Document generated on 06/20/2025 2:17 p.m.

Sociologie et sociétés

Repenser la sociologie du numérique à l'aune de notre vie commune avec les robots sociaux
Rethinking digital sociology through our shared life with social bots

Olivier Glassey

Volume 49, Number 2, Fall 2017

La sociologie numérique
Digital Sociology

URI: https://id.erudit.org/iderudit/1054274ar
DOI: https://doi.org/10.7202/1054274ar

See table of contents

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0038-030X (print)
1492-1375 (digital)

Explore this journal

Cite this article

Glassey, O. (2017). Repenser la sociologie du numérique à l'aune de notre vie commune avec les robots sociaux. *Sociologie et sociétés*, *49*(2), 59–82. https://doi.org/10.7202/1054274ar

Article abstract

The article focuses on the theoretical and empirical implications of humans and social bots copresence in digital social spaces. It examines how these entities are constructed trough media narratives and computer studies researches about the trouble they cause. The article shows how social robots' identification and categorization processes are also redefinitions of what constitute humanness. On this basis, the article defends the thesis that it becomes difficult to imagine a digital sociology that is exempt from the examination of these anthropomorphic nonhuman entities.

Tous droits réservés © Les Presses de l'Université de Montréal, 2018

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

https://apropos.erudit.org/en/users/policy-on-use/

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

https://www.erudit.org/en/

Repenser la sociologie du numérique à l'aune de notre vie commune avec les robots sociaux

OLIVIER GLASSEY
Laboratoire d'étude des sciences et des techniques (STS-Lab)
Université de Lausanne
Courriel : Olivier.Glassey@unil.ch

Les études sociales du numérique s'intéressent depuis quelques décennies déjà aux traces des interactions matérialisées et rendues accessibles en ligne. Les interprétations qualitatives et quantitatives de ces traces ont contribué à un large spectre de travaux sociologiques visant à explorer les pratiques sociales en ligne. L'avènement, au début du millénaire, d'un web dit « social » a popularisé cette perception des traces numériques comme expression, voire extension, des interactions sociales. Cette connexion presque naturalisée de ces traces et de la part du social qu'elles sont censées traduire demeure cependant problématique. Ce lien n'est en effet ni évident ni direct. Il se trouve, par exemple, sans cesse construit par des logiques algorithmiques qui en recomposent les modalités. En outre, l'analyse de ces données considérées comme matérialisation du social ne commence qu'à peine à rendre compte d'une dimension importante : le fait qu'une partie non négligeable de celles-ci sont produites ou suscitées par des entités non humaines. Le présent article prend comme point de départ la prolifération des « *social bots* », ou robots sociaux[1], avec qui nous partageons nos

1. Les robots « sociaux », en anglais *socialbots*, *social bots* ou encore *socbots*, ne sont pas une appellation contrôlée et figée. Nous garderons ici volontairement ce terme générique qui possède pour nous l'intérêt de la juxtaposition de « robot » et de « social ».

espaces de sociabilité numériques pour en interroger les formes et la portée. L'arrivée massive de ces robots numériques confère une actualité très concrète aux enjeux de la coexistence entre humains et non-humains, qui semblait jusqu'ici relever essentiellement de la science-fiction ou de la prospective.

La majorité des opérateurs des principaux réseaux sociaux en ligne reconnaissent la présence de tels robots sur leur plateforme. Une entreprise comme Twitter admet par exemple dans une déclaration officielle de 2014 qu'entre 5 % et 8,5 % de ses comptes utilisateurs se trouvent à l'actif de robots[2]. Des recherches plus récentes (Varol *et al.*, 2017) situent la population de *bots* entre 9 % et 15 % dans leurs estimations les plus basses. Pour le seul réseau Twitter, ces chiffres représentent potentiellement quelque 48 millions de robots sur une population de 319 millions d'utilisateurs actifs mensuels. À l'échelle de l'ensemble des plateformes numériques les plus populaires, c'est donc par centaines de millions que l'on peut estimer la présence des dispositifs de type robots sociaux. Alors que les robots mécaniques « traditionnels » œuvrant dans le monde physique progressent lentement sur les chemins de l'interaction avec l'humain, les robots strictement numériques sont déjà des interlocuteurs massivement à l'œuvre. Agissant dans l'univers des traces du social matérialisé en ligne, leur existence et les effets qui leur sont attribués interpellent à la fois la société et les sciences humaines. Les enjeux relatifs au pouvoir de manipulation potentiel de ces systèmes sont au cœur de préoccupations politiques (Morozov, 2012), de nombreux incidents étant rapportés un peu partout dans le monde (Woolley et Howard, 2017)[3].

Sur la base de ce constat préliminaire, notre première interrogation est en apparence simple : quelle est la nature de ce « social » numérique quand la part de la population non humaine est prise en considération ? Comment caractériser cette mixité pour en appréhender les effets ? Quelles réactions et quels ajustements cette situation provoque-t-elle ? La présence d'une multitude d'entités interagissant avec des humains doit-elle nous conduire à réviser en profondeur les formes de l'analyse sociologique du numérique ?

Pour répondre à ces questions, nous proposons de ne pas nous focaliser sur la seule dimension technique des robots sociaux, mais plutôt d'ouvrir la réflexion aux enjeux épistémologiques que dessine leur coprésence avec des humains en ligne. Dans un premier temps, nous convoquerons des concepts issus de la sociologie des sciences et des techniques pour identifier, à partir d'éléments théoriques, les contours de cette cohabitation. La deuxième partie explorera empiriquement les modalités de cette coprésence en examinant deux types de situations spécifiques. La première situation prendra pour objet les réactions des acteurs du web relativement au phénomène des robots sociaux. Leurs discours médiatiques relativement au trouble occasionné par les robots seront utilisés comme révélateurs des registres interprétatifs mobilisés pour caractériser la nature problématique de la présence de ces agents. La seconde situation

2. Déclaration à la SEC (Securities and Exchange Commission), USA.

3. Parmi les cas signalés par les auteurs : Allemagne, Brésil, Canada, Chine, États-Unis, Pologne, Russie, Taiwan, Ukraine.

abordera la manière dont les chercheurs des *computer studies* analysent cette présence en ligne et tentent d'y répondre. Nous avons, pour cela, constitué un corpus d'articles publiés à partir de recherches conduites spécifiquement sur la thématique des *robots sociaux*. Nous nous attacherons à illustrer la manière dont ces travaux incluent implicitement dans leur cadre conceptuel des propositions de formalisation croisée du social et du technique. Nous formulerons alors l'hypothèse que les prises de position des acteurs du web, ainsi que les protocoles d'investigation de ces recherches, dessinent une double lecture performative du technique et du social : les processus d'explicitation de la nature des robots sociaux engagent une description de ce qui caractérise l'humain en ligne. *In fine*, nous discuterons la nécessité d'une analyse de ces définitions croisées comme adjuvant à toute sociologie des pratiques du numérique.

DÉFINIR LES ROBOTS DITS SOCIAUX

Au sein de la riche tradition de l'étude du rapport de l'homme à son environnement technique, les objets techniques demeurent le plus souvent synonymes d'outils. L'histoire et l'anthropologie des techniques, les recherches sur l'innovation ou encore la sociologie de l'usage se sont intéressées à la manière dont ces dispositifs accroissent, dupliquent, accompagnent, encadrent, prescrivent ou libèrent l'action humaine. Les robots sociaux impliquent d'opérer un léger glissement par rapport à ces perspectives pour se focaliser sur un problème particulier : qu'advient-il quand la technologie prétend devenir un acteur et que l'outil se fait passer pour un humain ?

La première étape, qui consiste à définir les « robots sociaux », s'avère d'emblée une entreprise délicate. L'identification de ces dispositifs, c'est-à-dire le fait de les singulariser dans la masse des autres fonctionnalités existantes du web, est en effet problématique. Une telle démarche s'expose au danger d'hypostasier un mécanisme technique qui participe et dépend d'une chaîne de processus plus complexes (conception, programmation, diffusion, activation et désactivation). Évoquer ces robots en tant qu'agents autonomes risque de les extraire de leurs conditions de production et de les instituer en des entités qu'ils ne sont pas. En d'autres termes, les invoquer dans une analyse, c'est déjà participer activement à la construction de l'objet ainsi qu'à sa mythologie. Cette difficulté n'invalide pas la spécificité du phénomène, comme en attestent de nombreuses définitions (Leonard, 1997 ; Middlebrook et Muller, 2000 ; Ferrara *et al.* 2015). Leur lecture permet de constater schématiquement une convergence autour de trois dimensions clés que sont la personnalisation, le dialogue avec un humain et l'automatisation des scripts de réponses.

