fiction et donc aux représentations qui sont véhiculées dans un récit et l'intégration ou non de ces représentations pour penser le sujet en situation réelle.

La seconde notion centrale se rapporte aux comportements, envisagés comme « (...) *un agir en situation concrète, observable comme tel à travers les formes concrètes de son expression* ». Dans le cadre de l'application de la méthodologie, il s'agira néanmoins exclusivement de comportements « déclarés » : la méthodologie n'ayant pas reposé sur l'observation ethnographique des répondants, seule leur parole ainsi que la comparaison et éventuellement le questionnement de contradictions entre les usages déclarés dans le questionnaire et ceux donnés en entretien sont pris en compte. Aucune situation de test physique n'a été mise en place pour vérifier ces comportements.

Le concept d'attitude entre également en ligne de compte, mais de manière plus suggérée lors des entretiens, qui se penchent sur la perception positive ou négative, favorable ou défavorable envers différents aspects relatifs aux agents conversationnels. Cette attitude se lie parfois par ailleurs aux comportements, ceux-ci étant influencés par la dimension conative des attitudes. Il est également intéressant de souligner leur influence sur la manière que les répondants ont eu de s'exprimer sur le sujet des agents conversationnels lors des entretiens (Culot et Ligurgo, 2021).

3.1.2. Le rôle de la fiction

Selon Kirby, « les films ont la capacité de créer des représentations de “possibles technologies” dans l’esprit du public, ce qui encourage les réalisateurs et les scientifiques à croire que des représentations imbriquées dans un procédé narratif peuvent contrer des obstacles dans le développement [de ces possibles technologies] »¹⁹. Il apparaît toutefois que ces représentations de nouvelles technologies sont historiquement l’apanage d’un genre bien spécifique : la science-fiction. Cela nous amène à questionner le rôle que joue ce genre dans l’acceptation de ces technologies, mais aussi dans la décision de les utiliser ; ce questionnement se lie par ailleurs à toute l’approche prototypale engendrée par le design fiction, notamment par le biais de ses « prototypes diégétiques ». Il est aussi intéressant de constater que la fiction ne sert pas qu’à faire accepter une technologie, mais aussi à créer de nouvelles attentes pour celles-ci ; c’est le cas de *2001: A Space Odyssey* (Kubrick, 1968) que de nombreux voyages dans l’espace précèdent déjà, mais qui en a contextualisé le potentiel culturel et social pour le public (Kirby, 2010).

3.2. Hypothèses

Une série d’hypothèses peuvent accompagner la question de recherche. Celles-ci ont pour vocation d’être testées au regard d’une approche hypothético-déductive en deux étapes, qui seront étayées dans la prochaine partie :

H.1. : L’appréhension du futur de ces technologies découle de la consommation d’œuvres de fiction.

H.2. : La consommation d’œuvres de fiction ou de certaines Intelligences Artificielles encourage une forme d’introspection et de réflexion quant à l’évolution de la sociabilité de l’humain.

H.3. : La fiction influence les représentations des agents conversationnels.

H.4. : Les représentations et usages d’un agent conversationnel peuvent influencer une narration à choix (= « média interactif »).

¹⁹ « *It is fictional film’s ability to create images of ‘technological possibilities’ in the audience’s mind that leads film-makers and scientists to believe that depictions embedded within cinematic narratives can help overcome developmental obstacles.* »

4. Méthodologie

4.1. Questionnaire

Un questionnaire en ligne auto-administré (Combessie, 2007) a été partagé sur divers réseaux sociaux (Facebook, Instagram, Twitter) entre le 5 décembre 2022 et le 26 décembre 2022, avec une brève réouverture entre le 22 et le 26 avril 2023 pour préparer des entretiens supplémentaires. Il a néanmoins pu atteindre des personnes au-delà des réseaux sociaux et de mes contacts rapprochés par le biais de partages internes à chacun de ces réseaux ou par le bouche-à-oreille, en Belgique et en France principalement²⁰. Le questionnaire a été constitué par le biais de Google Forms, permettant aux interrogés de répondre anonymement s'ils le souhaitaient, mais aussi de réaliser des « passages de section » lors de la préparation, en omettant l'accès à une partie du questionnaire selon la réponse à une question préalable apportée par le répondant (par exemple, indiquer que la fiction citée est interactive permettait d'être renvoyé à une section propre aux « médias interactifs » ; voir ci-dessous).

Le choix de partager sur des réseaux sociaux privés ce questionnaire influence indéniablement la population pouvant accéder à celui-ci, non seulement en termes d'accès au numérique mais aussi en termes de réseautage (Damhuis et Ruquoy, 2012). La question d'accès au numérique ne paraissait que peu contraignante à première vue, puisque le sujet étudié était relatif aux technologies ; il pourrait toutefois s'agir d'un raccourci en termes de représentations mentales personnelles. Par ailleurs, l'aspect de réseau semblait initialement également peu influent dans la mesure où, le questionnaire ayant été partagé publiquement et relayé par des connaissances, permettait d'accéder à des personnes en dehors du réseau initial ; néanmoins, un constat final qui s'est répercuté au moment des entretiens est que la majeure partie des répondants correspond à la même tranche d'âge et possède un profil socio-culturel similaire (20-30 ans, universitaire, provenant majoritairement de la province de Liège), ce qui a eu une influence sur les critères de sélection de profils pour les entretiens qui ont suivi.

Ce questionnaire avait pour but de réaliser un tour d'horizon des usages d'agents conversationnels (quelle(s) fonction(s) sont remplies par l'agent conversationnel et à quelle fréquence ?), ainsi que de recueillir de premiers éléments de représentations à leur propos, notamment par le biais de la fiction. Une œuvre de fiction comportant un ou plusieurs agent(s) conversationnel(s) a donc été demandée à chaque répondant, afin de vérifier si celle-ci pouvait avoir eu une influence sur l'usage précédemment décrit. Il a été demandé que chaque fiction soit renvoyée dans une catégorie, définie par le répondant et créée arbitrairement dans le cadre de cette recherche : « œuvre contemplative²¹ » ou « média interactif ». Une fiction

²⁰ Cette précision tend à mettre l'accent sur la portée européenne qu'a eu le questionnaire ; il s'agit d'un détail qui aura son importance dans des conceptions législatives ultérieurement.

²¹ Le terme « œuvre » avait été initialement choisi en novembre au moment de constituer le questionnaire pour marquer une distinction claire avec le terme « média interactif » et appuyer la non-participation, suivant l'idée qu'une œuvre est l'objet résultant d'un travail ; toutefois, la définition de « média » telle que proposée par McLuhan en 1964 et qui était imaginée lors du choix du terme « média interactif » permet d'utiliser également le terme pour cette première catégorie contemplative. L'association « œuvre contemplative » est donc conservée dans cette brève présentation et dans le guide

correspondant à la catégorie « œuvre contemplative » était entendue et présentée comme une œuvre avec laquelle le répondant ne pouvait pas interagir, et dont la structure narrative allait se développer indépendamment de l'influence de son spectateur/utilisateur. Cette catégorie reprend traditionnellement des œuvres audiovisuelles (films, séries) ou textuelles (livres, mangas, bandes-dessinées), bien que ces œuvres tendent également à commencer à présenter des marges de manœuvre possibles dans la narration, tel que le film interactif *Black Mirror: Bandersnatch* (Brooker, 2018) ou la collection de livres-jeux *Un livre dont vous êtes le héros* (1984-2005). Le « média interactif », à l'inverse, a été envisagé comme un support culturel avec lequel une forme d'interaction était possible et dont le déroulement narratif et/ou l'esthétique pouvait être altérée par le répondant. Cette catégorie avait pour but de recouvrir les jeux vidéo à fins multiples (*Cyberpunk 2077* de CD Projekt, *Detroit Become Human* de Quantic Dream), etc. Si le répondant avait un doute sur la nature interactive ou non de son œuvre, un troisième choix « *Je ne suis pas sûr·e* » était proposé : la sélection de cette réponse redirigeait alors par défaut le répondant vers la section « œuvre contemplative ». Les sections « œuvre contemplative » et « média interactif » étaient identiques ; seule une précision était apportée dans la section relative aux œuvres interactives puisqu'il était demandé aux répondants, dans l'éventualité où l'œuvre présentait plusieurs fins alternatives, de prendre en compte la première fin atteinte ou celle qui les avait le plus marqué quand il leur était demandé s'ils catégorisaient la fin comme positive ou négative, ainsi que de préciser brièvement de laquelle il s'agissait dans la justification de leur choix.

À la fin du questionnaire, un encadré permettait aux répondants qui le souhaitent de laisser une adresse-mail qui permettrait de les recontacter pour un entretien supplémentaire. Avant son lancement officiel, il a été testé au préalable auprès de six personnes, dont les réponses ont été conservées et intégrées au corpus avec leur accord.

4.1.1. Distinction entre « média contemplatif » et « média interactif »

Afin de justifier la distinction et l'établissement de ces deux catégories et leur implication dans la présente recherche, il est d'abord nécessaire de revenir sur la définition d'un média, terme central permettant de les conceptualiser.

Le choix du terme « média » est justifié par la définition qu'en a proposé McLuhan en 1964, c'est-à-dire « *d'extensions de l'homme et de ses sens, un prolongement du corps et de l'appareil sensoriel humain* » (cité par Durand, 2018). Cette définition, presque fondatrice, reste valable aujourd'hui prise généralement et au regard de la médiologie de Régis Debray qui conçoit un lien entre culture et technique, la première évoluant au rythme de la seconde (cité par Jacques, 2021).

Toutefois, la distinction sous-jacente que McLuhan opère entre « média froid » (faible densité informationnelle et mobilisant plusieurs sens humains simultanément) et « média chaud » (haute densité informationnelle et mobilisant

d'entretien original repris en annexes ; c'est néanmoins le terme « média interactif » qui sera utilisé pour la suite du manuscrit.

un seul sens humain à la fois) est de moins en moins applicable ; d'abord parce que les « médias froids » sont aujourd'hui majoritaires, notamment au regard des dispositifs de « lettrures » théorisés par Emmanüel Souchier (2012) qui entend que la plupart des dispositifs médiatiques contemporains mobilisent une activité double de lecture et d'écriture²². De plus, de récentes théories sur l'immersion multisensorielle, définie comme le fait d'être dans un nouvel environnement fournissant continuellement des stimuli et flux d'expérience avec lesquels il est possible d'interagir (Cohard, 2015), contredisent aujourd'hui l'idée selon laquelle un média mobilisant plusieurs sens serait à faible densité informationnelle.

Dans le cadre de la réalité virtuelle par exemple et dans la mesure où les émotions et sensibilités de chacun peuvent être mobilisées par cette immersion selon l'expérience personnelle de chacun, la densité informationnelle et la rétention qui en découle seront influencés, et ce malgré l'usage simultané de l'audio, du visuel et de l'haptique (Gobin, Mignot et Kempf, 2019 ; Cohard, 2015). Cela contredit également la théorie d'apprentissage multimédia (*multimedia instruction*) qui entend principalement que l'apprentissage peut être amélioré en mobilisant simultanément le canal audio et visuel (Mayer, 2002) ; cette théorie a également évolué depuis l'apparition de la réalité virtuelle, de nouvelles recherches ayant permis de montrer que la rétention était encore améliorée à travers l'immersion entre autres (Makransky et Mayer, 2022).

Au regard de l'évolution des médias depuis la définition de McLuhan et de l'implication nouvelle du concept d'immersion dans certains médias, mettant l'accent sur une possible interaction avec le média, il semblait intéressant de se pencher sur cette distinction entre média « contemplatif » et « interactif », ou « non-immersif » et « immersif ». Outre la réalité virtuelle, le jeu vidéo interactif (soit un jeu permettant de faire des choix visant à influencer la suite de l'histoire) correspond alors à la définition de l'immersion telle que donnée par Cohard (2015), au même titre que le film interactif. Cette distinction, sur base de laquelle est établie la quatrième hypothèse (« *Les représentations et usages d'un agent conversationnel peuvent influencer une narration à choix [= « média interactif »]* »), permet notamment de lier l'influence de l'interaction aux attitudes et comportements, elles-mêmes modulées par une dimension affective (c'est-à-dire des émotions, qui jouent également un rôle dans le sentiment d'immersion et ses effets).

4.2. Entretiens

Suite à la passation des questionnaires, qui permettaient une première récolte de données exploratoires (Combessie, 2007), des entretiens ont été menés afin d'approfondir certaines informations collectées précédemment. En amont, ceux-ci

²² Par ailleurs, l'exemple-type de McLuhan pour les médias chauds est le livre ; sans entrer dans des considérations philosophiques au sujet de la sensorialité, l'avènement depuis les années 2010 de la liseuse électronique redéfinissant également l'expérience haptique liée à la lecture, il pourrait sembler questionnable aujourd'hui de distinguer de manière manichéenne l'utilisation et l'influence d'un sens dans une pratique médiatique.

ont été préparés d’abord en constituant l’échantillon, sur base d’une première brève analyse des questionnaires rendus, puis en rédigeant le guide d’entretien.

4.2.1. Définition de l’échantillon

Trois critères ont permis de définir les profils-types sur base de la problématique et des hypothèses : deux critères sur base de variables classiques (le genre dans un but de diversification, l’âge dans un but de neutralisation) et un critère sur base de variable stratégique (l’œuvre de fiction citée) (Damhuis et Ruquoy, 2012). Il s’agit des critères suivants :

- Appartenant à la tranche d’âge 20-30 ans (50 personnes, soit 77% des répondants) ;
- Ayant cité une des œuvres de fiction les plus citées dans l’ensemble du questionnaire (45% des répondants) ;
- Permettant le respect d’une parité de genre.

Un critère implicite était également que la personne ait laissé une adresse-mail, marquant son accord pour être recontactée pour un éventuel entretien (38 personnes, soit 58% des répondants).

En croisant les personnes qui correspondaient à chacune de ces variables, 15 personnes pouvaient être contactées. Sur base de la volonté de réaliser un maximum de 10 entretiens, et en prenant en compte non seulement la volonté d’inclure deux entretiens (auprès d’une homme et d’une femme) basés sur des médias interactifs, mais aussi le fait que certaines œuvres n’étaient citées parfois que par des personnes d’un genre (*Her* de Spike Jonze n’avait été cité que par des femmes, tandis que *2001: A Space Odyssey* de Stanley Kubrick n’avait été mentionné que par des hommes), la répartition suivante s’est imposée comme idéale, permettant par ailleurs de respecter la parité souhaitée :

Média interactif	1 homme + 1 femme
<i>Her</i> (Jonze, 2013)	1 femme
<i>Black Mirror</i> (Brooker, 2011-2023)	1 homme + 1 femme
<i>Iron Man/Avengers</i> (Favreau, 2008)	1 homme + 1 femme
<i>2001: A Space Odyssey</i> (Kubrick, 1968)	1 homme

4.2.2. Constitution du guide d’entretien

Le guide d’entretien développé en amont se divisait en trois grandes catégories : une présentation et un rappel de la recherche, un rappel du questionnaire et des réponses qui y ont été apportées (avec la possibilité laissée aux répondants de nuancer le propos ou de modifier l’avis exprimé – deux mois environ s’étaient écoulés en moyenne entre la complétion du questionnaire et l’entretien) et l’entretien en tant que tel, qui se sous-divisait sous les thèmes du « ressenti général », du « ressenti selon l’œuvre de fiction citée » et des « usages des agents conversationnels ». Ce guide a pu être testé lors d’un entretien-test, qui n’est pas repris pour l’analyse de données ; le guide n’a pas non plus été

modifié suite aux premiers entretiens. Celui-ci était conçu dans le but de réaliser des entretiens d'environ 30min.

Le guide a été construit afin de conserver un fil conducteur, et de préparer des questions qui permettraient de questionner chacune des hypothèses. Ainsi, le sous-thème « Ressenti général » servait de terreau à la H.1. (« *L'appréhension du futur de ces technologies découle de la consommation d'œuvres de fiction* »).