Dans leur acception la plus générique et au stade des développements techniques actuels, les robots « sociaux » sont des programmes informatiques qui participent de manière anthropomorphisée à des échanges en ligne avec des humains (Gehl et Bakardjieva, 2017). Cette esquisse de définition, quoique limitée, permet néanmoins d'opérer une première différenciation entre les robots sociaux et les multiples formes de dispositifs numériques automatisés. Les premiers acquièrent le statut de « social » par leur capacité à fabriquer pour autrui l'apparence d'une conversation. Les seconds,

fort nombreux, interviennent en toute autonomie sur les contenus en ligne. Ils peuvent ainsi parcourir la Toile et y apporter des modifications sans forcément s'incarner et exposer leur existence pour les humains. Cette distinction est essentielle, car elle institue le lien potentiel entre humains et non-humains comme une dimension fondatrice du *social bot*. En d'autres termes, le robot social s'apparente à un programme informatique équipé d'une capacité conversationnelle et d'une identité explicite qui agit sur le monde des traces numériques, environnement qu'il partage avec les humains. Une approche technique basée sur la seule dissection du code ne peut pas s'avérer suffisante pour saisir l'ontologie de ce lien. Elle séparerait artificiellement ce qui se trouve, dans la réalité des échanges en ligne, fortement entrelacé. En effet, c'est bien le mélange complexe et mouvant des contributions de nature humaine ou robotique qui nous intéresse. Afin d'appréhender un tel terrain hybride, il convient d'abord de prendre en considération les origines des robots sociaux actuels.

AUX ORIGINES DES *BOTS*

Les robots sociaux, considérés comme dispositifs conversationnels, s'inscrivent dans l'histoire du développement des interactions personnes-machines. Ils procèdent d'une tradition de mécanismes imitant la prise de parole et le dialogue à l'image des *chatbots* qui ont animé les messageries instantanées depuis la fin des années 1990 (Latzko-Toth, 2017). Ils sont également les héritiers directs des expérimentations de discussions avec un ordinateur, à l'exemple du célèbre programme ELIZA (Weizenbaum, 1966). Ce programme simple fondé sur la reconnaissance de formes propose un simulacre de l'approche en psychothérapie de l'écoute rogérienne. Il reformule, de manière interrogative, les messages des usagers[4]. Nonobstant le côté assez primaire et répétitif du système, ELIZA intrigue déjà, à l'époque, une partie de ses utilisateurs. Ce logiciel préfigure le potentiel de séduction des robots sociaux et donne son nom à un « effet ELIZA », caractérisant la tendance à assimiler inconsciemment le comportement d'un ordinateur à celui d'un être humain (Hofstadter, 1996).

Cette lecture rétrospective implique de remonter au cœur des réflexions relatives à l'intelligence artificielle conduites par Alan Turing au début des années 1950. Pour Turing, la possibilité qu'une entité masquée participant à un dialogue avec une personne puisse être perçue comme humaine valide le statut de son intelligence. Le principe du test qu'il propose est simple : une machine parvenant à se faire passer pour un interlocuteur humain lors d'un dialogue prouve son « intelligence » (Turing, 1950). Nous ne pouvons refaire ici l'histoire de ce test ni recenser les innombrables débats qu'il a impulsés[5] et les controverses qu'il continue à engendrer (Christian, 2011)[6].

4. Pour expérimenter une émulation de programme ELIZA : http://www.manifestation.com/neurotoys/eliza.php3 ; pour une explicitation du contenu du programme : http://www.filfre.net/2011/06/eliza-part-2/

5. Sur ce point, voir le contre-argumentaire de la chambre chinoise de Searle (1980).

6. Sur ce point, voir l'ouvrage témoignage de Brian Christian (2011) nommé juge pour Le prix Loebner, compétition annuelle qui couronne les agents conversationnels satisfaisant le mieux les critères du test de Turing.

Soulignons simplement que les robots sociaux reproduisent et démultiplient ce schéma interactionnel expérimental canonique : ils se proposent, au quotidien, comme partenaires de dialogue à des humains tout en demeurant tapis derrière le paravent des médiations propres aux échanges numériques. Turing emprunte l'idée de son test à un jeu de salon bourgeois du 19e siècle, l'*imitation game,* qui consistait à deviner, par l'entremise de missives anonymes, le genre d'un convive présent dans une autre pièce. Ce jeu de cache-cache mondain de l'identité préfigure une dimension clé des robots sociaux : l'explicitation ou, au contraire, l'occultation de leur nature mécanique. Dans cette logique, les programmeurs de *bots* vont exploiter une caractéristique bien identifiée du numérique, à savoir le doute toujours possible quant à l'identité réelle de l'interlocuteur au-delà de l'écran. Cette ambivalence n'est toutefois pas systématique, certains robots « sociaux » affichant en effet leur statut d'entité digitale. Il s'agit par exemple des assistants et majordomes automatisés, ou encore des robots « artistiques » qui proposent des devinettes, génèrent des cadavres exquis informationnels ou détournent à leur manière les flux d'information du web (Finley, 2013).

L'avènement du web social constitue à cet égard un moment clé dans la prolifération des robots. Grâce au déploiement à grande échelle des plateformes numériques sollicitant fortement les contributions des usagers en ligne (conversations, photographies, commentaires, évaluations), c'est un immense terrain de jeu qui s'offre à ces dispositifs automatisés. L'écosystème des fonctionnalités associées à l'usage des réseaux sociaux est un environnement très favorable à des agents programmés pour incarner un interlocuteur humain. Comme le souligne Bakardjieva (2015), la rationalisation des activités au sein de ces réseaux se traduit par la standardisation, l'uniformisation, la prédictibilité et la quantification des interactions sociales, autant de dimensions propices à un traitement robotisé de l'information. Les formats réduits des échanges et l'anonymat facilité offrent des leviers d'action. La capacité de se rendre visible par la publication de contenus numériques, les modalités minimales et mécaniques de la mise en scène du lien social (*like*, *follow*, *retweet*, hashtag, emoji, etc.) constituent autant d'« affordances » avec lesquelles les concepteurs de robots peuvent jouer. Le flux important des échanges, leur volatilité et la compétition entre chaque dispositif pour accaparer l'attention des utilisateurs sont autant de pourvoyeurs de distraction, alliés objectifs des robots (Qiu *et al.*, 2017). En résumé, dans ces espaces de communication reposant sur la matérialisation numérique des échanges sociaux, les *bots* trouvent les ressources permettant la construction de la vitrine d'une identité factice.

Les actions des usagers eux-mêmes contribuent à cette diffusion sociale des robots. Saisis dans les logiques de mise en visibilité de soi que ces plateformes promeuvent (Cardon, 2008 ; Granjon, 2011), ils plébiscitent toutes formes de renforcements (abonnés, suiveurs, amis, etc.). L'établissement de la nature des entités ainsi cooptées s'avère souvent moins important que la valeur symbolique et parfois économique de la popularité associée à leur nombre.

Considérés globalement, les robots sociaux se caractérisent par leurs manières de jouer implicitement ou explicitement avec les traces attribuées à des humains. La mesure de leur succès est moins liée à une quelconque intelligence artificiellement

simulée qu'à leur aptitude à rendre difficilement lisible la distinction entre humains et non-humains. C'est cette transition « *from Turing test to blurring test* » (Weil, 2017 : 21) qui constitue l'une des originalités de la situation contemporaine.

Le numérique n'offre pas seulement un univers interactionnel simplifié, il est aussi un extraordinaire pourvoyeur d'informations sur les pratiques sociales de ses utilisateurs. Depuis le programme ELIZA qui reformatait les propos de ses interlocuteurs, les robots sociaux exploitent souvent les productions spontanées des internautes pour nourrir leurs propres capacités conversationnelles. Le recyclage de tournures de phrases, d'expressions, du vocabulaire, de liens, des photographies de profil ou même de messages entiers s'apparente à la mobilisation d'autant d'accessoires de déguisement que les usagers mettent, sans forcément le savoir, à disposition des robots sociaux. La remise en circulation des contenus les plus populaires devient une manière simple et mécanique de donner l'illusion de la participation pertinente au flux des échanges (Freitas *et al.*, 2015 ; Marra, 2011). La stratégie des *bots* s'apparente à celle de la punaise réduve, insecte qui attache les carcasses vides de ses victimes sur son dos pour ne pas être identifiée comme proie (Jackson et Pollard, 2007). Les *bots* récupèrent les messages d'autrui (privés de leur contexte d'énonciation) comme technique de déguisement pour rentrer en contact avec des humains sans être détectés. L'extraction et la réutilisation des contenus fonctionnent comme des sources d'inspiration. Elles incarnent une des modalités de l'apprentissage, de l'humanisation et du camouflage des entités automatisées que sont les *bots*.

Ces observations liminaires nous conduisent à considérer avec attention la complexité d'un terrain d'investigation constitué d'une forte imbrication d'échanges en ligne dont une partie au moins est fondée sur le camouflage et le subterfuge. Elle souligne l'intérêt d'approches sociologiques capables de prendre pour objet cette situation d'encastrement des contributions humaines et des robotiques. L'enjeu théorique principal pourrait alors se formuler de la manière suivante : comment analyser cet espace d'interaction hybride sans en écraser les spécificités sociotechniques ?

LES *BOTS* COMME OBJET SOCIOLOGIQUE

Devant la complexité évoquée précédemment, plusieurs chercheurs mobilisent la théorie de l'acteur réseau afin de proposer une sociologie des *bots* (Latzko-Toth, 2017 ; Massanari, 2017 ; Muhle, 2017). Ce cadre théorique, dont l'une des ambitions consiste « à repeupler l'univers sociologique avec les objets d'une manière qui permette de les engager finement dans la construction de la société » (Barbier et Trepos, 2007 : 35), semble en effet correspondre aux enjeux posés par les robots numériques. L'absence de distinction entre des acteurs humains et non humains, les processus de traduction et les réseaux d'acteurs, parmi d'autres[7], apparaissent comme autant de concepts opérationnels pour aborder le fonctionnement des robots « sociaux » *in situ*. Ils offrent

7. Pour une vue d'ensemble des concepts clef de l'approche ANT, voir Akrich, Callon et Latour (2006).

l'opportunité de saisir des entités autonomes, des actants, dans leurs activités de communication.

L'attention accordée par cette sociologie aux porte-paroles, entités représentant une autre entité au sein d'un processus de négociation, fait écho à la situation des robots : ils agissent dans le flux des échanges en ligne et servent de proxy aux concepteurs qui leur délèguent, par avance, des actions à accomplir. Nous pouvons penser ici à la notion de ventriloquie telle que la propose François Cooren. Même si elle n'est pas explicitement mobilisée à propos des robots sociaux, elle offre une piste intéressante pour articuler les notions d'identité déléguée et de prise de parole. Cette métaphore permet de reconnaître l'agentivité de l'interlocuteur tout en montrant comment celui-ci fait aussi parler (souvent implicitement) quelque chose, qui se met alors à agir par le biais de sa performance (voir Cooren, 2010 : 40-41). Les *bots*, à l'instar de la poupée du ventriloque, sont à la fois les énonciateurs des scripts de leur programmation, mais aussi des porte-voix des expressions qu'ils empruntent à leurs interlocuteurs (ou au web). Il serait possible de les appréhender comme le dernier avatar industrialisé d'une pratique culturelle ancestrale comme la ventriloquie.

Les concepts de mobilisation, d'intéressement et d'enrôlement (Callon, 1986), qui décrivent les manières d'arriver à ce que d'autres actants d'un réseau d'acteurs s'alignent en fonction des intérêts singuliers nous paraissent, là encore, opérationnels. Le robot « social » se voit souvent attribuer la mission de capter l'attention des internautes, les inciter à lui répondre, voire les conduire à le suivre dans ses actions. Ils s'apparentent à des agents qui, pour emprunter le vocabulaire d'Anthony Giddens (1984), participeraient par leurs interactions, leurs routines et leur travail d'explicitation à la structuration de l'espace numérique des échanges.

Les approches inspirées de la théorie de l'acteur-réseau (ANT) permettent donc de poser un premier cadrage pertinent pour aborder et décrire le réseau hétérogène que composent *de facto* humains et robots. Il faut cependant souligner immédiatement une limite significative de cette adoption théorique de l'ANT appliquée aux robots en ligne. Si la distinction entre humains et non-humains est aplanie par cette grille de lecture, elle demeure une préoccupation fondamentale pour l'étude des robots sociaux. Formulé de manière lapidaire, l'argument est que ni la souris de laboratoire de Bruno Latour (Latour et Woolgar, 1988) ni les coquilles Saint-Jacques de Michel Callon (Callon, 1986) ne tentent de se faire passer pour des actants humains. Or, cette possible confusion sur l'ontologie des actants, exprimée par les tentatives systématiques de se présenter pour ce que l'on n'est pas, est centrale dans le phénomène qui nous intéresse. Les robots sociaux sont en effet souvent programmés pour entretenir le doute auprès des humains quant à leur nature. Alors que les approches strictement ANT des *bots* se construisent essentiellement à partir de la description de leurs actions, nous postulons que d'autres focales sont nécessaires afin de mieux appréhender les enjeux de cette anthropomorphisation digitale. Il nous semble par exemple éclairant d'intégrer dans l'analyse les arènes au sein desquelles les effets de sens de la présence réelle ou fantasmée des *bots* s'expriment. Nous proposons donc de nous pencher non seulement sur

des situations dans lesquelles cette confusion se réalise, mais aussi sur les instances où elle se trouve explicitée et prise comme objet. Concrètement, nous avons choisi comme terrain de recherche des lieux qui traitent de la coprésence entre humains et non-humains. Dans notre investigation, il s'agit respectivement de l'espace médiatique comme chambre d'écho des prises de position des acteurs du web, et des recherches dédiées à l'étude des robots sociaux.

MÉTHODE

Resté longtemps une thématique marginale, réservée à des spécialistes et peu visible dans l'espace public, le thème des *bots* attire depuis quelques années une attention médiatique grandissante. La reprise par les médias des discours publics issus d'une série de plateformes du web (Facebook, Instagram, Twitter, Snapchat) en ce qui a trait aux robots constitue notre premier thesaurus[8]. Nous partons de l'hypothèse que les représentations mobilisées dans ces discours pour caractériser le phénomène participent pleinement de la construction du sens de la coprésence des bots et des humains. D'abord placée sous les registres de l'anecdotique, de l'expérimental et du ludique, cette cohabitation soulève des craintes, des controverses ou encore des réflexions sur la nécessité de l'instituer en problème public. Une telle situation d'incertitude, de doute, de crise, voire de panique morale, envers des robots sociaux, offre un terrain d'investigation fertile à partir duquel nous avons mené la première partie de notre étude. Notre démarche s'inspire de l'amorce de l'enquête telle qu'elle est définie par Dewey qui s'engage à partir d'une situation de crise, un trouble (Voirol, 2008). Ce trouble, dans le cas qui nous intéresse, concerne l'existence au sein du web social d'une population de robots dont le nombre et le pouvoir d'influence demeurent indéterminés. Les discours médiatiques sur le sujet se construisent sur la base de sources aussi variées que les témoignages, les communiqués de presse des entreprises et la publication d'études scientifiques. Les problématisations que les médias proposent s'appuient fréquemment sur l'interprétation des résultats de recherches. Ces études, principalement issues des *computers sciences*, constituent un autre lieu fort de l'explicitation du rapport entre humain et robots. Pour leur réalisation, ces recherches doivent être en capacité de les distinguer. Les grilles d'analyse du phénomène mobilisées par ces travaux agencent des rôles et des caractéristiques attribués aux robots et aux humains. Elles procèdent ainsi inévitablement d'un travail de tri et de classification *a priori* ou *a posteriori* dont nous avons examiné les caractéristiques à partir d'un corpus d'articles scientifiques[9]. Nous avons d'abord analysé une série d'études retenues pour leurs cita-

8. Nous avons procédé à la récolte systématique des actualités comportant les termes « *bots* », « *social bots* », « *spambot* », « *socbot* », « robots sociaux » au printemps 2017. Outre la manière de présenter le phénomène, nous avons utilisé ces actualités pour identifier les dépêches issues des opérateurs du web ainsi que les études scientifiques convoquées.