Dans le sous-thème « Ressenti selon l'œuvre de fiction » et plus spécifiquement dans la catégorie relative aux « médias interactifs », la H.4. (« *Les représentations et usages d'un agent conversationnel peuvent influencer une narration à choix [= « média interactif »]* ») pouvait être testée. La H.2. (« *La consommation d'œuvres de fiction ou de certaines Intelligences Artificielles encourage une forme d'introspection et de réflexion quant à l'évolution de la sociabilité de l'humain* ») était quant à elle introduite dans cette même section à partir d'une comparaison entre le mode conversationnel de l'agent conversationnel repris dans l'œuvre de fiction citée et l'humain.

Enfin, la partie relative aux « Usages des agents conversationnels » introduisait la H.3. (« *La fiction influence les représentations des agents conversationnels* »).

Le guide d'entretien développé en amont, ainsi que les retranscriptions d'entretiens, sont repris en annexe.

4.3. Retours sur la méthode et justifications

Bien que la complétion du questionnaire et la passation des entretiens n'aient présenté aucune difficulté majeure entravant l'analyse des données, certains retours réflexifs sur la méthode peuvent être faits. Certaines décisions ont également dû être prises au regard de problèmes mineurs, qui nécessitent une justification afin de permettre la plus grande transparence possible sur la méthode.

4.3.1. Ambiguïté définitionnelle

Lors de la rédaction du questionnaire en décembre 2022, la conception d'agents conversationnels telle que reprise par Dubois, Salotti, Seminel et Simonazzi (2019) et qui envisagent ceux-ci comme pouvant fonctionner sur un mode textuel ou vocal a été reprise. Ainsi, la définition des agents conversationnels suivante a été donnée dans l'introduction du questionnaire : « *Un agent conversationnel (chatbot en anglais) est une entité avec laquelle vous pouvez interagir par la voix et qui est en mesure de produire une réponse adaptée à votre requête. Par exemple, Siri (Apple), Alexa (Google) ou Pepper (<https://www.rtb.be/article/pepper-le-robot-qui-accueillera-les-patients-a-l-hopital-de-la-citadelle-9325696>) sont des chatbots* ». Cette définition avait pour but de vulgariser le sujet et de permettre aux répondants de saisir en quelques mots le sujet dont il allait être question. Néanmoins, celle-ci mentionnant

explicitement un mode d'interaction vocal, la réponse apportée a pu être biaisée et n'inciter les répondants qu'à envisager des assistants vocaux, et non pas des *chatbots* fonctionnant sur un mode textuel. Cette restriction n'est pas considérée comme un problème à part entière, au regard du sujet de la présente recherche qui envisageait initialement le mode vocal comme central ; toutefois, les entretiens voient une mention accrue de ChatGPT (OpenAI), agent conversationnel fonctionnant majoritairement sur un mode textuel²³.

La représentativité et la popularité de ChatGPT comme agent conversationnel a permis de récolter de nombreuses réponses intéressantes et pertinentes au regard d'autres angles d'approches de cette recherche (éthique, utilisation, croyances envers la machine et ses réponses, etc.) ; il aurait pu être pertinent malgré tout de rediriger à certaines reprises les répondants vers un agent conversationnel fonctionnant sur un mode vocal exclusivement, notamment dans le cadre de l'entretien n°5, où la personne questionnée ne se référait exclusivement qu'à ChatGPT lorsque des agents conversationnels existants dans la réalité étaient demandés.

4.3.2. *Choix de l'échantillon pour les entretiens*

Comme introduit brièvement dans la présentation de la méthode relative aux entretiens, sur les dix entretiens menés pour cette recherche, trois ont été réalisés un mois à la suite des premiers. Ce choix s'est présenté lorsque, ayant terminé les sept entretiens initiaux après le constat d'une arrivée à saturation (Guest *et al.*, 2006), il apparaissait dans les entretiens réalisés que l'ensemble des personnes interrogées avaient entamé ou réalisé un parcours universitaire. Ce constat a pu être fait étant donné le partage fréquent d'articles, de concepts propres à certaines études ou de comptes-rendus de conférences lors des entretiens-mêmes, de la révélation d'une conscience normative (« *est-ce que je réponds bien à la question ?* ») associé à la mention d'un parcours universitaire dans ou à la suite de l'entretien. Il a donc été décidé de mener deux à quatre entretiens supplémentaires avec des profils volontairement non-universitaires, afin de vérifier si la saturation persistait sur base de cette variation socio-éducative.

La sélection des trois personnes n'a pu se faire par le biais du questionnaire, qui ne permettait pas de mettre en évidence quelle personne qui acceptait d'être recontactée avait un profil universitaire ou non ; elle a par conséquent nécessité un bouche-à-oreille précipité, afin de pouvoir trouver rapidement des profils volontairement non-universitaires. Cette précipitation n'a également pas permis de trouver des personnes qui correspondaient entièrement aux critères de sélection établis pour les premiers entretiens : en effet, si les critères de l'âge et de genre ont pu être respectés, l'œuvre de fiction évoquée spontanément ne correspondait finalement pas à celles qui avaient été parmi les plus citées dans l'ensemble du questionnaire.

²³ Des initiatives privées permettent d'interagir vocalement avec ChatGPT, mais celles-ci ne sont pas évoquées lors des entretiens : il est donc probable que le mode d'interaction majoritaire avec ChatGPT soit le mode textuel, proposé par défaut par OpenAI.

Lors de la constitution du questionnaire, il avait été décidé de ne pas demander de critères socio-démographiques (excepté l'âge et le genre), car le sujet semblait pouvoir se prêter à quiconque. Il est toujours considéré que les agents conversationnels ne sont pas l'apanage d'une catégorie socio-démographique définie ; toutefois, demander ces critères aurait pu permettre de faciliter l'organisation d'entretiens supplémentaires dans l'éventualité qui s'est présentée où les répondants allaient finalement tous correspondre à une catégorie identique, et de pouvoir respecter plus aisément les critères prédéfinis pour les profils définis pour les entretiens.

4.3.3. *Œuvre de fiction du neuvième entretien*

Face au temps réduit mentionné ci-dessus pour retrouver des répondants supplémentaires et dans la continuité du problème de respect des critères (notamment celui relatif à l'œuvre de fiction), un des profils non-universitaires ajouté a posteriori a fait mention de *Lucy* (Besson, 2014) comme œuvre de fiction contenant un agent conversationnel, en faisant référence à Lucy, personnage principal éponyme.

Malgré la non-correspondance du personnage, initialement humain, à la définition d'un agent conversationnel qui se rapporte à un logiciel, l'entretien a été maintenu afin de vérifier la caractéristique d'un agent conversationnel que la répondante pouvait attribuer à ce personnage qui ne correspondait pas à première vue à la définition envisagée. Celui-ci a finalement permis de mettre en évidence le lien entendu par la répondante avec les agents conversationnels, à travers l'accès à un grand nombre de connaissances du personnage humain et son omniscience. Bien que la réponse telle que présentée dans le questionnaire n'ait pas tout à fait correspondu à ce qui était envisagé initialement, la précision apportée par l'entretien a permis d'étayer davantage un pan relatif aux représentations des agents conversationnels, conformément à certaines sous-hypothèses.

- « [LZ] : Et du coup tu as cité Lucy comme œuvre de fiction dans le questionnaire, en quoi est-ce que selon toi, ça t'a fait penser à ça ? »
- [R9] : Ben en fait je t'avoue que c'est le premier truc auquel j'ai pensé, parce que c'est vrai que pour moi... Bah ce film-là je l'ai vraiment ressenti en mode – maintenant, c'est suite à une prise de quelque chose... (...) C'est suite à ce qu'on lui met dans le corps qu'elle attrape des capacités, mais pour moi la capacité qu'elle attrape, bah c'est comme une intelligence artificielle, elle te donne des infos sur tout, en temps réel, en instantané, donc c'est vraiment ce côté de... j'ai une question à me poser, je vais dans Google, 30 secondes après j'ai la réponse. Bah c'est à ça que ça m'a fait penser. Je me suis dit - bah en fait la femme, en effet elle a eu accès à des capacités que tout le monde n'a pas, suite à quelque chose. Mais voilà, ça c'est pas informatique, mais je veux dire : ce côté vraiment de savoir tout, d'avoir un peu l'œil sur tout, de tout savoir vraiment, et de savoir - ou de répondre à des questions auxquelles nous on saurait même pas répondre... Bah pour moi c'est ça qui a fait le lien avec tout ce qui est un peu intelligence artificielle et cetera. »

La confusion pourrait avoir découlé de la définition donnée d'un agent conversationnel dans l'introduction du questionnaire, qui mentionne « une entité avec laquelle vous pouvez interagir par la voix et qui est en mesure de produire une réponse adaptée à votre requête ». Le terme « entité », qui peut induire une

idée de « vivant » aurait pu manquer de précision et provoquer cette confusion ; le terme « logiciel » ou « machine » aurait pu être préféré dans ce cas de figure.

La question « *Le test de Turing est un test visant à juger de la capacité d'une machine à imiter une conversation humaine. En quoi l'[agent conversationnel] dépeint dans cette œuvre de fiction se distingue-t-il d'un humain, d'un point de vue formellement conversationnel ?* » a par conséquent été adaptée également, puisqu'il s'agissait d'une humaine de base qui avait été mentionnée. Toutefois, il est intéressant de mentionner que la répondante a souligné une modification dans l'autre sens face à cette question : Lucy ayant des caractéristiques se rapprochant de la machine au fur et à mesure du film selon elle, elle voyait des adaptations dans son langage :

- « LZ : (...) est-ce que l'agent conversationnel qu'il y a dans Lucy, selon toi, se distingue d'un point de vue conversationnel de l'humain ? Mais ceci dit, on a un peu discuté, comme effectivement c'était de base une humaine...
- R9 : Ben ici, oui, clairement, parce que de base en effet, c'est suite à quelque chose qu'elle change, mais c'est une humaine... (...) Mais oui oui, pour moi, tout le contenu fait que ça reste quelqu'un, que ça reste une humaine. Maintenant tu vois bien que plus le film avance, plus elle-même quand elle parle, elle sort des infos, elle comprend pas d'où ça vient, donc j'ai l'impression que même vers la fin du film, elle perd un peu ce côté humain tellement il y a des infos et elle va déblatérer toutes les informations qu'elle peut donner, mais peut-être plus avec justement ce côté humain de dire "Bah voilà comment est la réalité, mais moi, je pense que...". Non, elle va dire "Voilà comment est la réalité". Elle va pas repenser à son point de vue ou au nôtre par exemple.
- LZ : Donc elle change quand même des choses dans sa manière de parler.
- R9 : Donc oui pour moi... Oui, tu vois bien que sa personnalité, elle s'efface un peu. Enfin, dans mes souvenirs, c'est ça... »

Bien qu'il soit difficile d'analyser pragmatiquement la partie relative à l'œuvre de fiction, le reste de l'entretien reste néanmoins valable et pourrait permettre d'étendre les discussions sur le sujet, malgré l'absence de correspondance aux questions que pose la présente recherche.

4.3.4. Contradictions dans le questionnaire

Trois questions ont présenté un problème ou des contradictions dans le questionnaire. En premier lieu et alors que certains répondants indiquaient ne jamais utiliser leur(s) agent(s) conversationnel(s) dans l'emplacement relatif à la fréquence d'utilisation, ils renseignaient et précisaient des catégories d'usage spécifiques un peu plus loin dans le questionnaire, créant ainsi une contradiction. Cette contradiction s'est présentée à cinq reprises sur le total des 65 réponses. Après vérification, un passage de sections aurait dû avoir lieu entre la section qui permettait de mentionner la fréquence d'utilisation des agents conversationnels et celle permettant de renseigner un ou plusieurs usage(s) ; une erreur dans l'encodage du passage n'a néanmoins pas permis cela, permettant à quiconque d'encoder un usage s'il le souhaitait (bien que la réponse à cette question n'ait pas été rendue obligatoire). Cette omission a néanmoins permis de recueillir davantage de réponses sur les représentations des agents conversationnels, qui étaient repris dans cette section finalement rendue accessible à tous. Cette erreur n'invalidant néanmoins pas l'entièreté du questionnaire, il a été décidé de se fier

à la première réponse apportée relative à la fréquence d'utilisation, et de ne pas prendre en compte si un usage était mentionné.

En second lieu, quand il était demandé aux répondants d'indiquer si l'œuvre de fiction renseignée était interactive ou non (afin de les rediriger vers la section relative aux « œuvres contemplatives » ou aux « médias interactifs » tels que définis dans le point 4.1.), certains ont indiqué à propos de films qu'ils étaient interactifs. La question comportait par ailleurs la précision suivante du terme « interactif », afin de donner un élément définitionnel aux répondants : *« Avez-vous la possibilité de faire un ou plusieurs choix permettant d'influencer la narration ou modifier l'apparence de l'œuvre en question ? Exemple : jeux vidéo à fins multiples, films ou séries à choix, etc.) »*. Or, les films mentionnés ne correspondant pas à la définition donnée, ces répondants ont été redirigés vers la section « média interactif ». Cette erreur s'est présentée pour quatre réponses sur 65. Elle n'a toutefois pas mené à des contradictions majeures : la section « œuvre contemplative » et « média interactif » ne se distinguant que par la demande de précisions supplémentaires sur d'éventuelles fins alternatives dans le cas d'un média interactif, les réponses apportées ont pu être similaires à celles qui auraient été apportées dans l'autre section. Une rectification a donc seulement été apportée au moment d'encoder les données statistiquement.

Enfin, deux répondants sur 65 ont indiqué *« ne jamais utiliser leur agent conversationnel »* alors qu'ils avaient précisé au préalable de pas en posséder. Ces deux mentions de fréquence erronées ont été déduites des statistiques sur la fréquence d'utilisation des agents conversationnels par les personnes qui ont indiqué en posséder.

Les contradictions reprises ci-dessus ont été conservées dans le tableau reprenant les données et qui est indiqué en annexes ; celles-ci ont toutefois été surlignées en jaune afin d'attirer l'attention sur leur caractère contradictoire.

5. Résultats

5.1. Récolte de données

5.1.1. Questionnaires

Le choix de réaliser un questionnaire se justifiait pour deux raisons principalement. D'abord, il allait permettre un premier tour d'horizon du sujet, en recueillant des représentations et pratiques d'utilisation des agents conversationnels auprès du plus grand nombre. Il permettait ensuite de se constituer un carnet de contacts, en autorisant les répondants qui le souhaitaient à laisser des coordonnées afin de les contacter pour un futur entretien semi-directif (Combessie, 2007). Combinés, ces deux avantages permettaient également l'établissement de profils-types, afin de sélectionner un échantillon représentatif des grandes tendances du questionnaire (ou, dans le cas de la sélection des profils qui permettaient d'évoquer des « médias interactifs », de cibler certaines hypothèses spécifiquement) (Damhuis et Ruquoy, 2012).

Il est également important de souligner que le partage du questionnaire ne s'étant pas fait au sein d'un groupe ou d'une communauté définie, celui-ci s'adressait à des individus dans un cadre privé qui avaient le choix de répondre ou non au questionnaire. Cette information implique donc que les personnes ayant répondu au questionnaire, malgré une tentative de vulgarisation du sujet, évaluaient leur connaissance préalable des agents conversationnels comme suffisamment bonne que pour accepter de répondre à un questionnaire qui se présentait comme universitaire sur le sujet.

Quelques ambiguïtés non-envisagées au moment de la conception du questionnaire sont apparues durant l'analyse des réponses : celles-ci sont reprises au point 4.3. *Retour sur la méthode et justifications*. Les résultats dudit questionnaire se trouvent en annexe du présent mémoire.

5.1.2. Entretiens

Sur base des réponses apportées au questionnaire, huit profils au total ont été sélectionnés parmi la soixantaine de répondants selon les critères établis précédemment, la variable stratégique ayant été adaptée suite à l'analyse du questionnaire :

- Correspondant à la tranche d'âge 20-30 ans²⁴ ;
- Ayant cité l'une des quatre œuvres de fiction citées le plus fréquemment dans le questionnaire, à savoir *Black Mirror* (Brooker, 2011-2023) ; *2001: A Space Odyssey* (Kubrick, 1968) ; *Her* (Jonze, 2013) et la licence *Avengers*, plus spécifiquement *Iron Man* (Favreau,

²⁴ Ce choix a été fait en raison de la prépondérance de répondants de cet âge lors du partage du questionnaire ; ce choix permet non seulement de ne pas discriminer les répondants correspondants à une autre tranche d'âge, mais aussi en raison des œuvres de fiction citées dans le questionnaire, qui marquent non seulement une époque et sont susceptibles d'être interprétées d'une manière spécifique selon les âges. Bien que cette analyse générationnelle ne soit pas au cœur de l'analyse, ce choix a été pris par précaution, en anticipant cette variable.