9. Le nombre d'articles scientifiques et d'actes de conférences retenus selon cette procédure est de 39. Ils appartiennent pour l'essentiel au domaine de l'informatique et des sciences de la communication. Aucune sélection n'a été effectuée en termes d'origines géographiques ou disciplinaires.

tions des articles de presse en ligne évoquant les robots sociaux. Le corpus a ensuite été étoffé à partir des références mobilisées dans ces articles qui traitaient explicitement de la détection de robots. Cette procédure visait à mieux contextualiser les approches convoquées par les travaux médiatisés au sein de leurs disciplines et permettait d'identifier à la fois les origines et les spécificités des méthodes proposées. Sur la base des publications scientifiques issues de ces recherches, nous nous sommes intéressés aux définitions opérationnelles des robots sociaux : sous quelles formes apparaissent-elles dans les expérimentations conduites ainsi que dans l'interprétation des résultats ? Le but de notre démarche ne vise pas une forme d'exhaustivité ni de représentativité. Il s'agit d'une étude exploratoire visant à comprendre ce qui se joue dans les entreprises d'interprétation du rapport entre robots et humains. Avec ces terrains médiatiques et scientifiques, nous nous détachons également d'une étude directe des dispositifs techniques. Cette prise de recul comporte le risque d'une analyse moins précise de ce qui se joue au niveau de chaque interaction avec les robots. Nous estimons que cette approche, complémentaire, possède cependant l'avantage de contextualiser ces interactions en explicitant comment elles se trouvent socialisées dans les médias et opérationnalisées scientifiquement. Dans une telle perspective, les interactions hommes robots ne sont plus de simples épiphénomènes, des rencontres singulières, mais s'inscrivent d'emblée dans des processus sociologiques de plus grande ampleur.

LA MÉDIATISATION DES ROBOTS SOCIAUX : INVASION, EXCLUSION ET ÉRADICATION

La première modalité de cette structuration commune qui s'affiche notamment dans les discours médiatiques s'exprime par la métaphore de l'invasion : « *L'invasion des robots du Web*[10] », « *Attention, les Social Bots débarquent !*[11] » ou encore, plus définitif, « *Chatbots Are Back And They're About to Take Over*[12] ». Inattendus, cachés, nombreux et indésirables : les surgissements des robots sociaux sont mis en scène comme une menace. C'est la figure de l'envahisseur des espaces publics de discussion en ligne (« *Bots unite to automate the presidential election*[13] »), du faux ami des réseaux sociaux (« *Your Facebook friends may be evil bot*[14] ») ou de l'imposteur des sites de rencontre (« *Scam : une série de bots tente de piéger les utilisateurs de Tinder*[15] ») qui s'affiche en de très nombreuses occurrences. Ces métaphores de l'invasion se déclinent sur différents

10. *Le Monde*, Technologies, 24.02.2014 ; https://www.lemonde.fr/technologies/article/2014/02/24/l-invasion-des-robots-du-web_4372458_651865.html ; consulté le 3 mars 2018.

11. Cnetfrance.fr, 25.10.2013 ; http://www.cnetfrance.fr/news/attention-les-social-bots-debarquent-39795060.htm, consulté le 3 mars 2018.

12. Time, Tech 26.01.2016 ; http://time.com/4194063/chatbots-facebook-messenger-kik-wechat/, consulté le 3 mars 2018.

13. Wired, 15.05.2016 ; https://www.wired.com/2016/05/twitterbots-2/, consulté le 3 mars 2018.

14. InfoWorld, 08.04.2014 ; https://www.infoworld.com/article/2614095/social-networking/your-facebook-friends-may-be-evil-bots.html, consulté le 3 mars 2018.

15. Bitdefender.fr, 07.04.2014 ; https://www.bitdefender.fr/actualite/scam-:-une-serie-de-bots-tente-de-pieger-les-utilisateurs-de-tinder-2902.html, consulté le 3 mars 2018.

registres qui associent le phénomène décrit, entre autres, au vocabulaire militaire, biologique et médical (Tassin et Kull, 2012).

Dans le répertoire militaire, le bot est d'emblée dans la position du vecteur d'une attaque volontaire, une menace extérieure à affronter. Dans le cadre plus large de la cyberguerre, il devient un fantassin de la propagande en ligne destinée à influencer, voire dénaturer, le processus politique du pays cible (Howard et Parks, 2012). Il fonctionne aux côtés des trolls humains stipendiés pour saturer les espaces d'échanges numériques ou encore des virus.

Présentés comme une espèce nuisible, les robots sociaux sont aussi comparés à une infestation biologique parasitaire qui menace l'écosystème des échanges en ligne. Faisant face à la prolifération des *bots* indésirables, les acteurs du web préconisent des initiatives pour les juguler sous forme de battues : « *Instagram part à la chasse aux faux influenceurs et aux bots distributeurs de likes*[16] », « *Chasser les robots pour limiter les fakes news*[17] ». Face à ce type d'invasion, l'une des stratégies utilisées consiste à créer des pièges à *bots*. En disposant en ligne, littéralement, des « pots de miel » (*honeypot*), le but est de les attirer. Ainsi écartés de l'accès aux contenus qu'ils visaient, ils sont rendus inoffensifs.

Considérés comme un agent pathogène, les *bots* sont aussi représentés comme une menace pour la santé d'un organisme plus grand. Ils s'apparentent alors à une infection qui compromet l'intégrité et la qualité des échanges en ligne et donc, potentiellement, la pertinence ou la survie de tout le système. De manière symptomatique, certains grands réseaux sociaux comme Facebook appellent « système immunitaire » l'ensemble des mesures prises pour faire barrage aux *bots* sur leurs plateformes[18]. Ces systèmes de défense évolutifs développent, pour chaque espèce reconnue et considérée comme indésirable, une sorte d'anticorps technologique sur la base de l'identification des signatures comportementales des robots. Ce recours à la métaphore invasive pour les robots sociaux constitue sans doute également une manière d'appréhender, traduire et simplifier un phénomène complexe dont les mécanismes exacts demeurent largement hermétiques pour le public (Chew et Laubicher, 2003).

Il peut sembler singulier d'entamer cette description du vivre-ensemble par cette liste de dénonciations virulentes des robots. Pourtant, nous pouvons aussi lire ces titres comme autant d'affirmations de leur présence, certes problématique, dans les espaces numériques des humains. Ce vivre-ensemble conflictuel révèle que ces entités purement techniques cumulent les tares et mettent à mal les catégorisations habituelles. Espèce nuisible, entités pathogènes, mais aussi agents manipulateurs pouvant influencer les esprits, ils restent difficiles à saisir. Toutes ces identités hétérogènes participent

16. *Mashable,* 5.07.2017 ; http://mashable.france24.com/medias-sociaux/20170705-instagram-influenceurs-bots-instagress, consulté le 3 mars 2018.

17. *24heures*, 27.06.2017 ; https://www.24heures.ch/high-tech/chasser-robots-limiter-fakes-news/story/25976673, consulté le 3 mars 2018.

18. Par exemple, le FIS pour *Facebook Immune System* qui passe en revue plusieurs centaines de milliers d'actions par secondes (Stein, Chen et Mangla, 2011).

de la même logique : elles bousculent « l'entre-soi » des humains communiquant en ligne et soulignent la fragilité intrinsèque de ces échanges qu'ils peuvent perturber. Les discussions autour des *bots* de Twitter illustrent d'une autre manière la nature profonde de cette fragilité en évoquant littéralement le corps *(invasion of the body snatchers*[19]*)*. La possibilité du vol d'identité numérique et de sa réincarnation robotique constitue un phénomène commenté par de nombreux observateurs du web. L'utilisation de photographies dérobées ainsi que d'autres informations aspirées en ligne comme éléments de camouflage pour déguiser la nature des robots participe de ces processus de vampirisation numérique. Les nombreux cas avérés de « bimbots », robots incarnant le profil d'une femme à partir de montage d'informations issues d'individus réels, en sont une illustration[20]. Ce n'est donc pas seulement l'altérité radicale des mécanismes des robots sociaux qui engendre les craintes de l'envahisseur, mais aussi le fait que, pour paraphraser David Vincent dans une série télévisée de la fin des années 1960, « ils sont parmi nous »[21]. Les envahisseurs extraterrestres de cette série ont d'ailleurs en commun, avec les robots sociaux en ligne, de bénéficier d'une apparence humaine pour se mêler à nous et nous tromper.

Cette situation étrange, mais en vérité déjà banale de notre vie commune avec les robots sociaux, instaure une crise du sens quant à la nature « sociale » de ces plateformes numériques. La part de robots dans la population des utilisateurs de certains réseaux sociaux fait l'objet de débats récurrents. Leur présence et leur nombre constituent un sujet délicat pour les promoteurs des grandes plateformes commerciales du web, car ils influencent l'appréciation économique de la valeur de leurs dispositifs déterminant la part des usagers et du trafic généré par de « vrais » utilisateurs humains. Ces derniers hésitent d'ailleurs souvent entre d'un côté une tolérance pour les robots « sociaux », de nature à augmenter les statistiques de fréquentation, et de l'autre des actions de nettoyage destinées à garantir la nature véritablement sociale des échanges. Comme le formule un dirigeant d'Instagram (en échos à des propos similaires de responsables d'autres réseaux sociaux) avant de lancer une large campagne d'éradication de robots, « *keeping Instagram authentic is critical — it's a place where real people share real moments…*[22] ».

Il y aurait beaucoup à déconstruire de la notion d'« authenticité » mobilisée ici. La présence massive des robots sociaux pose effectivement la question de l'authenticité, celle des interlocuteurs, des moments partagés et du statut de nos expériences en ligne. Relativement à une situation qui semble déjà inextricable, s'enclenche une dialectique

---

19. Par exemple, The art of the Twitter bot, le 30.11.2015 http://www.alphr.com/twitter/1002130/the-art-of-the-twitter-bot/page/0/2, consulté le 25.04.2018.

20. Pour un exemple de ces dépossessions, *cf.* « Who's That Woman In The Twitter Bot Profile ? », Fastcompany, 08. 2012, https://www.fastcompany.com/3000064/whos-woman-twitter-bot-profile, consulté le 25.04.2018.

21. *Les Envahisseurs* (*The Invaders*) série télévisée de science-fiction diffusée du 10 janvier 1967 au 26 mars 1968, *cf.* http://www.imdb.com/title/tt0061265/, consulté le 3 mars 2018.

22. Compte Tumblr d'Instagram « 300 Million : Sharing Real Moments », 10.12.2014, http://blog.instagram.com/post/104847837897/141210-300million, consulté le 25.04.2018.

entre crainte de l'invasion, opérations de bannissement massif et volonté de refonder un monde commun entre humains au sein du numérique. Cette question quant à l'ontologie des collectifs en ligne et ses implications se retrouve systématiquement dans un large nombre de domaines d'activités (par exemple les jeux vidéo, le poker en ligne[23], les sites de recommandation ou de rencontres, etc.). Dans chacune de ces situations, la question est identique : la présence des robots dénature-t-elle la raison d'être de ces sites ? Cette question offre simultanément de multiples terrains d'expérimentation de nouveaux robots comme des occasions de mise en œuvre de tests de détection et de procédures d'éviction (Hingston, 2012). Elle est à l'origine de plusieurs études que nous proposons d'examiner en nous attachant plus spécifiquement à la manière dont ces travaux, en cherchant à comprendre les robots, proposent une lecture de l'humain (Bollmer et Rodeley, 2017).

DISTINGUER LES ROBOTS EN DÉFINISSANT LES HUMAINS

L'identification des robots sociaux a engendré de nombreux travaux (Wang, 2010 ; Chu *et al.*, 2010 ; Davis *et al.*, 2016 ; Ferrara *et al.*, 2016) qui proposent des méthodes pour les distinguer du reste de la population des utilisateurs. Le fait que les *bots* mobilisent souvent des scripts rudimentaires pourrait conduire à estimer, *a priori*, qu'ils sont aisément repérables. À l'examen des modes de fonctionnement décrits dans les articles scientifiques, on constate qu'aujourd'hui encore, une grande majorité des robots sociaux restent rudimentaires. Ils ne résisteraient pas plus de quelques minutes à un test de Turing que l'on ferait passer en face à face dans les règles de l'art. La difficulté reconnue de leur identification tient moins à l'intelligence présumée de ces dispositifs qu'à la manière dont ils exploitent l'environnement au sein duquel ils déploient leurs actions. En conséquence, un des principaux obstacles auxquels doivent faire face ces études consiste à saisir les robots sociaux *in situ*.

Ainsi qu'évoqué précédemment, cet environnement s'avère très propice à diverses formes de mixité informationnelle avec les humains. Pour les chercheurs, c'est au sein de ces bases de données qu'il convient d'opérer un travail de tri, tout en ignorant si ces traces sont la traduction d'une intentionnalité humaine directe ou des productions automatiques. Le travail de discrimination des non-humains passe par la détection de signatures comportementales jugées significatives. Les études font référence à une très large palette des critères mobilisés pour la détection de cette signature robotique permettant de distinguer les humains des non-humains, comme la syntaxe des messages, leur sémantique, le comportement de l'usager, ses liens avec son voisinage numérique, la temporalité de ses actions (Subrahmanian *et al.*, 2016).

C'est un véritable travail de détective statistique qui débute alors. La syntaxe des messages et les conversations de l'utilisateur portent-elles des traces d'autogénération de message à l'exemple d'ELIZA ? Le nombre de hashtags, la quantité de liens présents

23. Site Cardschat « Rise Of The Machines : How Poker Bots Infiltrated The Online Game » https://www.cardschat.com/poker-bots.php, consulté le 3 mars 2018.

dans un message ou encore la manière de terminer un message sont d'autres indices révélateurs. Sur le plan sémantique, les sujets abordés, leur récurrence, la cohérence dans le temps des opinions émises ainsi que le caractère tranché de ces opinions servent de tamis de triage possible. La temporalité offre une autre catégorie de radar : la fréquence de la génération du message, la distribution de ces messages en fonction des moments de la journée sont prises en compte. La réactivité, la rapidité à composer un message, le nombre de suiveurs et de comptes suivis, l'absence ou la surabondance d'activité constituent d'autres critères encore, exploitables pour ce travail de catégorisation. Le profil de l'entité considérée peut également être un lieu d'investigation de choix pour évaluer la non-humanité : le profil est-il récent ? Est-il équipé de photographie(s) ? Est-il associé à une URL ? Le nom semble-t-il produit de manière procédurale ? C'est aussi en utilisant ces critères avec des pondérations spécifiques que les responsables des réseaux sociaux décident des procédures d'éviction massive des comptes jugés suspects. Il n'existe malheureusement pas à notre connaissance de recension du nombre de faux positifs, c'est-à-dire d'humains classés et exclus comme des robots. Il existe cependant de nombreux témoignages d'utilisateurs ayant reçu des messages indiquant que leur compte est bloqué, car leur comportement semblait être celui d'un robot[24].

La détection des robots ne se limite pas à constituer un domaine de recherche prolifique ou un moyen d'action sur la population des plateformes numériques. Il devient aussi un savoir-faire pratique. Des sites permettent d'auditer les comptes Twitter et offrent de calculer le pourcentage de chances que derrière le profil se cache une entité automatisée[25]. Le web contient de nombreux guides destinés à aider l'utilisateur à la détection des profils suspects[26]. Ces sites s'inspirent d'expériences subjectives, mais reprennent aussi, sous la forme de conseil pratique, plusieurs recherches. Le profil du suspect s'esquisse alors en de multiples termes. Un utilisateur faisant un usage excessif de citations d'autrui ou de liens hypertextes est suspect. Un compte qui poste très peu, qui n'affiche pas de photo de profil (ou qui garde l'image proposée par défaut) peut éveiller des soupçons. Un profil dont le dernier message remonte à plus de 30 jours, une entité qui suit beaucoup plus de personnes qu'elle ne se trouve suivie peut s'avérer douteuse. Au titre des éléments potentiellement à charge, relevons également des publications trop régulières (postée aux mêmes heures) ainsi que la répétition des mêmes propos. Finalement, une photographie de profil trop « belle » doit aussi être de nature à éveiller les soupçons[27]. Plusieurs conclusions peuvent être tirées de ces inventaires de comportements suspects. Premièrement, notons qu'ils proposent des critères

24. Voir par exemple le site de Gizmodo « Tons of People Can't Tweet Because Twitter Thinks We're All Robots [It's Fixed!] », le 02.10.2015, https://gizmodo.com/tons-of-people-cant-tweet-because-twitter-thinks-were-a-1685072195, consulté le 3 mars 2018.