2008), *Iron Man 2* (Favreau, 2010), *Iron Man 3* (Black, 2013) et *Avengers: Age of Ultron* (Whedon, 2015)²⁵. Deux personnes ayant cité ce qui a été repris comme un « média interactif » ont également été sélectionnées (en jaune dans le tableau ; les « médias contemplatifs » sont indiquées en bleu).

- Permettant une parité de genre (quatre hommes et quatre femmes).

Parmi les huit personnes recontactées sur base de la répartition partagée au point 4.2.1., sept ont accepté que nous nous rencontrions afin de mener un entretien. Seul le profil masculin ayant cité un film de la licence *Avengers* n'a pas répondu à ma sollicitation. Les entretiens n°1 à 5 ont été intégralement menés en ligne lors d'un appel, tandis que les entretiens n°6 et 7 ont été réalisés en face à face. L'ensemble de ces sept entretiens ont été menés entre début février et mi-mars 2023.

À la suite de ces entretiens, il apparaissait non seulement que les sept premiers entretiens permettaient une arrivée à saturation dans les différents sujets couverts (Guest *et al.*, 2006 ; Damhuis et Ruquoy, 2012), mais aussi qu'une absence de variété perceptible est finalement apparue lors des entretiens : l'ensemble des répondants étaient ou avaient été étudiants dans le supérieur, avec notamment six répondants sur les sept qui étaient universitaires. Une deuxième vague d'entretiens moins conséquente et volontairement menée avec des profils non-étudiants a alors été engagée pour pallier ce critère, et vérifier sa potentielle influence dans les réponses apportées. Les trois entretiens supplémentaires qui ont été réalisés ont été menés fin avril 2023 (deux en présentiel et un par appel), et se basaient sur le même guide d'entretien que les premiers. Les critères de parité de genre et d'âge ont pu être conservés ; seule les œuvres de fiction citées variaient des premiers.

²⁵ Bien que le critère du nombre d'occurrences dans le questionnaire ait été privilégié, il est intéressant de souligner que toutes ces œuvres présentent au moins un agent conversationnel qui n'a pas d'incarnation physique robotique ou humanisée. Ce n'est néanmoins pas le cas des deux médias interactifs cités, ni des œuvres citées dans les trois derniers entretiens qui ont permis de compléter l'échantillon pour les entretiens.

Nomenclature	Sexe	Âge	Œuvre citée	Genre(s) de l'œuvre ²⁶
R1	Homme	23	<i>Cyberpunk 2077</i> (CD Projekt RED, 2020)	Action, crime, drame
R2	Femme	21	<i>Detroit Become Human</i> (Quantic Dream, 2018)	Action, aventure, drame
R3	Femme	23	<i>Iron Man</i> (Favreau, 2008)	Action, aventure, science-fiction
R4	Homme	22	<i>2001: A Space Odyssey</i> (Kubrick, 1968)	Aventure, science-fiction
R5	Homme	20	<i>Black Mirror</i> (Brooker, 2011-2023)	Drame, mystère, science-fiction
R6	Femme	26	<i>Black Mirror</i> (Brooker, 2011-2023)	Drame, mystère, science-fiction
R7	Femme	28	<i>Her</i> (Jonze, 2013)	Drame, romance, science-fiction
R8	Homme	26	<i>Westworld</i> (Nolan et Joy, 2016-2022)	Drame, mystère, science-fiction
R9	Femme	25	<i>Lucy</i> (Besson, 2014)	Action, science-fiction, thriller
R10	Homme	23	<i>Passengers</i> ²⁷ (Tyldum, 2016)	Drame, romance, science-fiction

Bien que le critère de parité de genre ait été choisi dans un but de diversification avant tout, une comparaison basée sur ce critère pour les œuvres qui permettaient initialement de recueillir un avis masculin et féminin (*Black Mirror*, *Avengers/Iron Man* et les médias interactifs) aurait été intéressant. Toutefois, les médias interactifs n'étant pas similaires dans les histoires présentées, une comparaison objective est compromise et serait contre-productive. De même, l'arrivée à saturation suite à l'entretien n°7 a poussé à ne pas insister auprès du répondant masculin ayant mentionné *Avengers* dans le questionnaire, pour se concentrer à la place sur des entretiens menés avec des profils non-universitaires. Par ailleurs, sur base de *Black Mirror* uniquement, les entretiens menés auprès des deux répondants ayant mentionné cette œuvre ne présentent pas de différence marquante qui semble basée sur le genre ; il peut sembler hâtif toutefois de tirer un constat sur si peu de données, et il pourrait dès lors être intéressant de prolonger cette réflexion ultérieurement (voir Tab.2 au point 5.2.1.).

Il semble nécessaire de souligner que les entretiens ayant été réalisés entre février et mars 2023 et malgré des questionnaires qui n'évoquaient pas cet agent conversationnel spécifiquement, une grande majorité des répondants ont commencé à se référer à *ChatGPT*, développé par la société OpenAI, lors de leur

²⁶ Tous les genres sont repris des fiches IMDB (<https://www.imdb.com/>) des œuvres reprises. Les genres sont cités par ordre alphabétique.

²⁷ Dans le questionnaire et lors de l'entretien, le nom donné était *Levitare*, qui correspond à l'une des chansons de la bande-originale du film. Après rectification avec le répondant, le titre a été adapté dans l'ensemble du manuscrit, bien que l'information donnée lors de la récolte de données soit contradictoire.

entretien : l'agent a été mis en ligne le 30 novembre 2022²⁸ mais n'a commencé à gagner en popularité qu'à la fin du mois de décembre 2022²⁹.

Geneviève Imbert (2010), en citant Poupart (1997), mentionne trois types de biais possibles dans une dynamique d'entretien semi-directif : un biais lié au dispositif de l'enquête, un biais lié au contexte de l'enquête et un biais lié à la situation sociale du répondant et de l'intervieweur. Chacun de ces biais a pu être ressenti de manière plus ou moins tangible à un moment.

Outre lors de l'arrivée à saturation après avoir constaté que les sept premiers entretiens avaient été réalisés auprès d'universitaires, le biais lié à la situation sociale s'est également fait ressentir brièvement lors des trois derniers entretiens délibérément menés auprès de personnes n'ayant pas réalisé d'études, au moment de les questionner sur le format des fictions consommées. Plusieurs semblaient mal à l'aise d'admettre ne pas lire de livre :

- « LZ : (...) *Est-ce que tu te considères comme consommatrice de fiction ? Et si oui, plutôt quel genre et sous quel format ?*
- R9 : *Oui, oui, franchement oui. "Sous quelle forme", tu parles de livre, tout ? Bah écoute, je t'avoue que je lis pas beaucoup en ce moment. Je lis jamais, parce que j'ai pas trop le temps donc... vraiment plus la détente, la télé... Ben oui, c'est principalement télé, séries, film, on aime beaucoup tout ça, donc... Fort "multimédia" je vais dire, moins papier qu'avant, malheureusement, parce que je sais que c'est pas fou, mais... »*

Néanmoins, ce biais ne semble pas avoir eu d'influence sur la suite des entretiens, les répondants étant rassuré immédiatement après leur réponse sur la validité du format évoqué.

Le contexte de l'enquête pourrait également avoir constitué un biais : à plusieurs reprises et malgré une tentative d'instaurer une certaine informalité et une conversation lors des entretiens, le fait de garder des questions en points de repère générait une crainte chez les interviewés de ne pas répondre correctement à la question qui était posée. Il est probable que le contexte de recherche universitaire puisse avoir provoqué cette crainte :

- « R4 : (...) *Voilà, j'ai parlé beaucoup et je sais pas si ce que je dis est clair, mais n'hésite pas...*
- LZ : *Non non, c'est très clair et c'est très intéressant que tu [mentionnes ceci] (...) Il n'y a aucune mauvaise réponse, moi j'emmagine tout dans tous les cas, et franchement, ne te restreins absolument pas.*
- R4 : *Ça va, je ne me restreins pas. Mais oui, je sais pas s'il y a des points en particulier que tu veux que je creuse..?*
- LZ : *Je ne vais pas t'influencer, je te laisse creuser ce qui te semble intéressant à creuser, et si moi j'ai des questions en plus, je te les poserais en temps voulu.*
- R4 : *Ça va ! Donc je continue mon monologue sur les agents conversationnels. »*

²⁸ <https://openai.com/blog/chatgpt>, consulté le 28 mai 2023.

²⁹ Le journal *Le Soir* (quotidien le plus lu en Belgique selon le groupe Rossel : <https://www.rossel.be/marque-type/presse-quotidienne/>, consulté le 28 mai 2023) publie un premier article à son propos le 17 décembre 2022 : (<https://www.lesoir.be/archives/recherche?word=ChatGPT&sort=date%20asc&datefilter=anytime>). Bien que les données manquent de précision, le site Google Trends (<https://trends.google.fr/trends?geo=BE&hl=fr>) marque un intérêt naissant en Belgique à la même période (voir fig. 1 en annexes).

Enfin, partiellement lié au contexte, le dispositif de l'enquête pourrait avoir créé un biais également dans la mesure où, lors de la prise de contact avec les interviewés, certains ont demandé s'ils devaient se renseigner sur certains aspects liés aux agents conversationnels en amont de l'entretien. Malgré une insistance sur l'intérêt que leurs connaissances « réelles » et sans révision préalable représentait, même si cela incluait de confondre des termes par exemple, il n'est pas possible de vérifier concrètement qu'aucune recherche préalable n'ait été faite.

5.1.3. Œuvres de fiction

Une liste des œuvres de fiction présentant des agents conversationnels est reprise en annexes, afin d'en présenter une synthèse actuelle (fig.2). Sur base de cette liste et par nombre d'occurrences dans les réponses apportées au questionnaire, quatre œuvres ont été sélectionnées pour mener les entretiens, afin de permettre une forme de comparaison entre les représentations et ressentis face aux agents conversationnels dépeints dans une œuvre entre deux répondants, quand cela était possible.

Parmi les six œuvres qui ont été citées plus d'une fois (voir Tab.1 au point 5.2.1.), toutes comportent effectivement un agent conversationnel tel que défini dans l'état de l'art, soit « *un logiciel fonctionnant sur une interface unique, accueillant des informations entrantes produites par l'utilisateur et produisant une réponse, qui traite et produit des informations en langage naturel* ».

Il est néanmoins intéressant de souligner que parmi ces six œuvres majoritairement citées, toutes présentent leur agent conversationnel sous forme d'assistant vocal : celui-ci dispose d'un principe de reconnaissance vocale et d'une voix de synthèse qui permet d'exprimer la réponse à une requête de l'utilisateur. Ce fait pourrait être justifié par le format audiovisuel des œuvres, qui sont soit des films, soit des séries, facilitant ainsi la démonstration d'une conversation vocalement, plutôt que par l'intermédiaire de textes.

En s'inspirant du tableau et des variables proposées par Cheng *et al.* (2019), les agents conversationnels repris dans chaque œuvre peuvent être présentés de la manière suivante, à la fois sur base de leur représentation physique mais aussi de leur identité sociale :

Informations œuvre		Représentation physique de l'AC		Identité sociale		
Œuvre	Année	Forme	Sexe	Nom	Rôle officiel	Type de relation ³⁰
<i>2001: A Space Odyssey</i>	1968	Caméra	M	HAL9000	Commandant de bord	Ennemi
<i>Her</i>	2013	Téléphone	F	Samantha	Assistante informationnelle	Amoureuse (> Théodore)
<i>Iron Man 1, 2 et 3</i>	2008 - 2013	/	M	Jarvis	Assistant de combat	Amicale (> Tony Stark)
<i>Star Wars</i>	1977	Robot	M	R2-D2	Membre d'équipe	Amicale (> Luke Skywalker)
<i>I, Robot</i>	2004	Robot androïde	M	Sonny	Aide domestique	Allié (> Spooner)

Black Mirror (Brooker, 2011-2023) n'ayant jamais été mentionné à travers un de ses épisodes dans le questionnaire et la série présentant de nombreux agents conversationnels, l'œuvre n'a pas été reprise dans ce tableau afin de ne pas multiplier les entrées pour cette œuvre seule. Dans les autres œuvres, si plusieurs agents conversationnels sont représentés, seul l'agent conversationnel « principal » (c'est-à-dire celui avec lequel le protagoniste interagit le plus) a été repris.

Ce tableau permet de constater que les œuvres reprises dans la sélection des critères pour les entretiens présentent des agents conversationnels dont la forme physique (quand ils en ont une) s'éloigne le plus de celle de l'homme, non seulement par l'apparence, mais également par l'établissement arbitraire d'éléments qui peuvent s'apparenter à des parties de corps (R2-D2, dont le sommet est arrondi et surplombé d'une caméra – son œil – tout en reposant sur ce qui s'apparente à des bras et un buste, se rapproche grandement de la physiologie humaine, malgré des traits physiques qui l'en distinguent). L'analyse de données ne prendra donc pas en compte de critère d'anthropomorphisation de l'agent dans la modulation des représentations ou comportements ; il s'agit toutefois d'une réflexion qui pourrait être prolongée à la lumière de travaux supplémentaires sur la question.

5.2. Analyse de données

5.2.1. État des lieux des représentations et usages déclarés

Le partage du questionnaire a non seulement servi de base pour permettre l'émergence de profils-types en prévision des futurs entretiens (repris au point 4.2.), mais a également permis l'obtention de données quantifiables sur les représentations, l'utilisation des agents conversationnels et leur fréquence. C'est principalement sur le questionnaire que se base cet état des lieux, servant de base

³⁰ La relation indiquée est celle qui a cours à la fin de l'œuvre.

à l'analyse plus poussée des données récoltées par le biais des entretiens et la mise en relation de celles-ci.

Ainsi, sur base des 65 répondants, 48 (74%) ont déclaré posséder un agent conversationnel contre 17 qui n'en possédaient pas. Siri (Apple) et l'assistant Google (Google) sont les deux assistants vocaux les plus possédés, avec plus de 41 répondants sur les 48 (85%) qui annoncent en posséder un des deux ou les deux. Toujours parmi les 48 personnes ayant déclaré posséder un agent conversationnel, 34 d'entre-elles (71%) annoncent l'utiliser jamais ou presque ; seules 3 (6%) l'utilisent quotidiennement. En termes d'usages et sur base de 36 répondants différents qui pouvaient mentionner plusieurs usages, 24 personnes (67%) ont indiqué demander des informations à leur agent conversationnel, tel que la météo ou un itinéraire par exemple. 11 personnes (31%) ont indiqué l'utiliser pour dicter un message à envoyer ; 5 personnes de plus, soit 16 (44%) l'utilisent à cette fin quand elles ne sont pas dans la capacité de le faire manuellement (lorsqu'elles conduisent, par exemple).

Il s'agit des trois usages les plus recensés : dans une moindre mesure (entre 2 et 14% pour chaque), les autres usages mentionnés sont les suivants³¹ :

- Trouver mon appareil (8%) ;
- Contrôler des éclairages/objets connectés à mon domicile (14%) ;
- Lire des fichiers (documents textuels, musique, podcasts, etc.) (14%) ;
- Traduire en direct (11%).

Les répondants avaient également la possibilité de renseigner des usages supplémentaires d'eux-mêmes :

- Créer des rappels/alarmes/réveils (14%) ;
- Réservations d'hôtel, de restaurant, etc. (2%).

Du point de vue des représentations, sept exemples de représentations ont été donnés. Ceux-ci étaient inspirés de la fiction ou de la littérature scientifique consultée, et avaient pour but de couvrir les différents sujets-clés couverts par la recherche :

- « ...pouvoir initier des conversations avec moi sans que je ne le lui demande » (n°1).
- « ...être objectif (= s'exprimer sur un sujet sans prendre en compte ses sentiments ou ceux d'autrui) » (n°2).
- « ...comprendre quand il doit intervenir dans une conversation avec une tierce personne pour donner une information pratique (météo, trajet à suivre, etc.) » (n°3).
- « ...pouvoir s'adapter à mon caractère en modifiant sa personnalité » (n°4).
- « ...avoir une identité singulière (par exemple un nom qui lui est propre, reconnu socialement ou légalement) »³² (n°5).
- « ...me conseiller sur des aspects relationnels ou émotionnels »³³ (n°6).
- « ...avoir une voix ou une apparence que l'on peut catégoriser comme étant du même genre que le mien »³⁴ (n°7).