25. Voir par exemple : BotOrNot (http://botornot.co), consulté le 25 avril 2018.

26. Voir par exemple : http://www.madmoizelle.com/instagram-inconnus-bots-775665, consulté le 3 mars 2018.

27. http://botornot.co/blog/index.php/how-to-spot-fake-twitter-followers-bots/, consulté le 3 mars 2018.

qui constituent autant de dimensions où s'affrontent des stratégies de détection et de furtivité. À l'image de l'évolution des armes et des armures, aux astuces permettant d'identifier les *bots* répondent de nouvelles formes de camouflage sociologique. Les programmateurs de *bots* affinent les formes de mimétisme comportemental que leurs créations mobilisent. Certains *bots* fabriquent l'impression de dormir (ils ne publient pas pendant 6 à 8 heures), leur réactivité est amoindrie (latence aléatoire avant réponse), et ils semblent hésiter (la vitesse de frappe de leurs messages réplique un rythme humain), etc.

Ces indices conduisent inévitablement de nombreux usagers de ces plateformes à remplir tout ou partie de ces critères. Les traces de leur comportement les placent alors à la lisière de l'humanité numérique telle qu'elle est définie ici. Ainsi, toute tentative de distinction entre humains et robots sur la seule base de leurs traces numériques produit, mécaniquement, non pas deux, mais bien quatre catégories : 1) des humains dont les traces ressemblent à celles d'humains ; 2) des robots laissant des traces qui correspondent à celles d'un robot ; 3) des robots qui produisent des traces similaires à celles des humains ; et 4) des humains laissant des traces qui les associent à une identité de robot.

Un site de comparaison de poésies humaines et robotiques[28] permet de mieux comprendre cette situation : après lecture d'un court poème, les visiteurs sont appelés à en identifier son auteur, humain ou robot. Le palmarès distingue le poète humain le mieux identifié comme tel, le robot le plus à même de se faire passer pour un humain et le robot dont la nature ne fait pas de doute. La dernière catégorie, celle du poète humain dont les visiteurs considèrent qu'il écrit comme un robot, nous semble essentielle même si encore peu étudiée. Dans ce cas de figure, ce n'est plus seulement la qualité d'une performance ou d'une imitation qui se trouve évaluée mais une ressemblance par défaut : l'humain confondu avec une machine. Cette catégorie pose des questions clés : quelles sont les caractéristiques nécessaires au comportement d'un humain pour être reconnues comme telles ? Comment ces caractéristiques évoluent-elles avec les nouvelles fonctionnalités des bots ? Quelles sont les conséquences (sociales et techniques) du fait d'être considéré par erreur comme une machine ? Autrement dit, ces situations de confusion ontologique illustrent la manière dont les robots sociaux interrogent les spécificités de l'expression de notre identité en ligne. À cause d'eux, nous devons de plus en plus souvent faire la preuve de notre humanité. Ainsi, notre rapport aux robots sociaux s'avère singulier, non seulement parce qu'il s'agit de dispositifs qui peuvent revendiquer une forme numérique anthropomorphisée crédible, mais parce que ce faisant, ils déplacent les balises de notre identité.

Les contenus traitant de l'expression des opinions sont typiquement considérés comme une piste fertile pour détecter les robots (Dickerson, Kagan et Subrahmanian, 2014). Au terme de leur étude, les chercheurs tirent les conclusions suivantes : « les robots changent moins souvent d'avis que les humains ; en termes d'opinion, quand

28. Voir un autre site intitulé ButOrNot : http://botpoet.com, consulté le 3 mars 2018.

les humains expriment des opinions positives, ils ont tendance à exprimer des opinions positives plus fortes que les robots, la même tendance (légèrement plus nuancée) tient pour l'expression d'opinions négatives par des humains et, finalement, les humains sont plus en désaccord avec la moyenne de l'ensemble de l'application Twitter » (Dickerson, Kagan et Subrahmanian, 2014).

Ces résultats s'avèrent particulièrement intéressants, nous permettant de saisir la manière dont ce type de travaux, centré sur l'identification des robots, participe en creux d'une définition de l'humain. Le portrait de l'humain qui se dessine est étonnant. Il se singularise par la vigueur de ses opinions et aussi par son écart à la *moyenne de la population* de Twitter ! Un peu comme si l'humain n'était, déjà, plus au centre de la « normalité » des échanges en ligne.

Illustrant symboliquement cette perte de terrain de l'humain, des études privées révèlent que le trafic généré par les *bots* devient majoritaire. Mis en minorité, les utilisateurs se trouvent devant une situation dans laquelle les opérations de distinction se complexifient[29].

La généralisation de l'usage de robots conversationnels « légitimes » complique encore cette situation : les apports d'agents automatiques bénéfiques, les « *good bots* », ne doivent pas se confondre avec ceux des robots malfaisants. Il ne s'agit plus simplement de distinguer entre humains et non-humains, mais entre bons robots et mauvais. Le « DARPA[30] Twitter BoT Challenge » illustre les difficultés soulevées par ces impératifs de distinction. Il est organisé sous la forme d'une compétition, où six équipes de chercheurs s'affrontent pour découvrir les méthodes les plus efficaces pour détecter des cohortes de robots influenceurs[31]. Cette compétition propose une vision théorique du monde assez simple. D'un côté, se trouvent les entités dont le déploiement est considéré comme légitime et positif, de l'autre, les robots qui peuplent les mêmes espaces, mais dont les finalités sont jugées comme une menace. La nature de la menace demeure assez prosaïquement définie : ils « posent un réel danger à la liberté d'expression » (Subrahmanian *et al.,* 2016 : 38). Leur éradication s'exprime donc comme une nécessité. Aucun détail n'est malheureusement fourni sur la nature de cette liberté d'expression en ligne ou sur la manière de définir la légitimité, l'éthique ou la responsabilité de l'action des robots sociaux.

Le principe de la compétition implique qu'il doit être possible de détecter et de réagir contre les actions d'un ensemble coordonné d'agents néfastes. L'intérêt de ce cas, pour notre propos, tient au fait que les robots ne sont plus évoqués comme des entités

29. Source : Imperva Incapsula Bot Traffic Report 2016, https://www.incapsula.com/blog/bot-traffic-report-2016.html, consulté le 3 mars 2018.

30. DARPA : Agence pour les projets de recherche avancée de défense, États-Unis.

31. Le challenge utilise un monde virtuel commun artificiellement peuplé, entre autres, d'une collection de robots influenceurs « nocifs » connus. Cet environnement fait office de vérité terrain (*ground truth*) servant de base pour évaluer l'efficacité des méthodes de capture proposées par les équipes en compétition.

solitaires, mais comme des collectifs[32]. Dans le contexte des débats sur la manipulation des citoyens en périodes électorales, le collectif de robots sociaux devient le suspect idéal. Agents masqués, souvent séducteurs dotés en théorie d'une réactivité inégalable ainsi que d'une capacité à se démultiplier rapidement, ils constituent une multitude anonyme, mais potentiellement omniprésente (Murthy *et al.*, 2014). Le collectif de *bots* s'apparente à une variante mécanique d'une forme sociale de type *Anonymous* : une foule indéterminée, « ouverte », innombrable et non identifiée, mobilisable à tout instant (Kaufmann, Luque et Glassey, 2016). Dans ce contexte, la stratégie des équipes concourant pour le challenge de la DARPA consiste à en capturer quelques-uns puis — là se situe la composante centrale de cette compétition — d'en extirper les traits saillants. C'est à partir de ces traits qu'il devient possible d'extrapoler une identification systématique des complices dans la masse des suspects possibles.