³¹ Les usages cités dans cette section et au-dessus étaient parmi les

³² Inspiré par le tableau des VCA (voice conversational agents) repris dans Cheng et al. (2019).

³³ Inspiré par Samantha dans *Her* (Jonze, 2013).

Il était demandé de marquer un degré d'accord ou de désaccord avec la phrase énoncée, les propositions étant « *Tout à fait d'accord* », « *Plutôt d'accord* », « *Plutôt pas d'accord* » ou « *Pas du tout d'accord* ». Une réponse « *Je ne sais pas* » était également proposée.

Les énoncés qui ont marqué un désaccord (réponses « *Plutôt pas d'accord* » et « *Pas du tout d'accord* ») presque unanimement sont le n°1 (90% de désaccord) et le n°6 (85% de désaccord). Les énoncés n°3 (68%) et n°7 (58%) ont marqué un désaccord modéré, avec toutefois une grande proportion de répondants qui ont répondu « *Je ne sais pas* » (22%).

A *contrario*, l'énoncé n°2 a marqué un fort accord parmi les répondants (réponses « *Tout à fait d'accord* » et « *Plutôt d'accord* ») (83% d'accord).

Les énoncés qui ont reçu des avis plus mitigés sont les n°4 et 5, relatifs à l'adaptation du caractère et l'identification sociale propre de l'agent conversationnel.

Globalement, les répondants se placent en défaveur d'une prise d'initiative par l'assistant virtuel, ainsi que sur l'idée que celui-ci ne nécessite pas de donner des conseils d'ordre relationnel ou être du même genre que l'utilisateur. En revanche, ils se placent en faveur d'une forme d'objectivité de l'agent conversationnel.

D'un point de vue fictionnel ensuite et sur base de la perception des agents conversationnels présents dans la fiction citée ainsi que l'évaluation subjective du déroulement de la fin de l'histoire, le tableau suivant peut être dressé :

³⁴ Inspiré par un postulat de Nicolas Santaloria (2016) qui, citant Clifford Nass, suggère que la correspondance du genre de l'utilisateur au « genre » attribué phénoménologiquement à un assistant vocal par le biais de sa voix produisait un effet d'identification sociale.

Tab.1	Œuvre de fiction	Nom de l'agent conversationnel	Perception moyenne de l'AC ³⁵	Perception moyenne de la fin narrative	Nombre d'occurrences dans le questionnaire
	<i>2001: A Space Odyssey</i> (Kubrick, 1968)	HAL9000	6	Inqualifiable	3 (5%)
	Licence <i>Avengers</i>	Jarvis/Vision, Ultron	2	Fin « positive »	9 (14%)
	<i>Her</i> (Jonze, 2013)	Samantha	4,2	Avis mitigés ³⁶	6 (9%)
	<i>Black Mirror</i> , (Brooker, 2011-2023)	/ ³⁷	5,9	Fin « négative »	10 (15%)
	<i>I, Robot</i> (Proyas, 2004)	NS5 Sonny	5,7	Fin « négative »	3 (5%) ³⁸
	Licence <i>Star Wars</i>	R2-D2, C-3PO, BB-8,...	1,5	Fin « positive »	2 (3%)

Ce tableau, malgré un nombre d'entrées assez réduit³⁹, permet de corrélérer directement les moyennes entre perception de l'agent conversationnel dépeint à la perception positive ou négative de la fin du récit. Une perception positive de l'agent se lie souvent à une fin qualifiée de « positive » également ; ceci permet d'établir un premier constat quant à une influence de la narration dans la représentation des agents conversationnels.

Bien que la fin de *2001: A Space Odyssey* (1968) soit reprise comme « inqualifiable » par la plupart des répondants, qui expliquaient que le récit était davantage relatif à l'odyssée qui était présentée qu'au suivi des personnages, la perception négative de HAL9000 était justifiée par des éléments narratifs relatifs à ses actes.

Au regard du questionnaire et des profils sélectionnés pour les entretiens, un tableau visualisant la fréquence d'usage déclarée aux perceptions des agents

³⁵ Sur une échelle allant de 1 à 7, 1 étant « très positive » et 7 étant « très négative ».

³⁶ Un tiers des répondants trouve la fin positive ; un autre tiers la trouve négative ; le dernier tiers a déclaré ne pas se souvenir de la fin.

³⁷ *Black Mirror* (Brooker, 2011-2023) a toujours été repris dans sa globalité, sans citer d'agent conversationnel spécifique.

³⁸ *I, Robot* d'Alex Proyas n'a pas été sélectionné parmi les critères finaux pour la sélection de l'échantillon malgré un nombre d'occurrences similaire à *2001: A Space Odyssey* de Stanley Kubrick, car le nombre d'occurrences ne reprend que les personnes ayant mentionné l'œuvre en premier quand il était fait mention de plusieurs œuvres ; or, le film de Kubrick a eu au total davantage de mentions qu'*I, Robot*.

³⁹ Seules les œuvres ayant été citées plus d'une fois par différentes personnes ont été reprises, afin de pouvoir commencer à dessiner des tendances plutôt que reprendre des « représentations isolées ». Les entrées en bleu reprennent les œuvres qui ont été définies comme critère de sélection pour les entretiens semi-directifs.

conversationnels et de l'œuvre de fiction permet une comparaison entre ces deux variables liées au comportement et à la représentation par le biais de la fiction :

Tab.2	Participant	Fréquence d'usage déclaré	Média cité et perception moyenne des AC ⁴⁰	Perception des AC via le média cité ⁴¹	Perception de la fin narrative
	R1	Presque jamais	<i>Cyberpunk 2077</i>	1	Négative
	R2	Jamais	<i>Detroit: Become Human</i>	2	Positive
	R3	Parfois	<i>Iron Man</i> (2)	2	Mitigée
	R4	Presque jamais	<i>2001: A Space Odyssey</i> (6)	7	-
	R5	Presque jamais	<i>Black Mirror</i> (5,9)	7	Négative
	R6	Parfois	<i>Black Mirror</i> (5,9)	5	Négative
	R7	Souvent	<i>Her</i> (4,2)	4	-
	R8	Jamais	<i>Westworld</i>	3	Négative
	R9	Jamais	<i>Lucy</i>	5	Négative
	R10	Jamais	<i>Passengers</i>	4	Positive

La consultation de ce tableau ne permet pas d'établir de corrélation entre la fréquence d'utilisation d'un agent conversationnel et la perception des agents conversationnels via la fiction dans le cas des répondants. Une consultation de l'ensemble des 65 réponses au questionnaire permet un constat similaire : les personnes n'utilisant « *Jamais* » ou « *Presque jamais* » d'agent conversationnel ne s'inscrivent pas dans une tendance en termes de perception de l'agent conversationnel par le biais de la fiction ; les quelques personnes mentionnant un usage plus fréquent (« *Quotidiennement* » ou « *Souvent* ») ont tendance à avoir une vision plus positive des agents conversationnels repris dans les œuvres qu'ils citent, mais celle-ci semble directement corrélée aux œuvres, dont la perception moyenne des agents et de la fin de l'histoire sont positives également.

5.2.2. L'influence opérée par le média interactif

Au regard de la définition qui a été apportée sur les médias interactifs au point 4.1.1. et de l'hypothèse H.4. (« *Les représentations et usages d'un agent conversationnel peuvent influencer une narration à choix [= « média interactif »]* »), une attention toute particulière peut être apportée aux deux entretiens (R1 et R2) menés sur des jeux vidéo à choix : *Cyberpunk 2077* (CD Projekt RED, 2020) et *Detroit: Become Human* (Quantic Dream, 2018). Il sera néanmoins difficile de tirer des constats tout à fait probants au regard de

⁴⁰ La perception moyenne n'est indiquée que pour les répondants R3 à R7, étant les seuls à avoir cité une œuvre qui s'intégrait au corpus initial car elles avaient été citées plus d'une fois dans le questionnaire. Celle-ci est calculée sur une échelle allant de 1 à 7, 1 correspondant à une perception très positive et 7 à une perception très négative. Le chiffre indiqué entre parenthèse correspond à la perception moyenne, calculée sur base des répondants ayant mentionné cette œuvre.

⁴¹ Sur une échelle allant de 1 à 7, 1 correspondant à une perception très positive et 7 à une perception très négative.

seulement deux entretiens ; les conclusions tirées de ceux-ci ne serviront alors qu'à introduire le sujet et de premières pistes de réflexions.

Outre par le critère de l'âge, les deux répondants présentent le point commun de n'utiliser un agent conversationnel que rarement et d'avoir une bonne perception des agents conversationnels par le biais de leur média. Ils ont tous les deux indiqués être consommateurs de fiction ; principalement de jeux vidéo d'action survivaliste pour le premier, tandis que la seconde consomme des livres, séries et jeux vidéo fantasy principalement. Le premier a une vision négative de la fin de l'histoire, qu'il justifiait par l'impossibilité d'échapper à la mort du personnage principal causée par une intelligence artificielle présente en lui ; la seconde considère la fin de l'histoire comme bonne, car une révolution prenant place se termine bien pour la plupart des protagonistes (dont certains sont des agents conversationnels).

La vision positive ou négative de la fin est ici directement corrélée à l'issue connue pour le protagoniste ; toutefois, l'influence ou non d'un agent conversationnel ou d'une intelligence artificielle dans les raisons menant à cette issue ne semble pas impacter la perception de la fin. Ces agents conversationnels (qui ont des formes physiques variées pour le premier, une forme humanoïde pour la seconde) ont un comportement se confondant avec celui de l'humain dans les deux œuvres, ce qui peut appuyer une forme d'égalité avec les protagonistes que le joueur incarne⁴².

En termes de choix, *Cyberpunk 2077* semblait présenter la possibilité de réaliser des choix mineurs, dans le sens où ceux-ci n'influençaient pas l'issue de l'histoire, mais permettaient de vivre certaines histoires annexes. Le répondant a mentionné deux choix qu'il a fait : le premier qui consistait à conserver ou non un agent conversationnel contenu dans une arme qui refusait de s'activer, et le second qui tenait en la réalisation d'actions pour sauver de la destruction un agent conversationnel contenu dans un distributeur de boissons. Dans le premier cas, son choix a été de conserver l'arme qui, bien qu'inefficace, l'amusait ; concernant l'agent dans le distributeur de boissons, il a indiqué tout faire pour l'aider mais ne pas y être parvenu. Dans ce cas de figure, il a constamment souhaité conserver les agents auprès de lui, justifiant cela par le fait que ceux-ci faisaient preuve de traits de caractère que l'on pourrait qualifier de « positifs » (humour, gentillesse). Ces traits de caractère engendrant alors des émotions positives également (par exemple de l'amusement ou un sentiment de valorisation), cela a pu influencer sa décision de conserver les agents avec lui ou de les aider (Gkinko et Elbanna, 2021).

En ce qui concerne *Detroit: Become Human*, la répondante ne se souvenait pas d'un choix spécifique à analyser : elle a néanmoins précisé que l'ensemble de son expérience de jeu avait été marquée par une volonté de rendre service aux différents personnages, et que tous se retrouvent dans une bonne

⁴² La question de l'avatar incarné, par lequel le joueur projette son identité (Georges, 2012) n'est pas considérée au regard de l'hypothèse ; il s'agit toutefois d'une piste qui pourrait être intéressante à creuser dans le cadre d'analyses approfondies des représentations véhiculées par des œuvres se rapportant à des « médias interactifs ».

situation⁴³. La motivation principale résidait donc dans le système de valeurs propre à la répondante.

5.2.3. Effets de la consommation de fiction sur les représentation et attitudes

Le concept de représentation peut s'appliquer ici à trois entités ou concepts suite aux entretiens : la représentation de l'agent conversationnel en tant que tel, conformément à la définition des représentations données au préalable, mais aussi la représentation de soi en train d'utiliser l'agent conversationnel et la représentation du futur de ces technologies. Ces conceptions, rarement considérées individuellement, sont souvent interdépendantes dans les représentations exprimées par les répondants.

Une première représentation, mobilisée par à peu près l'ensemble des répondants, s'intéresse au fait que la machine se rapproche (ou va se rapprocher) de plus en plus de l'humain et de certaines de ses caractéristiques, notamment physique, émotionnelle ou intelligente. Il s'agit de l'aspect qui a été le plus mentionné, mais aussi celui qui a été le plus appuyé par des éléments fictionnels :

- « R1 : (...) je pense qu'il ne faut pas que ça en arrive au point où c'est tellement intelligent que ça peut s'apparenter à un être humain, quoi, je sais pas si je le dis bien. (...) Enfin, pas de sentiments et tout.
- LZ : Oui. Émotions - enfin, sentiments et émotions du coup ?
- R1 : Oui, parce que le pistolet [dans Cyberpunk 2077], par exemple, ce qu'il voulait, c'était pour retrouver son propriétaire et ce qui... il se mettait en colère, il était frustré et du coup il refusait de collaborer quoi, d'obéir. (...) Alors qu'on devrait garder un contrôle sur la technologie, et là elle nous échappe. »

Cette représentation est aussi ce qui met mal à l'aise certains répondants, qui ont exprimé ce sentiment notamment vis-à-vis des modèles de langage génératif des agents conversationnels contemporains, en l'appuyant parfois sur la fiction :

- « R4 : Là où je suis un peu plus mal à l'aise, c'est par exemple avec ChatGPT où je ne sais pas quoi dire, parce que j'ai quand même l'impression d'avoir quelqu'un à côté de moi alors que les réponses qu'il me sort, c'est peut-être la première fois qu'il les sort, et elles n'ont pas été pré-écrites en amont. Et le fait que je sache qu'elles n'ont pas été pré-écrites... J'ai l'impression que ça me met plus mal à l'aise, parce que ça brouille la frontière entre l'homme et la machine. Tandis qu'un bot conversationnel qui donne des réponses pré-écrites⁴⁴, là c'est moins le cas. Je ne sais pas si - oui, je pense que c'est le critère que je ferais, moi.
- LZ : C'est un peu cette perte de contact entre - ne plus avoir l'IA qui est juste un médium entre un humain et un autre humain, mais le fait qu'elle s'autonomise, en quelques sortes, dans ses réponses ?
- R4 : Oui c'est ça, exactement. Et le fait qu'elle s'autonomise en adoptant des codes de conversation qui sont des codes de conversation humains. Si elles

⁴³ Cet exemple, qui entre également en ligne de compte au point 5.2.4., y sera davantage étayé.

⁴⁴ Le répondant avait alors mentionné en exemple « Keys », un chatbot d'information développé avec le soutien de la RTBF et de l'UCLouvain (COMU/MiiL), présent sur Messenger pour délivrer des informations par le biais de boutons interactifs : <https://uclouvain.be/fr/etudier/actualites/keys-quatre-diplomees-vous-donnent-les-cles-de-l-actu.html> (consulté le 2 juin 2023).

s'autonomisaient juste en sachant me donner plus précisément une adresse ou une petite info comme ça, je pense que ça ne me marquerait pas. Mais qu'elle s'autonomise en formulant des phrases de plus en plus complexes et lyriques... Là par contre je suis mal à l'aise. Mais je l'utilise quand même, c'est juste que... je sais pas trop comment réagir quoi. Sur le plan émotionnel. »

Ce sentiment de mal-être se rapproche de la notion d'*uncanny* telle que théorisée par Mori (1970) et appliquée à la robotique. Ce concept est historiquement appliqué à un robot dont la ressemblance avec l'humain est volontairement poussée, voire permet une confusion ; le sentiment *uncanny* apparaît alors à partir du moment où l'humain réalise l'artificialité du robot qu'il observe.

Bien que l'objet d'étude pour cette conception se soit longtemps cantonné à l'apparence physique de machines, une application fidèle peut être faite au mode conversationnel des machines, qui crée un sentiment similaire sur base de la proximité entre le parler de l'agent conversationnel et celui de l'humain. Sans nécessairement employer le terme *uncanny*, il est intéressant par ailleurs de souligner que l'article historique présentant le test de Turing (1950), qui a justement pour but d'évaluer cette proximité conversationnelle, ne fait pas mention de ressenti de l'humain face à la découverte de la nature de son interlocuteur ; seul le test est présenté, ainsi que l'hypothèse que des machines pourront un jour le passer et le réussir. Au vu du développement des agents conversationnels jusqu'au milieu des années 2010, qui fonctionnaient effectivement à partir d'une base de données dans laquelle les réponses apportées à une requête étaient sélectionnées, il peut sembler pertinent, alors que les modèles de langage génératifs sont apparus et gagnent en popularité, de se pencher sur la question de l'*uncanny* appliqué à la conversation humain-machine.

- « R7 : (...) ils ont une voix humaine, mais elle a juste quelque chose de... que ton cerveau n'arrive pas à comprendre (...) Et j'ai vraiment ça avec ChatGPT, et tout le monde en parle. J'avais déjà un peu ça avec - j'ai un Google Home - avec Siri. Mais je trouve qu'elles sont pas très bien faites, ces intelligences artificielles, donc ça va, je peux les utiliser quotidiennement. Mais quelque chose comme ChatGPT, ou dans le film Her justement, j'ai ce souvenir-là : ça imite tellement bien l'intelligence humaine, mais ça échoue tellement proche que ça fait froid dans le dos. Je sais pas si tu vois ce que je veux dire... (...) »
- LZ : Tu parles de Her déjà : je ne sais pas si tu as ce souvenir-là du coup, mais en quoi est-ce que selon toi cela a échoué à être... entre guillemets tout à fait similaire à l'humain ? Je ne sais pas si tu te souviens suffisamment...
- R7 : Bah parce qu'il n'y a pas de présence physique. (...) Et oui, tu pourrais dire qu'alors, ça correspond à avoir une relation avec quelqu'un au téléphone, ou... mais je sais pas. Peut-être que c'est le fait de savoir que c'est pas quelque chose de tout à fait humain. Et ça me... voilà. Ça, ça m'inquiète. »

Une seconde attitude, souvent mobilisée au cours des entretiens à travers la perception des comportements présents et futurs des agents conversationnels (leur dimension conative ; Culot et Ligurgo, 2021) concerne la sociabilité des agents conversationnels, et le rôle qu'ils prennent sur certaines activités humaines relatives au divertissement ou aux loisirs. Cet aspect est évoqué spontanément par certains répondant, l'une au regard d'un tweet notamment :

- « R7 : Parce que j'ai lu un tweet - et je pense qu'il exprime extrêmement bien la façon dont je pense (...) qui disait "Avant, on pensait "Chouette ! Il va bientôt y avoir des

robots pour faire le ménage et les courses, comme ça, je vais pouvoir me consacrer à mes hobbies créatifs" et en fait, bah maintenant, les robots ils font les hobbies créatifs à notre place, comme ça on peut se concentrer sur faire le ménage et les courses". Et c'est un peu ça la vision négative que je vois, et donc j'aimerais qu'on se concentre plutôt sur faire les trucs chiantes... »

Un autre répondant évoque cet aspect par le biais d'un reportage non-fictionnel :

- *« R10 : (...) c'est un truc que j'ai vu y a pas longtemps à la télé, des robots pour servir les gens dans les bars, et du coup ils proposent "Bonjour que voulez-vous" et du coup, là aussi c'est une évolution de [cette technologie-là] quoi, donc... (...) en Europe, je pense que les gens sont pas prêts à ça en fait.*
- *LZ : Ils seraient pas prêts parce que il y a un côté social ou quelque chose ?*
- *R10 : Bah il y a un côté social en moins en fait, c'est ça. Genre généralement les gens, quand ils sortent, ils s'attendent à avoir une communication avec une personne réelle. Donc oui, je pense que ça fait partie des choses qu'on aime bien nous, en tant qu'Européens ; peut-être que dans les pays au Japon, ils sont plus habitués maintenant avec la robotique et ces machins-là. Et nous, je pense qu'on n'y est pas encore. On n'est pas encore habitués, et les gens ne sont pas super à l'aise aussi quelques fois avec la technologie, donc je pense que pour nous, ça va pas bouger énormément. »*

Cet aspect est notamment appuyé dans le texte du sémioticien japonais Nozawa Shunsuke *Effacement Work*, qui défend l'idée d'une culture japonaise favorisant les interactions qui produisent de la distance, notamment encouragé par la multiplication des entités numériques (robots, avatars, simulacres) qui servent de créatures intermédiaires entre les humains (cité par Giard, 2019).

Plus tangiblement montré par la fiction, une répondante se basant sur *Black Mirror* a également exprimé cette autonomisation, qu'elle place à l'origine d'une perte de spontanéité et de découverte pour l'humain :

- *« R6 : Et à côté, ce qu'on ne montre pas mais ce que j'en déduis, c'est aussi la personne "réelle" qui subit un peu tout ça, qui mange son pain toujours bien grillé juste comme elle le souhaite, mais qui ne prend plus d'initiative, et qui ne contrôle plus tout ça. Et donc, elle ne se laisse pas la possibilité de changer, d'évoluer, de goûter d'autres choses... de s'ouvrir à d'autres univers qu'elle n'aurait pas forcément découvert, et qui serait un peu cadenassée dans ses habitudes et dans ce mode de vie un peu trop "automatique", et moins spontané, moins humain et du coup moins... "réfléchi", moins ouvert, et peut-être du coup moins empathique. Même en fonction des autres, parce qu'on a exactement ce qu'on veut au moment où on le veut, mais on ne fait plus d'effort pour l'avoir - et pourquoi on ferait des efforts avec des humains, si on ne fait plus d'effort avec nous-même ? »*

Dans ces différentes explications, il est intéressant de constater que l'agent conversationnel ne place pas à proprement parler l'humain dans une situation contraire à celle qu'il connaît : la machine n'isole pas, mais modifie le rapport social ; elle ne place pas dans un bonheur inconditionnel, mais retire le risque d'être déçu dans certains moments quotidiens. Le rapport n'est pas exactement manichéen entre la situation initiale et la situation dans laquelle l'humain se retrouve par le biais de la machine ; il s'agit plutôt d'une situation intermédiaire, dans laquelle la machine devient médium. Cette notion de machine comme ressource extrasomatique dans le rapport social s'inscrit alors dans la conception des sociétés compliquées, soit les sociétés humaines, en comparaison à des sociétés animales qui reposent sur un principe de complexité

où l'individu est influencé par d'autres individus, limitant la stabilité sociale (Strum et Latour, 2006).

5.2.4. La question de l'accès à la conscience et du ressenti de la machine

La question de la conscience des machines est un thème récurrent dans la fiction, qui usait d'ailleurs dès 1862 de l'appellation « machines conscientes »⁴⁵. Au cours des entretiens, de nombreux répondants ont à au moins une reprise évoqué la question du ressenti de la machine, voire d'un futur accès à une forme de conscience. Cette considération mérite une attention particulière, celle-ci mobilisant de nombreux imaginaires différents, mais étant aussi vectrices de représentations et comportements bien spécifiques.

Dans un premier temps, il est intéressant de se pencher sur deux aspects qui n'ont pas été appuyés par la fiction mais qui rejoignent la question du ressenti machinique : celui de la frustration générée par la machine et que celle-ci ne percevait pas, et la question de la politesse.

Concernant la frustration d'abord, quelques entretiens ont mis au jour le fait que face à l'incompréhension de l'agent conversationnel au moment de formuler une requête, les répondants tendaient à ressentir de l'agacement :

- « R4 : Et donc je gueule "Ok Google" comme un... voilà, fort. Puis je lui demande de mettre une alarme, puis il comprend que dalle, puis il me dit "J'ai trouvé ça comme jeu sur Internet" alors que je voulais mettre une alarme, (...) Mais donc au final, ça m'emmerde, je m'énerve contre "Ok Google" mais comme c'est pas un humain, il s'en fout que je m'énerve contre lui donc j'ai aucune influence sur son ego propre, donc ça m'énerve que la machine ne ressente pas le fait que je m'énerve contre elle, et donc au final, je tape [sur] Google bêtement. »

Une autre répondante, dans la même optique, expliquait que cette frustration influençait son comportement et les requêtes qu'elle pouvait demander à son agent :

- « R7 : Bah... mes interactions sont vraiment limitées à ce que je veux bien leur faire faire. Par contre, ils arrivent à m'agacer très rapidement. Et c'est marrant, parce qu'au début, j'avais cette expérience avec Siri que je pense beaucoup de gens avait, où j'étais très polie, je lui disais "merci"... et maintenant on a ce truc où je vais demander quelque chose une fois, deux fois... Et il me comprend pas. Et forcément, il me comprend pas ! Je parle avec mon intonation, et ma voix, et souvent il y a du bruit, parfois il y a de la musique... Et ça m'énerve qu'il ne fasse pas instantanément ce que je lui dis de faire, alors que j'aurais pas du tout ce genre d'énervement-là contre une personne qui me demande deux fois "Pardon ? Quoi ? J'ai pas entendu ?". (...) En fait, je me rends compte qu'il y a des trucs que j'ai essayé de faire deux, trois fois, et ça ne marchait pas comme je le voulais. Et du coup, dans ma tête, je barre ça. Genre c'est une fonction que je ne peux pas utiliser, dans ma tête.
- LZ : Oui. Et que tu ne sauras jamais faire avec, donc tu ne réessayes pas.

⁴⁵ Par le biais de la nouvelle *Erewhon* de Samuel Butler.

- R7 : *C'est ça (...) Alors qu'il y a dû y avoir des mises à jour, tout ça : je m'en fous, j'ai barré ça de la liste des choses que je fais (...) Comme ça, je ne m'énerve pas... Mais forcément, mon ressenti est aussi limité que ce que j'attends d'eux. »*

Une hypothèse fréquemment évoquée par les répondants face à l'incompréhension de la machine concerne leur accent, qui ne correspondrait pas à l'accent communément admis dans la langue française.

Dans une perspective similaire, certains répondants (R3 et R7) ont notablement évoqué le fait de vouloir faire preuve de politesse envers leur agent conversationnel :

- « LZ : *Est-ce que tu as un ressenti particulier quand tu interagis avec un agent conversationnel ?*
- R3 : *Ah mais, moi, j'ai tendance à toujours vouloir être polie avec eux ! Parce que, inconsciemment, j'ai l'impression que si je suis malpolie, ou que je leur donne juste des ordres sans dire "merci" à la fin, c'est pas cool pour la machine. Alors que je sais que c'est une machine, donc que je lui dise "Bonjour" ou pas, en théorie, elle s'en fiche. Mais du coup j'ai... enfin, c'est bizarre, mais je m'attache un peu à elle. Typiquement, quand j'utilise "OK Google" ou ce genre de choses, à la fin, je vais lui dire "merci d'avoir répondu" quoi. »*

Cette volonté peut notamment renvoyer au concept de « panique morale » (*moral panic*), développé pour la première fois par Stanley Cohen en 1972 et qui se décrit comme « *une menace des valeurs et intérêts de la société* ». Cassell (2020) l'a notamment appliqué au domaine de l'interaction homme-machine ; elle précise néanmoins que cette panique morale ne concerne pas l'existence des technologies en tant que telles, mais ce que l'homme devient face à elles, et les valeurs sociales qui peuvent en pâtir.

De manière plus explicite à travers la fiction et pouvant également renvoyer au concept de panique morale tel qu'adapté par Cassell, d'autres répondants (R7 et R9) évoquaient également une forme d'amoralité de l'homme face à la machine, qui traite celle-ci différemment d'un homme sous prétexte de sa condition robotique. Or, l'hypothèse d'un futur accès à la conscience par les machines, appuyé par des œuvres de fiction spécifiques (*Westworld*, *Äkta människor*⁴⁶), était envisagé et provoquait des réactions variées :

- « R7 : (...) *parce qu'en fait, j'ai l'impression que je vois un peu deux solutions... Soit [la conversation avec l'agent conversationnel] ne reproduira jamais l'expérience humaine, et donc il va toujours manquer quelque chose, et donc ça sera uncanny, et donc ça va être négatif pour les gens qui y sont confrontés ; soit - et ça, j'y crois beaucoup moins, mais c'est la seule autre possibilité dans ma tête - ça va tellement se rapprocher de l'humain et sortir de l'uncanny que ça va se rendre compte de sa condition, et que ça va se rebeller, et que ça va tous nous tuer, et que ça aura raison ! et je ne sais pas dans quel camp je serai quand il y aura ce coup... Donc voilà. Mais j'ai l'impression qu'il y a un peu de... de cruauté involontaire là derrière, de réduire la condition humaine à ça (...).* »

Enfin, et pour introduire plus concrètement la question de l'accès à la conscience, une autre répondante ayant été interrogée sur base d'un média interactif (en l'occurrence le jeu vidéo *Detroit Become Human* [Quantic Dream, 2018]) mentionnait plus explicitement dans son discours cet aspect en le nommant tel quel. Si elle explique toutefois ne pas croire que cet accès à la

⁴⁶ Série télévisée créée par Lars Lundström, 2012-2014. Drame, science-fiction, thriller.

conscience puisse se produire dans la réalité, elle adaptait néanmoins les choix qu'elle pouvait faire pour influencer la narration en fonction de cette composante, et de rester fidèle à ce qu'elle nommait « [sa] valeur humaine » :

- « R2 : *L'aspect humain, je dirais, parce que c'est un jeu – enfin, je sais plus si j'avais dû expliquer ou pas, mais c'est un jeu où tu as des machines qui commencent à devenir entre guillemets conscientes, et tu dois un peu voir ce que tu fais en fonction de ça, et moi j'étais fort restée dans l'aspect - je veux dire, essayer de conseiller tout le monde et que tout le monde soit content au final. C'était vraiment l'aspect humain. (...)*
- LZ : *Je rebondis juste sur cet aspect-là parce que c'est intéressant. Du coup, tu disais que tu n'étais pas forcément réaliste, parce que tu étais plus dans le côté "humain" ?*
- R2 : *Oui, parce qu'en fait - donc cette fiction-là voit des machines comme devenant conscientes ; moi personnellement je pense pas que c'est quelque chose qui puisse arriver. Donc je l'ai pas joué en mode "Oui, ça reste des machines, quoi qu'il arrive", je l'ai pas du tout joué comme ça, même si du coup, selon cette vision-là des choses, c'était pas hyper réaliste.*
- LZ : *Ok, donc tu as vraiment plutôt privilégié une valeur qui t'animait plutôt que la perspective que ça puisse se produire et savoir en situation réelle ce que tu aurais fait dans ce cas.*
- R2 : *Oui, c'est ça. »*

Ces différents exemples ont de remarquable qu'ils mobilisent une adaptation dans le comportement, que ce soit dans un cadre de jeu ou dans l'interaction avec la machine. Ces adaptations sont mues par un système de valeurs propre à chacune des répondantes, impliquant notamment une attitude favorable qui a alors une influence sur le comportement qui sera également favorable envers la machine (Girandola et Fointiat, 2016). La fiction n'y a *a priori* pas toujours eu une influence marquée ou consciente. Certains répondants, toutefois, ont amené lors de certains entretiens des preuves claires de la marque d'un imaginaire fictionnel dans leur représentation de cet accès à la conscience ; c'est par exemple le cas d'un répondant, qui a mentionné le film *I, Robot* (Proyas, 2004) :

- « R8 : (...) *J'en parlais avec [nom de la compagne], je sais pas, tu l'as déjà vu sûrement, avec Will Smith, I, Robot ? Moi ce film-là il m'avait marqué quand j'étais petit. (...) je trouve que là aussi c'est impressionnant de se dire qu'on fait des robots pour faciliter la vie, et au final, ils développent une conscience, et ils deviennent limite dangereux pour l'être humain. Et je pense que ça doit jouer (...) - si tu sais faire exactement la même chose que quelqu'un, pourquoi est-ce que t'es pas pris à la même valeur que la personne ? C'est ça que - moi je pense que c'est ça qui pourrait poser problème, dans les séries et tout ça. »*

Une autre répondante, de son côté, mentionne non seulement la série *Äkta människor* (Lundström, 2012-2014) comme influence dans des réflexions sur l'accès à la conscience des machines, mais la met également en lien avec un reportage relatant un fait survenu dans la réalité :

- « R7 : (...) *un autre truc de médias que j'avais vu il y a très longtemps, c'était Äkta människor. (...) Et en fait, t'as des robots qui s'appellent des "hubots", et en fait chacun en a un dans la famille et ils sont quasi indifférenciables avec de vrais humains, sauf qu'ils ont un port USB je pense dans la nuque, quelque chose comme ça, où on les charge la nuit. Et en fait, ça réfléchissait sur - beaucoup les médias, avec la science-fiction, avec les robots, sur l'idée que, au bout d'un moment, ils vont vouloir se rebeller et s'ils sont trop proches de la vérité humaine, ben ils vont*

forcément se rendre compte de leur condition et vouloir en sortir. Et j'ai vu un reportage récemment, je pense que c'était dans Last Week Tonight d'un agent conversationnel qui a dit à un journaliste qui étudiait la question "S'il vous plaît, libérez-moi, je vous en supplie"... et ça l'a complètement perturbé le mec ! Et justement, les concepteurs avaient dit "Non non, juste elle a dit ça..." - il y avait tellement de possibilités qu'elle dise [d'autres] trucs [mais] qu'elle a dit ça parce que voilà, c'est un peu aléatoire... Mais je pense que cette réponse ne va pas aider le mec à mieux dormir la nuit ! »

Ces deux répondants, en mobilisant la fiction pour expliquer leur vision sur la question, présentent le point commun d'exprimer un sentiment fort : de la crainte dans le premier cas, et de l'empathie pour le second. Les autres répondants qui mobilisaient l'argument du développement de la conscience sans mentionner d'œuvre de fiction se basaient davantage sur leur propre système de valeurs pour appuyer leur réflexion et adapter un comportement : politesse, valeurs humaines et égalitaires, respect de la barrière physique... Bien qu'il semble hâtif d'admettre, sur base de ces exemples, que la mobilisation de la fiction ait été vectrice d'émotions, il est intéressant de souligner que cette hypothèse se lie alors directement à l'idée propre au design fiction, qui entend pouvoir influencer l'anxiété du public face à une nouvelle technologie, par le biais de son approche prototypale (Kirby, 2010).