Considérées dans leur ensemble, ces recherches sur l'identification de robots apparaissent comme des entreprises de fabrication de collectifs basés sur des traces numériques. Elles proposent des regroupements d'entités fondés sur un faisceau d'indices jugés comme similaires. Pour les robots comme pour les humains, c'est à l'aune des traces numériques des autres entités que leur signature est établie et que leur identité est validée. Ils partagent alors des destinées algorithmiques similaires (être reconnus comme robots, être considérés comme humains) dans cet univers numérique commun.

Cette similarité de destinée ne constitue pas une fatalité synonyme d'un écrasement ontologique inévitable. Elle s'opère dans le cadre du numérique et il reste toujours possible de tenter de faire jouer humains et robots de leurs différences pour les singulariser. C'est par exemple le principe du test de défi-réponse utilisé en informatique, pour s'assurer qu'une réponse n'est pas générée par un ordinateur : le CAPTCHA, acronyme de « Completely Automated Public Turing test to tell Computers and Humans Apart ». L'histoire de ce dispositif et son évolution constante pour tenter d'endiguer la venue de robots est exemplaire. Dans l'une de ses variantes les plus modernes, le reCAPTCHA de Google traduit une promesse moins en termes de séparation entre humain et non-humain qu'en termes de discrimination. À l'aide du slogan « Tough on Bots Easy on Human », le reCAPTCHA de Google est en effet présenté comme conçu de manière à mener la vie dure aux robots, tout en améliorant l'accessibilité des humains aux sites web[33]. Pour l'entreprise qui proclame sur le même site « Stop a bot. Build a bot », les réponses des humains qui attestent de leur qualité d'humain servent aussi à nourrir le développement de la nouvelle génération d'intelligence artificielle. Alors que le but de ce dispositif est explicitement une volonté de reconnaissance des différences, il participe simultanément à des circulations d'informations entre humains et robots. Accessoirement, mais ce n'est pas anodin, le géant

32. L'analyse de l'action collective d'un réseau de robots, un botnet, constitue un domaine d'investigation à part entière (Abokhodair, Yoo et McDonald, 2015).

33. « reCAPTCHA knows when to be easy on people and hard on bots », site Google reCAPTCHA https://www.google.com/recaptcha/intro/android.html, consulté le 3 mars 2018.

du web en profite pour faire travailler des millions de personnes au service de programmes de reconnaissance d'image.

Distinguer entre robots et humains n'est donc pas une opération aisée, simple ou linéaire. Tout travail de catégorisation des robots produit et institue, *de facto*, une catégorisation de la « nature » de l'humain, elle-même difficile à cerner en ligne. La situation s'avère encore plus complexe quand on examine les situations non conflictuelles et les collaborations explicites ou implicites déjà à l'œuvre.

LA VIE COMMUNE AVEC LES ROBOTS SOCIAUX

Le territoire commun des robots et des humains ne s'épuise bien évidemment pas dans les seules logiques d'invasion, de confrontation et de détection. Des formes de cohabitation plus symbiotique sont observables. Quand les réseaux sociaux comme Twitter ou Instagram décident de nouvelles vagues de suppression de robots sociaux, certains usagers s'en émeuvent. C'est ainsi le cas lors de la procédure d'élimination massive lancée en 2015 par Instagram, connue sous le nom de « *The Instagram Rapture* » (L'enlèvement d'Instagram). Cet épisode conduisit à l'exil numérique de près de 18,9 millions de *followers* soupçonnés d'être des *bots* (Bollmer et Rodley, 2017).

Cette émotion n'exprime pas un attachement affectif pour ces entités automatisées (dont ils ignorent souvent le nombre dans la liste de leurs contacts). Plus prosaïquement, la perte de nombreux suiveurs est considérée par les utilisateurs comme une atteinte à leur notoriété numérique. Agents stipendiés ou accidentels de leur visibilité sociale, ces entités participent de fait à la construction de leur popularité numérique[34]. Les réactions relativement aux dévaluations de réputation causées par la suppression de robots sont riches d'enseignement pour la sociologie. Elles montrent par exemple que, pour certains internautes, il vaut mieux être accompagné d'un robot que de donner l'image d'être seul (Li, 2015). Dans la dérive inflationniste de la notoriété en ligne, de nombreuses études soulignent la part importante de faux suiveurs qui constituent l'audience fidélisée des personnages publics en ligne (Furnas et Gaffney, 2012 ; Martosko, 2013), sans que des mesures drastiques ne semblent travailler à supprimer ce phénomène. En l'état, c'est un peu comme si nous avions accepté qu'une part de notre identité sociale en ligne se joue, à la marge et pour le bien de nos réputations, en collaboration avec les robots.

Symétriquement, la popularité des robots est établie avec l'apport des humains. L'évaluation de la qualité d'un robot social se trouve régulièrement mesurée à l'aune de sa capacité de séduction des humains (nombre de *followers,* de *retweet,* ou likes générés, etc.) (Freitas et Benevenuto, 2015). Les humains victimes de ces sirènes numériques ou partenaires volontaires participent ainsi du benchmark de l'efficacité robotique[35]. De la

34. Voir, par exemple, l'article du *New York Times* du 6 juin 2017 intitulé « How Bots Are Inflating Instagram Egos ».

35. Sur ce point, voir l'exemple du projet Web Ecology et sa compétition pour robots sociaux « *Social bots 2011, Help Robots Take Over The Internet* », http://www.webecologyproject.org/studies/ et McMillan (2012).

même manière, les traits de caractère que les humains préfèrent chez les *bots* sont explorés (Akinori et Hayashi, 2016). Les deux catégories d'entités se retrouvent finalement associées aux mêmes métriques de la notoriété (typiquement le score Klout[36]) au sein desquels leur visibilité numérique est calculée.

Les formes de collaboration explicite impliquent des moments où humains et non-humains se relayent. Concernant les tentatives de manipulation, les chercheurs soulignent que les situations qui hybrident volontairement humains et robots sociaux échappent facilement aux tentatives de détection. Typiquement, l'humain intervient lors de la création du compte, afin d'aider le *bot* à franchir la barrière d'un CAPTCHA. Ce sont aussi les exemples de coopérations hybrides entre *trolls* et *bots* dont les actions conjointes procèdent souvent d'un renforcement mutuel (effet démultiplicateur pour les *bots*, effets de validation par les *trolls*). Wikipedia illustre cette collaboration explicite à une très large échelle. Depuis 2002, des milliers de *bots* sillonnent l'encyclopédie, agissant le plus souvent dans l'ombre, sans construire de dialogue direct avec les utilisateurs humains. Ils façonnent pourtant le contenu constamment et participent pleinement aux échanges entre les contributeurs. Ils possèdent une signature, une identité et s'impliquent même dans des conflits (Tsvetkova *et al.*, 2017). Plus largement, cette collaboration requiert de définir les caractéristiques de robots compatibles avec une socialisation avec les humains. Les recherches qui posent la question d'un possible vivre-ensemble avec les robots déterminent de possibles prémisses d'une nétiquette robotique. Au titre de ces règles, on trouve par exemple l'affichage visible du statut de robot, l'interdiction d'amorcer des échanges informationnels sans le consentement des humains concernés ou encore la transparence quant aux usages des données générées pendant les échanges (Maréchal, 2014).

L'annonce de nouvelles arrivées massives de robots (« Facebook Messenger se prépare à l'invasion des robots »[37]) offre un signe de l'ampleur de ce repositionnement collaboratif. Ici, le terme d'invasion change de connotation et devient positif. Motivés par les occasions de rationalisation et d'économie d'échelle qu'offre l'automatisation personnalisée des interactions avec leurs clients, les géants du web plébiscitent les robots conversationnels.

Le développement de robots sociaux occupe une place de choix dans les plans stratégiques et la planification des principaux acteurs du web[38]. Ceux-ci encouragent les développeurs indépendants à produire des robots au sein de leur écosystème. Ils ouvrent même l'accès à une partie de leurs algorithmes d'intelligence artificielle[39]. Associés au discours relatif au big data et au Deep Learning, une partie des *bots*

36. Klout est une application qui utilise des informations extraites de Twitter, Facebook, Google+, LinkedIn, Foursquare, Wikipedia, et Instagram pour établir l'activité sur les médias sociaux et calculer une « influence » sous la forme d'un score.