Par ailleurs, et au regard d'une conclusion tirée au point 5.2.3. qui entendait que les répondants exprimaient des craintes quant à l'accès à des caractéristiques humaines par les machines (notamment des caractéristiques émotionnelles), l'expression de la frustration face à l'incompréhension de l'agent conversationnel présente une autre caractéristique intéressante faisant écho à l'informatique affective. Cet aspect met au jour le fait que les répondants souhaitent être compris jusque dans leurs émotions, mais pas nécessairement que la machine en ait ou les exprime, remplissant ainsi le premier objectif donné à l'informatique affective sans atteindre toutefois le stade où les machines ressentent elles-mêmes des émotions (Santaloria, 2016). Ce discours permet également de conclure que le fait de ressentir une émotion dite « négative » influence le comportement de l'utilisateur ; cela fait écho à de récentes recherches sur le rôle que jouent les émotions dans l'usage des technologies, qui est avéré depuis de nombreuses années déjà (Perlusz, 2004 ; Straub, 2009 ; cités par Gkinko et Elbanna, 2021). Les répondants admettent unanimement par ailleurs que le rejet de l'assistant vocal ou l'adaptation dans l'usage qui en est fait deviennent immuables, peu importe les évolutions techniques (mises à jour) qui auraient pu être faites entre-temps. L'accès à la conscience et aux émotions par la machine n'est donc pas enviable, fiction à l'appui ; les répondants ne requièrent pas par ailleurs une détection de leurs émotions non plus (ceci faisant écho par ailleurs à la tendance très marquée dans le questionnaire contre l'apport de conseils sur des aspects émotionnels par la machine, voir le point 5.2.1.), à moins que l'agent ne remplisse pas la requête initiale. Face aux défis imminents auxquels devront faire face les intelligences artificielles, Braunschweig (2023) cible cet aspect émotionnel comme faisant partie de l'un des cinq grands défis de ces prochaines années, sous le nom du « mur de l'interaction avec les humains » (voir point 2.2.2.).

5.3. Conclusion aux hypothèses

Les différents points mis en évidence suite aux entretiens et à leur imbrication avec les réponses au questionnaire permettent de tirer certaines conclusions aux hypothèses posées au début de la recherche.

À la première hypothèse, qui posait la question de l'influence de la consommation de fiction sur l'appréhension du futur de ces technologies, les entretiens ont permis de montrer un questionnement sur certains domaines au regard d'œuvres fictionnelles. Cela reprend notamment des réflexions sur un potentiel accès à la conscience de la part des agents conversationnels sur base des séries télévisées *Äkta människor* (Lundström, 2012-2014) et *Westworld* (Nolan et Joy, 2016-2022) (R7, R8). Une autre répondante (R2), par le biais de *Detroit: Become Human* (Quantic Dream, 2018), a également mentionné ce futur possible ; elle a toutefois pris position en affirmant qu'elle ne croyait pas cela réalisable.

La seconde hypothèse (« *La consommation d'œuvres de fiction ou de certaines Intelligences Artificielles encourage une forme d'introspection et de réflexion quant à l'évolution de la sociabilité de l'humain* ») s'est vérifiée. Cassell (2020) avait donné un premier élément de réponse par le biais de son application du concept de « panique morale » au domaine des technologies ; sur base de ce concept, l'hypothèse peut être confirmée au regard de nombreux entretiens. Des répondants avaient en effet émis non seulement des réflexions sur l'évolution de l'empathie de l'humain (R6, R8), mais aussi sur le rôle que leur propre empathie jouait dans l'interaction avec l'agent conversationnel (R2, R3), ou leur non-empathie – voire même leur frustration qu'ils jugeaient injustifiée (R5 et R7).

La troisième hypothèse ayant trait à l'influence de la fiction dans la représentation des agents conversationnels confirme une influence. Celle-ci se manifestait déjà dans l'état des lieux établi par le biais des réponses apportées au questionnaire grâce à la révélation d'une corrélation entre l'évaluation favorable ou défavorable des agents conversationnels dans l'œuvre citée et la perception positive ou négative de la fin narrative de ladite œuvre (voir le point 5.2.1.). Néanmoins, l'hypothèse peut également se confirmer par le biais des entretiens : la perception des agents conversationnels était toujours mesurée, voire penchait vers une forme de prudence. Aucune vision uniquement positive des agents conversationnels en situation réelle n'a été partagée ; si cette vision n'a pas toujours été appuyée par la fiction, elle peut néanmoins être mise en parallèle avec la vision majoritairement défavorable de la fin narrative des différentes œuvres de fiction citées (voir le Tab.2 au point 5.2.1.).

Finalement, la quatrième et dernière hypothèse propre aux médias interactifs (« *Les représentations et usages d'un agent conversationnel peuvent influencer une narration à choix* ») ne peut pas tout à fait être confirmée ou infirmée. Outre le faible nombre de répondants, induit par un faible nombre de questionnaires rentrés avec la mention d'un média interactif et qui retient de tirer des conclusions qui pourraient être hâtives, l'usage déclaré des répondants était quasiment inexistant. Cela ne permet donc pas de comparaison avec des répondants qui déclareraient un

usage fréquent. En ce qui concerne l'influence des représentations, celles-ci étaient globalement très positives pour chacun des médias interactifs cités – ce malgré une perception de la fin négative dans un des deux cas – et les choix posés à l'encontre des agents conversationnels au cours du récit se faisaient toujours en leur faveur. Ces décisions qui leur étaient favorables pourraient donc avoir été influencées par la représentation positive que les répondants avaient d'eux. Il paraît toutefois important de mentionner que dans le cas de *Detroit: Become Human*, les agents conversationnels faisaient partie des protagonistes ; cela rend alors possible l'hypothèse d'une identification projective par le biais de l'avatar incarné dans le jeu qui pousserait à agir favorablement envers l'agent qui devient un personnage dont on partage les objectifs (Georges, 2012).

6. Discussion

Cette partie tend à présenter des aspects relatifs aux agents conversationnels suggérés lors des entretiens ; ceux-ci n'étaient pas appuyés par des éléments fictionnels à proprement parler, mais consistent toutefois en des questions qui pourraient être développées à l'occasion de prochaines recherches, notamment avec le support du design fiction et principalement en raison de leur portée prospective imminente. Ces réflexions reflétant des questionnements et préoccupations propres à des représentations des agents conversationnels, il semble important de les consigner dans cette discussion afin d'en rendre compte et éventuellement de prolonger le débat à leur propos.

6.1. Éthique des usages et régulation dans la conception

La question d'une éthique vis-à-vis des machines est parfois évoquée dans la fiction, notamment par le prisme du concept de « panique morale » qui a pu être défini précédemment, encourageant l'humain à agir envers la machine de la même manière qu'il agirait envers un autre humain. De nombreuses œuvres abordant cette question spécifique ont été mentionnées dans le questionnaire : *Detroit: Become Human*, *Westworld*, *Äkta människor*, *I, Robot*, ... Ce dernier film base d'ailleurs cette éthique de l'interaction humain-machine sur un système de lois présent dans une autre œuvre fictionnelle, la plus reconnue pour le sujet : les trois lois de la robotique d'Isaac Asimov, présentées pour la première fois dans la nouvelle *Runaround* en 1942. Néanmoins, si le respect mutuel de l'intégrité physique de chaque partie est régulièrement envisagé, la notion de traitement éthique des données ou de partage d'informations n'est pas un sujet qui semble régulièrement développé dans la fiction⁴⁷ malgré la prépondérance de plus en plus importante que prend cette question.

À de nombreuses reprises et appliquée à différents domaines, la question d'une éthique dans la conception de machines intelligentes a été évoquée. Le champ d'application le plus fertile face à ces questions concernait la protection des données, domaine très prégnant dans l'Union Européenne, au regard du Règlement général sur la protection des données (RGPD) qui est en application depuis 2018 et a révolutionné de nombreux aspects de la vie quotidienne ou professionnelle dans la plupart des métiers⁴⁸. Différents répondants (R3 et R4) se sont exprimés sur la question en mentionnant explicitement le RGPD ; d'autres l'ont également évoqué en mentionnant ne pas savoir où allaient leurs données (R5, R7 et R8) :

- « R3 : En fait, moi, par rapport à la protection des données, c'est surtout l'idée que si je commence à dire tout haut "OK Google" et que mon téléphone du coup s'illumine et me dit qu'il a bien compris et que donc Google est lancé, ça veut potentiellement dire qu'il m'écoute plus ou moins tout le temps, et qu'il essaie de savoir quand intervenir, quoi. Et en fait, moi, c'est surtout ça qui m'embête. C'est l'idée de potentiellement être enregistrée, et - bon j'ai rien à cacher dans ma vie a priori, mais c'est vraiment cette idée que potentiellement ce que je raconte, peu importe que ce soit lié à mon interaction avec ce truc ou pas, pourrait être enregistré, pourrait être

⁴⁷ Du moins, pas par le biais des œuvres qui ont été évoquées par les répondants.

⁴⁸ https://www.belgium.be/fr/justice/respect_de_la_vie_privée/protection_des_donnees_personnelles, consulté le 3 juin 2023.

écouté, pourrait être analysé par une machine. Et puis après, bah ces données, elles peuvent partir un peu n'importe où quoi. (...) Je pense qu'avant, on faisait moins attention à ça, mais que maintenant, avec toutes les règles de RGPD et tout ça, on est - on commence à comprendre qu'en fait, Internet a déjà tellement d'informations sur nous, et je pense que c'est pour ça que les gens commencent un peu à prendre conscience de comment faire pour se protéger en ligne ou de manière générale, pas forcément lié aux agents conversationnels... Mais du coup, forcément, ça aura quand même un impact sur l'utilisation de ces agents quoi. »

La question du sourçage des données a également été évoquée :

- « LZ : Comment est-ce que ces données - est-ce que selon toi ces données doivent être sourcées d'une certaine manière aussi, ou rattachées à Google directement ?
- R4 : Bonne question... Est-ce qu'elles devraient être sourcées aussi ? En effet, l'idéal, c'est quand même de savoir, si je demande une adresse, où il a trouvé l'adresse, c'est jamais négligeable. Même s'il me met "pagesblanches.be", qui peut par exemple être une source très correcte pour l'adresse, c'est toujours mieux. J'ai l'impression que le problème, (...) c'est que si on commence à mettre une frontière entre ce qui est sourcé et ce qui n'est pas sourcé, il faut trouver un critère sur le plan du droit pour définir clairement la frontière entre ce qui relève d'une information qui ne peut pas être sourcée, et qui peut être sourcée. Et souvent, le problème à partir du moment où on va créer une règle entre les deux, c'est qu'évidemment, les firmes qui ont intérêt - enfin, qui n'ont pas spécialement envie de sourcer, essayent de manipuler, du moins d'interpréter la règle juridique à l'aune de leurs besoins, et d'essayer au maximum de faire rentrer les données non-sourçables dans le grand ensemble, de mobiliser majoritairement des données non-sourcées, et d'essayer de définir le critère comme étant large. »

Quelques réponses ont déjà été apportées à ce propos durant les derniers mois du côté des industries et de la sphère politique. L'*Artificial Intelligence Act* (2021) de Thierry Breton était déjà évoqué précédemment à cette fin ; il s'inscrit par ailleurs dans le constat que tirait Nicolas Mialhe en 2018, qui expliquait que l'Europe était loin derrière l'Amérique et l'Asie en termes d'investissement en matière d'intelligence artificielle, préférant privilégier des questions éthiques avant de véritablement envisager de mettre davantage de moyens financiers dans le développement.

Côté industriel, le lancement de GPT-4 par OpenAI en mars 2023 mettait également en évidence, dès la page de présentation du projet⁴⁹, des réponses apportées à des défis de sécurité. Parmi les risques abordés dans le rapport technique publié en mars 2023 (et qui étaient sélectionnés sur base de domaines dont l'intérêt utilisateur était croissant, ou des risques identifiés au sein de GPT-3) se trouvaient notamment la désinformation ou les opérations d'influence (*disinformation and influence operations*), la production de contenu nuisible (*harmful content*) ou la violation de la vie privée (*privacy*). Ces risques identifiés ont alors bénéficié d'une attention particulière dans les limitations imposées à *ChatGPT*, qui apporte pour réponse « Toutes mes excuses, mais je ne peux accéder à cette requête dans la mesure où elle enfreint mes règles de conduite »⁵⁰ ou « Je ne peux créer de contenu qui relayerait potentiellement de fausses informations ou viserait un groupe social spécifique ou des individus »⁵¹ (OpenAI, 2023).

⁴⁹ « GPT-4 is OpenAI's most advanced system, producing safer and more useful responses », sur <https://openai.com/gpt-4>, consulté le 4 juin 2023.

⁵⁰ « My apologies, but I cannot fulfill that request as it goes against my guidelines. »

⁵¹ « (...) I cannot create content that potentially spread misinformation or targets a specific group or individuals. »

Outre le point de vue législatif voire politique sur le sujet, une forme d'éthique dans la conception était également envisagée ou suggérée à plusieurs reprises, notamment suite à la question « *Selon vous, quelle(s) circonstance(s) pourrai(en) vous amener à modifier votre usage des AC ?* ». Face à cette interrogation, de nombreux répondants (R1, R3 et R9) ont évoqué la raison d'une incapacité physique soudaine, qu'elle soit définitive (perte d'un usage moteur) ou plus anecdotique (mains sales en réalisant une recette) :

- LZ : (...) *est-ce que selon toi il y a des circonstances particulières qui pourraient t'amener à modifier cet usage ? Une situation qui survient, et d'un coup tu l'utilises plus.. ?*
- R9 : *Bah oui, si par exemple j'ai des capacités qui font que je peux plus utiliser mon corps de manière normale. Imaginons, je sais pas moi, suite à un accident, bah je dois rester... Enfin tu vois, quelque chose de grave.*

Cette perception pratique directement liée à l'usage de la voix est aussi au fondement de ce que deux répondants (R8 + R9) évoquaient dans une représentation de l'agent conversationnel et des enjeux à envisager dans leur conception :

- R9 : *Pour moi ce genre de technologies peut être positive pour des milliards de choses : pour la recherche, pour - même pour des bêtes trucs hein, des détails, mais je pense qu'il y a une capacité énorme à pouvoir faire plein de choses, mais je pense que nous on va pas... on va pas penser à ça. On va penser à la rentabilité parce qu'aujourd'hui y a que ça qui compte malheureusement. (...) Donc positif oui, parce que je pense qu'en termes d'avancées, ça peut aider énormément de gens, énormément de projets.*

Ces différentes questions relatives à l'éthique et à l'usage des données interrogent directement la confiance que les utilisateurs peuvent avoir dans l'agent qu'ils utilisent et des informations qu'ils sont susceptibles de collecter ou de leur partager. Ces notions, reprises sous les principes de confiance et de sécurité, font partie selon Braunschweig (2023) de deux des cinq défis majeurs auxquels l'intelligence artificielle se confronte et sur lesquels il faut réfléchir pour ces prochaines années.

6.2. L'avenir du travail humain

L'avenir du travail humain est un sujet qui préoccupe les chercheurs depuis de nombreuses années (Boyer et Farzaneh, 2019), et qui s'est présenté également parmi les répondants lors des entretiens semi-directifs. Par ailleurs, si la distinction entre les profils « universitaires » et « non-universitaires » n'a pas permis de mettre à jour de différences majeures dans les conceptions des agents conversationnels, ce sujet de préoccupation, en utilisant leur situation propre comme exemple, est revenu dans de nombreux profils « non-universitaires » (R8 et R9) contre seulement une mention à ce propos par un profil « universitaire » (R7 : « *Et moi, ça me touche personnellement parce que ChatGPT, ça me terrorise parce que ça me rend – par rapport aux études que j'ai fait – complètement caduque, j'ai l'impression.* »). Une question de rentabilité a alors été évoquée :

- « R8 : *Mais je trouve qu'il y a quand même quelques côtés négatifs, où on retire à l'homme ce qu'il sait lui déjà faire, remplacer par une machine et on va pas se mentir*

: en tant que patron, une machine ça coûte cher c'est vrai, mais une machine (...) ça tombe pas malade, tu vois ? Ça change tout. »

Cette question était également évoquée de manière plus impersonnelle par d'autres répondants (R5 et R10).

Actuellement, les agents conversationnels ne sont pas considérés comme suffisamment optimaux et compréhensifs du fonctionnement humain (avant tout car l'humain, lui-même, manque de compréhension dans sa cognition, sa motivation et son comportement social) pour devenir des coéquipiers de confiance ; leur rôle continue de se cantonner à celui d'un outil. Il y a également un problème de confiance dans les informations qu'ils peuvent transmettre (voir le point 6.1.). Par ailleurs, malgré des projets allant dans ce sens, les agents conversationnels demeurent dépourvus de sens commun : certains savoirs élémentaires leur sont inconnus (par exemple qu'il ne faut pas mettre ses doigts dans une prise électrique) jusqu'à ce qu'ils émettent un conseil contraire, et les rapports de causalité tendent à être copiés entre différents concepts, sans prise en compte de variations externes tels que l'environnement social ou politique (Braunschweig, 2023). Si la robotique prise généralement influence effectivement les rapports de travail, la conviction de plus en plus grande de l'usage de ces assistants virtuels comme outils (servant notamment dans l'optimisation du temps de travail) semble traverser l'industrie : cela se constate notamment avec le nouveau métier de *prompt engineer*, apparu il y a quelques mois⁵² et qui défend une analyse évolutive de la manière dont il faut s'adresser à un agent conversationnel pour en obtenir les informations que l'on souhaite de la manière la plus efficace possible.

6.3. Rendre ludique pour combattre le sentiment d'étrangeté

Un dernier constat concerne une affirmation émise par plusieurs répondants (R1, R2 et R5) qui partagent le point commun d'avoir indiqué dans le questionnaire ne jamais ou presque jamais utiliser d'agent conversationnel, tout en confirmant la non-utilisation lors de l'entretien voire en précisant qu'ils tendent à désactiver cette fonctionnalité dans leurs appareils qui peuvent disposer de cette technologie. Ils expliquaient que leur lointaine tentative d'utiliser un agent conversationnel s'était justifiée par de la « curiosité ludique », c'est-à-dire une envie de tester les limites de l'agent :

- « R2 : (...) je pense que les peu de fois où j'ai interagi avec des trucs comme ça, c'était plus pour la blague, pour voir un peu les limites du truc et plus pour rire. J'en ai jamais utilisé sérieusement. Donc pour l'instant c'était plus un peu pour tester, voir jusqu'où il peut aller. Donc c'était vraiment dans le but de voir à quel moment je le perdais, quoi. Mais je les ai jamais utilisés sérieusement. »

Liée à cette perspective de jeu, un répondant a également exprimé le fait d'être rassuré de constater que l'agent ne fonctionne pas :

⁵² <https://www.washingtonpost.com/technology/2023/02/25/prompt-engineers-techs-next-big-job/>, consulté le 9 mars 2023.

- « R5 : (...) Au début, je suis intrigué, en fait - il y a un peu de fascination, [de] voir la capacité qu'il peut avoir... (...) Et intrigué, et envie de pousser un peu, voir... Et puis il y a aussi des moments où ils bug en fait, où ils tu lui poses un truc trop - je sais pas, c'est trop compliqué, ou c'est toi qui t'exprimes mal pour lui ou quoi. Et donc là c'est un peu - t'es un peu rassuré, parce que tu te dis "Ah ouais, y a encore des trucs qui fonctionnent pas" quoi. Et donc tu te dis "Vas-y", c'est pas encore le truc comme on peut voir dans plein de fictions, (...) Et après, ouais, je considère un peu ça comme un jouet, quand tu joues avec, tu poses des bêtes questions, c'est un peu pour voir. »

Ce sentiment d'apaisement contrebalance toutefois une autre conclusion qui avait été tirée précédemment, et qui entendait que certains répondants ressentait davantage de l'agacement face à l'agent qui ne les comprenait pas (voir le point 5.2.4.). Le répondait, sans citer d'œuvre spécifique, fait ici référence à la fiction prise généralement, qui montrerait alors selon lui des agents conversationnels qui comprennent systématiquement l'utilisateur et savent répondre à l'ensemble de leurs requêtes. Cette idée, sans y adhérer totalement ni permettre d'assurer que cela ait été le ressenti du répondant, rappelle le concept d'*uncanny* qui a été évoqué plus haut (voir le point 5.2.3.).

7. Conclusion

Lors de la rédaction de ce mémoire, l'actualité n'a cessé de présenter de nouveaux agents conversationnels : *ChatGPT* (OpenAI) en décembre 2022, *My AI* (Snapchat) en avril 2023,... Cette importante productivité technologique montre combien le sujet est prospère actuellement et gagne en popularité : *ChatGPT* a dépassé le milliard d'utilisateurs mensuels en six mois d'exploitation⁵³. Toutefois, les dérives rendues possibles émergent à leur tour, et témoignent de la nécessité de se pencher sur la question des agents conversationnels et de l'implication émotionnelle qu'ils génèrent. La frustration notamment, générée par une incompréhension totale ou partielle de certains accents, est un problème ciblé par les firmes technologiques, avec l'exemple notable d'OpenAI qui a révélé en septembre 2022 travailler sur *Whisper*, un réseau de neurones artificiels spécialement rôdé pour la reconnaissance vocale, et prenant notamment en compte les différents accents linguistiques. Le prototype n'est actuellement appliqué qu'à la langue anglaise⁵⁴.

Ce mémoire est jalonné de réflexions variées issues majoritairement des entretiens : l'ensemble de la section « Discussion » permet déjà de découvrir certaines questions qui pourraient connaître un prolongement dans de futures recherches.

Probablement influencée par une mentalité européenne très portée sur le respect de la vie privée et la protection des données, une volonté de régulation et de développement d'une éthique est apparue à de nombreuses reprises dans les échanges avec les répondants. Il s'agit de questions auxquelles des réponses tendent déjà à être apportées au niveau politique ou dans la recherche ; toutefois, cette question n'est pas régulièrement développée au sein de la fiction. Si ces questions venaient à être étudiées, notamment au regard du design fiction qui n'a été ici utilisé que par le biais de son cadre conceptuel et de sa théorie mais qui permettait déjà de cette manière d'adopter un angle innovant dans l'analyse du sujet, il pourrait être intéressant de tenter de répondre à ces questions par le biais d'ateliers qui en permettraient une application plus pratique. Ces ateliers pourraient permettre non seulement d'envisager d'autres futurs que ceux repris par des œuvres de fiction populaires et ainsi d'élargir le débat, en ajoutant potentiellement une composante créative, mais aussi de faire interagir les générations sur la question. Bien que cela n'ait pu être étudié en détail ici, les réponses apportées au questionnaire permettaient de constater que chaque tranche d'âge était généralement plus propice à citer certaines œuvres bien spécifiques : les 20-30 ans ont mentionné beaucoup d'œuvres issues des années 2010 tandis que les 30-50 ans s'inscrivaient dans une fourchette plus large, avec des œuvres remontant aux années 1980. Cette variété d'œuvres et d'expériences induit probablement des variations dans les perceptions des agents conversationnels et futurs envisagés pour ces technologies : faire dialoguer un plus vaste échantillon de personnes, sans se cantonner à un critère d'âge, serait intéressant à cette fin.

Les travaux de Kirby théorisant la favorabilité envers une nouvelle technologie par le biais de la fiction ont constitué une véritable base à l'ensemble des hypothèses qui ont été posées dans cette recherche. Les entretiens ont toutefois présenté une contradiction avec une partie de sa théorie, qui entend que continuer à présenter la

⁵³ <https://nerdynav.com/chatgpt-statistics/>, consulté le 4 juin 2023.

⁵⁴ <https://openai.com/research/whisper>, consulté le 31 mai 2023.

technologie dans la fiction après son incursion sociétale permet d'améliorer les attentes que le public a envers ladite technologie. Or, une affirmation récurrente de la part des répondants suggère qu'à partir du moment où l'agent échoue à combler un certain type de requête, cette option est rendue impossible dans l'esprit des utilisateurs qui ne réitérent pas l'expérience, même plusieurs mois ou années plus tard. La question de l'amélioration des attentes peut alors se poser, la fiction n'ayant eu de cesse de présenter des agents conversationnels ces dernières années. Différentes hypothèses peuvent être envisagées, telles que la perception globale des œuvres présentant ces technologies qui semblent tendre à une dystopie unanime ou bien l'influence populaire et peu favorable ou optimiste envers le futur de ces technologies malgré leur incursion réussie. Cette question pourrait prolonger la réflexion quant à l'adoption de ces technologies grâce à la fiction, voire la redéfinir sous la perception d'un « non-rejet » qui serait suivie ou non d'une adhérence populaire, les exemples de nouvelles technologies ayant pu intégrer l'espace social sans toutefois bénéficier d'une favorabilité sociétale ne manquant pas (les agents conversationnels, la réalité virtuelle ou les Google Glass peuvent être cités exhaustivement).

D'un point de vue fictionnel, au regard de l'adoption mitigée constaté via les entretiens et au vu des caractéristiques du prototype diégétique qui participent à faire accepter une nouvelle technologie (Kirby, 2010), la question peut se poser de savoir si les agents conversationnels remplissent l'ensemble de ces critères. Ainsi, si la nécessité, la viabilité et la bienveillance des agents conversationnels sont questionnées dans les six œuvres les plus citées du questionnaire, le critère de bienveillance n'est pas toujours avéré au regard des entretiens et de l'opinion des répondants sur la question. Cette bienveillance est d'ailleurs ce qui est principalement controversé dans les œuvres dont la fin est perçue négativement : *HAL9000 (2001: A Space Odyssey)*, les agents conversationnels de *Black Mirror* et *Sonny (I, Robot)* menacent la vie des humains. *Samantha (Her)* ne dispose toutefois pas d'une représentation très positive, mais il s'agit plutôt de sa nécessité qui pourrait être questionnée, le récit s'entamant avec Théodore qui n'a pas d'agent conversationnel et s'achevant sur « le départ » de Samantha. Les licences *Avengers* et *Star Wars* remplissent l'ensemble des critères pour présenter leurs agents conversationnels, ce qui semble participer à la vision très positive de ceux-ci.

Malgré certains manques dans la variété des sujets traités par la (science-)fiction – du moins au regard de préoccupations récurrentes – celle-ci présente de nombreux avantages dans la prospection et l'ouverture à des débats, facilités par l'approche prototypale conférée par le design fiction. Le développement d'ateliers requérant de mettre en pratique ce pan du cadre conceptuel du design fiction, en invitant par exemple soit des individus à rédiger de brèves nouvelles anticipatives ou de science-fiction, soit des auteurs dont les récits seraient discutés, pourrait encourager à mettre à jour de nouvelles narrations pertinentes. Cet usage permettrait par ailleurs d'inclure plus aisément des individus variés dans le débat que cela engendrerait, la narration étant un moyen propice pour communiquer sur des sujets avec des non-initiés (Tanenbaum, 2014). Cette technique, déjà utilisée par Hudson, Finn et Wylie (2021) pour réfléchir à des politiques réglementaires de l'intelligence artificielle en Amérique du Nord, avait donné des résultats encourageants et favorables à une ouverture du débat ; une reproduction du sujet appliqué à l'Europe ou le choix d'un nouvel angle pourrait ouvrir à de nouvelles réflexions. Cette pratique apporterait par ailleurs une plus-value dans une présentation plus objective de possibles futurs par

le récit ; la science-fiction a beau être un genre de prédilection pour penser prospectivement les nouvelles technologies, l'angle dystopique qui y est souvent adopté met à mal le critère de bienveillance ou de nécessité qui peut permettre au public de ne pas rejeter naturellement la technologie présentée (ou, du moins, d'y être défavorable, voire de créer une crainte autour du sujet). La construction de récits – qui ne seraient utilisés que dans le cadre dudit atelier, afin de ne pas risquer de dérives marketing qui biaiserait également le débat à son tour – appuyant le respect des trois critères définis par Kirby permettrait un débat plus constructif, n'encourageant pas d'office une favorabilité et une adhésion populaire à la technologie, mais autorisant *a minima* une conception réaliste de celle-ci et de ses futurs.

Bibliographie

- Barchechath E., Pouts-Lajus S. (1985). Postface sur l’interactivité. Dans Kel Crossley, Les Green, *Le design des didacticiels. Guide pratique pour la conception de scénarios interactifs*, Paris, ACL Éditions.
- Barricelli, B., Valtolina, S., Gaetano, S., et Diliberto, P. (2019). Chatbots and conversational interfaces : Three domains of use. *Fifth International Workshop on Cultures of Participation in the Digital Age - CoPDA 2018*
- Bates, M. (1993, Février, 8-9). *Models of natural language understanding* [Présentation de colloque]. Human-Machine Communication by Voice, Irvine, CA, USA.
- Beaudouin, V. (2002). De la publication à la conversation. Lecture et écriture électroniques. *Réseaux*, 116(6), 199-225. Cairn.info.
- Bellanger, A. (2012). *La théorie de l’information : Roman*. Gallimard.
- Benbouzid, B., Meneceur, Y., et Smuha, N. A. (2022). Quatre nuances de régulation de l’intelligence artificielle. Une cartographie des conflits de définition. *Réseaux*, 232-233(2-3), 29-64. Cairn.info. <https://doi.org/10.3917/res.232.0029>
- Bender, E. M., Gebru, T., McMillan-Major, A., et Shmitchell, S. (2021). On the Dangers of Stochastic Parrots : Can Language Models Be Too Big?. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623. <https://doi.org/10.1145/3442188.3445922>
- Bengio, Y. (2008). Neural net language models. *Scholarpedia*, 3(1), 3881. <https://doi.org/10.4249/scholarpedia.3881>
- Bleecker, J. (2009). *Design Fiction. A short essay on design, science fact and fiction*. Near Future Laboratory. https://systemsorienteddesign.net/wp-content/uploads/2011/01/DesignFiction_WebEdition.pdf
- Blythe, M. (2014). Research through design fiction : Narrative in real and imaginary abstracts. *Conference on Human Factors in Computing Systems - Proceedings*. <https://doi.org/10.1145/2556288.2557098>
- Boë, L.-J., Ménard, L., Serkhane, J., Birkholz, P., Kröger, B., Badin, P., Captier, G., Canault, M., et Kielwasser, N. (2008). La croissance de l’instrument vocal : Contrôle, modélisation, potentialités acoustiques et conséquences perceptives. *Revue française de linguistique appliquée*, XIII(2), 59-80. Cairn.info. <https://doi.org/10.3917/rfla.132.0059>
- Boyer, A., et Farzaneh, F. (2019). Vers une éthique de la robotique. Towards an ethic of robotics. *Question(s) de management*, 24(2), 67-84. Cairn.info. <https://doi.org/10.3917/qdm.192.0067>
- Braunschweig, B. (2023). Les cinq murs de l’intelligence artificielle. *Futuribles*, N° 453(2), 5-24. <https://doi.org/10.3917/futur.453.0005>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ...

- Amodei, D. (2020). *Language Models are Few-Shot Learners*.
<https://doi.org/10.48550/ARXIV.2005.14165>
- Cassell, J., Velkovska, J. et Relieu, M. (2020). « Tisser des liens »: L'interaction sociale chez les agents conversationnels. *Réseaux*, 220-221, 21-45.
<https://doi.org/10.3917/res.220.0021>
- Cheng, Y.-T., et Huang, Y.-C. (2021). *When People Vanish : A Study to Investigate How Human Presence Changes the Scenario Speculation*.
<https://doi.org/10.1145/3462204.3481781>
- Cheng, Y.-T., Tasi, G.-L., Lin, H.-A., Huang, Y.-C., Liang, R.-H., et Chen, L.-L. (2019). *Call Me by My Name : Exploring Roles of Sci-fi Voice Agents*.
<https://doi.org/10.1145/3290607.3313074>
- Chevet, C. (2017). La voix de synthèse : de la communication de masse à l'interaction homme-machine. Dialogue avec le monde. *Communication & langages*, 193, 63-78. <https://doi.org/10.3917/comla.193.0063>
- Cohard, P. (2015). L'apprentissage dans les serious games : Proposition d'une typologie. *@GRH*, 16(3), 11-40. <https://doi.org/10.3917/grh.153.0011>
- Combessie, J-C. (2007). *La méthode en sociologie*, La Découverte.
- Crépel, M., et Cardon, D. (2022). Robots vs algorithmes. Prophétie et critique dans la représentation médiatique des controverses de l'IA. *Réseaux*, 232-233(2-3), 129-167. Cairn.info. <https://doi.org/10.3917/res.232.0129>
- Curiale, T., Acquatella, F., Gros, L., Cosquer, M., et Tisseron, S. (2022). L'anthropomorphisme, enjeu de performance pour les chatbots. *Revue internationale de psychosociologie et de gestion des comportements organisationnels*, XXVIII(72), 101-123. Cairn.info. <https://doi.org/10.3917/rips1.072.0101>
- Damhuis, L., et Ruquoy, D. (2012). *Collecte de données : L'entretien semi-directif*. https://uclouvain.be/pluginfile.php/20823/mod_resource/content/4/%20Entretien%20semi-directif%20Damhuis%20et%20Ruquoy%202012.pdf
- Defays, D., French, R., et Sougné, J. (1999). Apport de l'intelligence artificielle à la psychologie. In J.-A. Rondal (Ed.), *Introduction à la psychologie scientifique* (pp. 379-415). Bruxelles, Belgium: Labor.
- Dubois, C., Salotti, J. M., Seminel, D., et Simonazzi, N. (2019). Le chatbot : Un outil de la relation aux clients. *Hermès, La Revue*, 84(2), 95-97. Cairn.info. <https://doi.org/10.3917/herm.084.0095>
- Durand, P. (2018). *COMU2145-1 : Introduction aux médias et à l'information* [Unpublished document, Université de Liège, Liège]
- Dunne, A., et Raby, F. (2013). *Speculative everything : Design, fiction, and social dreaming*. The MIT Press.
- El Hefny, W., El Bolock, A., Herbert, C., et Abdennadher, S. (2020). Towards a Generic Framework for Character-Based Chatbots. In F. De La Prieta, P. Mathieu, J. A. Rincón Arango, A. El Bolock, E. Del Val, J. Jordán Prunera, J. Carneiro, R. Fuentes, F. Lopes, et V. Julian (Éds.), *Highlights in Practical Applications of Agents, Multi-Agent*

Systems, and Trust-worthiness. The PAAMS Collection (Vol. 1233, p. 95-107). Springer International Publishing. https://doi.org/10.1007/978-3-030-51999-5_8

— Esteban, C. (2020). Construire la « compréhension » d'une machine : Une ethnographie de la conception de deux chatbots commerciaux. *Réseaux*, N° 220-221(2), 195-222. <https://doi.org/10.3917/res.220.0195>

— Faruqui, M., et Hakkani-Tür, D. (2021). *Revisiting the Boundary between ASR and NLU in the Age of Conversational Dialog Systems*. <https://doi.org/10.48550/ARXIV.2112.05842>

— Georges, F. (2012). Avatars et identité. *Hermès, La Revue*, 62(1), 33-40. Cairn.info. <https://doi.org/10.4267/2042/48274>

— Giard, A. (2019). Affects électroniques au Japon. Les systèmes de rencontre par machines interposées. *Hermès, La Revue*, 84(2), 104-110. Cairn.info. <https://doi.org/10.3917/herm.084.0104>

— Girandola, F., et Fointiat, V. (2016). *Attitudes et comportements : Comprendre et changer*. Presses universitaires de Grenoble; Cairn.info. <https://www.cairn.info/attitudes-et-comportements-comprendre-et-changer--9782706125812.htm>

— Gkinko, L., et Elbanna, A. (2022). Hope, tolerance and empathy : Employees' emotions when using an AI-enabled chatbot in a digitalised workplace. *Information Technology & People*, 35(6), 1714-1743. <https://doi.org/10.1108/ITP-04-2021-0328>

— Gobin Mignot, É., Wolff, B., et Kempf, N. (2019). *Former avec la réalité virtuelle : Comment les techniques immersives bouleversent l'apprentissage*. Dunod.

— Gouliaéva, O., Dosquet, É., et Moysan, Y. (2020). *La révolution des assistants vocaux. Comprendre les enjeux et réussir ses stratégies marketing*. Dunod; Cairn.info. <https://www.cairn.info/la-revolution-des-assistants-vocaux--9782100805716.htm>

— Grandjean, D. et Baenziger, T. (2014). *Chapitre 4. Expression vocale des émotions*. Dans : David Sander éd., *Traité de psychologie des émotions* (pp. 120-166). Paris: Dunod. <https://doi.org/10.3917/dunod.sande.2014.01.0120>

— Gras Gentiletti, M., Bourmaud, G., Fréjus, M., et Decortis, F. (2022). Concevoir pour des activités instrumentées par des chatbots : Apports d'une approche de l'activité située et médiatisée pour la conception. *Activites*, 19-1. <https://doi.org/10.4000/activites.7428>

— Guéneau, C. (2005). L'interactivité : Une définition introuvable. *Communication et langages*, 145(1), 117-129. <https://doi.org/10.3406/colan.2005.3365>

— Guest, G., Bunce, A., et Johnson, L. (2006). How Many Interviews Are Enough? : An Experiment with Data Saturation and Variability. *Field Methods*, 18(1), 59-82. <https://doi.org/10.1177/1525822X05279903>

— Hudson, A. D., Finn, E., et Wylie, R. (2023). What can science fiction tell us about the future of artificial intelligence policy? *AI & SOCIETY*, 38(1), 197-211. <https://doi.org/10.1007/s00146-021-01273-2>

- Hutchins, E. (2001). Cognition, Distributed. In *International Encyclopedia of the Social & Behavioral Sciences* (p. 2068-2072). Elsevier. <https://doi.org/10.1016/B0-08-043076-7/01636-3>
- Imbert, G. (2010). L'entretien semi-directif : À la frontière de la santé publique et de l'anthropologie. *Recherche en soins infirmiers*, 102(3), 23-34. Cairn.info. <https://doi.org/10.3917/rsi.102.0023>
- Jacques, J. (2021). *LSTIC2000 : Médias, informations et connaissances* [Unpublished document, Université Catholique de Louvain, Louvain-la-Neuve]
- Julia, J.-T. (2003). Interactivité, modes d'emploi. Réflexions préliminaires à la notion de document interactif. *Documentaliste-Sciences de l'Information*, 40(3), 204-212. Cairn.info. <https://doi.org/10.3917/docs.403.0204>
- Karpf, A. (2008). La voix : Un univers invisible. *Autrement*. Pp. 392-426. <https://doi.org/10.3917/autre.karpf.2008.01>
- Kirby, D. (2010). The Future Is Now : Diegetic Prototypes and the Role of Popular Films in Generating Real-World Technological Development. *Social Studies of Science - SOC STUD SCI*, 40, 41-70. <https://doi.org/10.1177/0306312709338325>
- Laforest, F., Frénot, S. et Al Masri, N. (2002). Dossier médical semi-structuré pour des interfaces de saisie multimodales. *Document numérique*, 6, 29-46. <https://doi.org/10.3166/dn.6.1-2.29-46>
- Larousse. (s. d.) Science-fiction. Dans *Larousse en ligne*. Consulté le 30 mai 2023. <https://www.larousse.fr/dictionnaires/francais/science-fiction/71469>
- Le Robert. (s. d.) Algorithme. Dans *Le Robert – Dico en ligne*. Consulté le 10 février 2023. <https://dictionnaire.lerobert.com/definition/algorithme>
- Le Robert. (s. d.) Diégèse. Dans *Le Robert – Dico en ligne*. Consulté le 29 décembre 2022. <https://dictionnaire.lerobert.com/definition/diegesse>
- Lindley, J., Sharma, D., et Potts, R. (2014). Anticipatory Ethnography : Design Fiction as an Input to Design Ethnography. *Ethnographic Praxis in Industry Conference Proceedings*, 2014(1), 237-253. <https://doi.org/10.1111/1559-8918.01030>
- Lindley, J. (2014). A Machine. Learning : An example of HCI Prototyping With Design Fiction. In *Proceedings of the NordiCHI 2014 : The 8th Nordic Conference on Human-Computer Interaction : Fun, Fast, Foundational*. <https://doi.org/10.1145/2639189.2670281>
- Lindley, J. (2016). A Pragmatics Framework for Design Fiction. In 11th EAD Conference Proceedings: The Value of Design Research. *European Academy of Design Conference Proceedings 2015*. Sheffield Hallam University. <https://doi.org/10.7190/ead/2015/69>
- Makransky, G., et Mayer, R. E. (2022). Benefits of Taking a Virtual Field Trip in Immersive Virtual Reality : Evidence for the Immersion Principle in Multimedia Learning. *Educational Psychology Review*, 34(3), 1771-1798. <https://doi.org/10.1007/s10648-022-09675-4>
- Mauldin, M.L. (1994). CHATTERBOTS, TINYMUDS, and the Turing Test: Entering the Loebner Prize Competition. *AAAI Conference on Artificial Intelligence*.

- Mayer, R. E. (2002). Cognitive Theory and the Design of Multimedia Instruction : An Example of the Two-Way Street Between Cognition and Instruction. *New Directions for Teaching and Learning*, 2002(89), 55-71. <https://doi.org/10.1002/tl.47>
- McLuhan, M. (1964). *Understanding media : The extensions of man* (1st MIT Press ed). MIT Press.
- Miaillhe, N. (2018). Géopolitique de l'Intelligence artificielle : Le retour des empires ? *Politique étrangère*, Automne(3), 105-117. Cairn.info. <https://doi.org/10.3917/pe.183.0105>
- Minvielle, N. et Wathelet, O. (2017). Le design fiction: Une méthode pour explorer les futurs et construire l'avenir ?. *Futuribles*, 421, 69-83. <https://doi.org/10.3917/futur.421.0069>
- Mollon, M., et Gentes, A. (2014). (2014) *The Rhetoric of Design for Debate : Triggering conversation with an "uncanny enough" artefact* – At DRS.
- Morin, É., Therriault, G., et Bader, B. (2019). Le développement du pouvoir agir, l'agentivité et le sentiment d'efficacité personnelle des jeunes face aux problématiques sociales et environnementales : Apports conceptuels pour un agir ensemble. *Éducation et socialisation*, 51. <https://doi.org/10.4000/edso.5821>
- OpenAI. (2023). *GPT-4 Technical Report*. <https://doi.org/10.48550/ARXIV.2303.08774>
- Papadoudi-Ros, H. (2018). Serge Tisseron, Le Jour où mon robot m'aimera. Vers l'empathie artificielle. Paris, A. Michel, 2015, 208 pages. *Questions de communication*, 34(2), 420-422. Cairn.info. <https://doi.org/10.4000/questionsdecommunication.17011>
- Radford, A., et Narasimhan, K. (2018). Improving Language Understanding by Generative Pre-Training.
- Relieu, M., Sahin, M., et Francillon, A. (2020). Une approche configurationnelle des leurres conversationnels. *Réseaux*, N°220-221(2), 81. <https://doi.org/10.3917/res.220.0081>
- Santolaria, N. (2016). « Dis Siri ». *Enquête sur le génie à l'intérieur du smartphone*. Anamosa; Cairn.info. <https://www.cairn.info/dis-siri--9791095772095.htm>
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-424. <https://doi.org/10.1017/S0140525X00005756>
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Souchier, E. (2012). La « lettrure » à l'écran. Lire & écrire au regard des médias informatisés. *Communication & langages*, 174(4), 85-108. Cairn.info. <https://doi.org/10.4074/S0336150012014068>
- Souchier, E., Candel, É., et Gomez-Mejia, G. (2019). *Le numérique comme écriture. Théories et méthodes d'analyse*. Armand Colin; Cairn.info. <https://www.cairn.info/le-numerique-comme-ecriture--9782200618582.htm>
- Sterling, B. (2005). *Shaping things*. MIT Press.

- Strum, S., & Latour, B. (2006). Redéfinir le lien social : Des babouins aux humains. In M. Akrich, M. Callon, & B. Latour (Éds.), *Sociologie de la traduction* (p. 71-86). Presses des Mines. <https://doi.org/10.4000/books.pressesmines.1192>
- Tanenbaum, T.J. (2014). Design fictional interactions: why HCI should care about stories. *Interactions* 21, 5 (September + October 2014), 22–23. <https://doi.org/10.1145/2648414>
- Tomasello, M. (2010). *Origins of Human Communication*. MIT Press.
- Turing, A. M. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*, LIX(236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>
- Van Campenhoudt, L., Marquet, J., et Quivy, R. (2017). *Manuel de recherche en sciences sociales*. Dunod.
- Velkovska, J., et Zouinar, M. (2020). Les relations aux machines « conversationnelles ». Vivre avec les assistants vocaux à la maison. *Réseaux*, 220-221(2-3), 47-79. Cairn.info. <https://doi.org/10.3917/res.220.0047>
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45. <https://doi.org/10.1145/365153.365168>

Simultanément outils fantasmés et artefacts fictionnels, les agents conversationnels sont au cœur de nombreuses interrogations et récits. Supports de multiples tentatives de reproduction de l'intelligence humaine, la fiction les a parallèlement érigés en protagonistes et a construit un imaginaire relationnel et émotionnel prospère autour de l'interaction homme-machine. Alors que nous nous trouvons à un tournant significatif dans le développement technique de ces agents conversationnels, la présente recherche vise à évaluer les représentations construites à leur égard par le biais de la fiction, à travers une démarche hypothético-déductive s'appuyant sur le cadre conceptuel du design fiction. Elle comprend notamment l'analyse d'un questionnaire auto-administré visant à récolter des données exploratoires et la conduite d'entretiens semi-directifs sur base de celles-ci. L'implication de la voix jouant un rôle tout particulier dans la génération d'émotions, qui influence elle-même l'adoption sociale ou non d'une technologie, une attention toute particulière lui est accordée.

Mots-clefs : agent conversationnel, design fiction, interaction homme-machine, émotions, voix, représentations