37. *Cf.* l'express L'Expansion du 13.04.2016, consulté le 3 mars 2018.

38. Citons, pour les principales plateformes de développement des *bots* en 2017 : Amazon Echo, Android, Cisco Spark, Discord, IMessage Facebook Messenger, Skype, Slack, Telegram, Twitter, Viber, Wechat.

39. Par exemple l'initiative TensorFlow de Google ou l'initiative d'OpenAI.

deviennent des dispositifs légitimes, les « bons » robots dont nous parlions plus haut. Ce sont alors typiquement des agents qui prennent en charge de manière automatisée la relation aux usagers clients à l'image du robot coach du site de rencontres Meetic ou de l'interface conversationnelle de WeChat en Chine. Dans les scénarios d'usage de ces dispositifs, un robot qui se trouve devant une demande trop complexe ou imprévue peut alors métaphoriquement passer la main à un opérateur humain. Les coulisses des marchés gris des validations sociales, les activités conjointes avec les *trolls*, sur Wikipedia, les agents conversationnels offrent autant d'exemples du fait qu'humains et robots sociaux interagissent déjà sans cesse de manières implicites ou explicites. Ces collaborations et leurs développements futurs ouvrent un autre chapitre important de notre rapport aux machines, celui de la répartition des tâches. Les robots sociaux, par leur capacité à traiter automatiquement certaines tâches, mais aussi par leur capacité d'en déléguer tout aussi automatiquement certaines autres aux humains, joueront sans doute un rôle central et ambivalent dans la future définition de notre rapport au travail.

Au prisme du rapport entre humains et non-humains en ligne, la problématique de la cohabitation évolue sensiblement. L'enjeu n'est plus la stricte capacité à discriminer ou à camoufler la nature de ces entités. Il s'agit d'explorer les possibilités offertes par la mixité explicite, assumée et collaborative d'agents humains et robots. Des scénarios explorés actuellement comme ceux impliquant des robots sociaux qui incitent les humains à tenir des discussions constructives et démocratiques (Graham et Ackland, 2017) proposent d'autres modalités de la présence des robots au sein des collectifs d'humains. Dans ces visions, les robots ne seraient à terme plus seulement appelés à imiter le social, mais également à le cultiver.

CONCLUSION

L'examen des rapports entre humains et non-humains au sein de nos infrastructures numériques en est encore à ses balbutiements. Longtemps, cette approche s'est particulièrement concentrée sur une lecture individuelle du rapport entre une entité-personne et une entité-machine, avec une focale sur le thème de l'intelligence de cette dernière. Riches en enseignements, ces travaux s'intéressent peu au fait que les rencontres entre humains et non-humains se déroulent dans une grande diversité de contextes qui les influencent très directement. Plus gênant, ils font aussi souvent l'impasse sur le fait que ces dialogues avec des machines se déploient au sein de collectifs d'humains et de machines qui en transforment sans cesse la portée et le sens. De nombreuses recherches sociologiques débutent sans doute là où se termine la présente contribution. La question de la réception, de la production du sens chez les personnes interagissant avec des systèmes conversationnels automatisés constitue un domaine clé, parmi d'autres, que nous n'avons malheureusement pas traité ici. Nous n'avons pas non plus fait référence à de possibles travaux qui s'intéresseraient aux formes d'exclusion de l'humain que peuvent engendrer les robots quand ils interagissent entre eux. Par leur réactivité, ils produisent ainsi souvent, au sens littéral du terme, des *no man's*

land, c'est-à-dire des univers interactionnels au sein desquels l'humain n'a plus la capacité physiologique de s'immiscer (Madrigal, 2014).

En amont de ces développements, notre contribution s'interrogeait sur la nécessité de considérer la coprésence des humains et des robots sociaux comme une dimension clé de toute réflexion sociologique sur le numérique. À la lumière de nos explorations, il nous semble difficile d'envisager une approche du social en ligne s'exonérant de l'examen de ces entités. Pourtant, tant sur le plan des concepts que des méthodes, beaucoup reste à inventer. Comment intégrer dans la sociologie des usages du numérique le fait que le social en ligne est largement travaillé par des entités non humaines anthropomorphes ? Comment identifier, qualifier, quantifier et pondérer cette présence dans les analyses sociologiques qui s'intéressent aux rapports entre humains ? En équilibre instable entre les tentations de « technologisation » du social et, à l'inverse, celles d'une socialisation du technique, les robots sociaux nous permettent de prendre précisément cet entre-deux comme objet. C'est déjà, sans doute, une contribution à porter à leur crédit.

RÉSUMÉ

L'article se concentre sur les implications théoriques et empiriques de la coprésence d'humains et de robots sociaux au sein des espaces de sociabilité numériques. Il explore les modalités de cette coprésence en examinant deux types de situations que sont les discours médiatiques relativement au trouble occasionné par les robots et la manière dont les chercheurs des *computer studies* analysent cette présence en ligne. L'article montre que ces processus d'identification et de catégorisation des robots sociaux participent simultanément d'un travail de redéfinition de ce qui est le propre de l'humain. Sur cette base, il défend la thèse qu'il devient difficile d'envisager une sociologie des usages du numérique qui s'exonère de l'examen de ces entités non humaines anthropomorphes.

Mots clés : Robots sociaux, anthropomorphisation, sociabilité numérique.

ABSTRACT

Rethinking digital sociology through our shared life with social bots

The article focuses on the theoretical and empirical implications of humans and social bots copresence in digital social spaces. It examines how these entities are constructed trough media narratives and computer studies researches about the trouble they cause. The article shows how social robots' identification and categorization processes are also redefinitions of what constitute humanness. On this basis, the article defends the thesis that it becomes difficult to imagine a digital sociology that is exempt from the examination of these anthropomorphic nonhuman entities.

Keywords : social bots, anthropomorphization, digital sociability

RESUMEN

Este artículo se centra en las implicaciones teóricas y empíricas de la copresencia de humanos y robots sociales dentro de los espacios de sociabilidad digitales. Aquí se explora las modalidades

de esta copresencia, examinando dos tipos de situaciones que son los discursos mediáticos en relación con el trastorno causado por los robots y la forma como los investigadores de *computer studies* analizan esta presencia en línea. El artículo muestra que estos procesos de identificación y categorización de los robots sociales participan simultáneamente en un trabajo de redefinición de lo que es propio del ser humano. Sobre esta base, defiende la tesis por la cual resulta difícil imaginar una sociología de los usos de lo digital que se exonera del examen de estas entidades no humanas antropomorfas.

Palabras clave: Robots sociales, antropomorfización, sociabilidad digital.

BIBLIOGRAPHIE

Abokhodair, N., D. Yoo et D. W. Mcdonald (2015), « Dissecting a Social Botnet : Growth, Content and Influence in Twitter », *in* Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15), ACM, New York, p. 839-851.

Akinori, A. et M. Hayashi (2016), « On communication assistance via bots — towards IMDJ », *in Procedia Computer Science,* vol. 96, p. 1657-1665.

Akrich, M., M. Callon et B. Latour (dir.) (2006), *Sociologie de la traduction : textes fondateurs*, Paris, Presses des Mines, coll. « Sciences sociales ».

Bakardjieva, M. (2015), « Rationalizing Sociality : An Unfinished Script for Socialbots », *The Information Society*, vol. 3, n° 3, p. 244-256.

Barbier, R. et J. Y. Trepos (2007), « Humains et non-humains : un bilan d'étape de la sociologie des collectifs », *Revue d'anthropologie des connaissances*, vol. 1, n° 1, p. 35-58.

Bollmer, G. et C. Rodley (2017), « Speculation on the sociality of Socialbot », *in* Gehl, R. W. et M. Bakardjieva (dir.), *Socialbots and their friends*, Londres, Routledge, p. 147-163.

Callon, M. (1986), « Éléments pour une sociologie de la traduction. La domestication des coquilles Saint-Jacques dans la baie de Saint-Brieuc », *L'Année sociologique*, vol. 36, p. 169-208.

Cardon, D. (2008), « Le design de la visibilité. Un essai de cartographie du web 2.0 », *Réseaux*, vol. 6, n° 52, p. 93-137.

Chew, M. K. et M. D. Laubichler (2003), « Natural Enemies — Metaphor or Misconception ? », *ScienceMag*, vol. 301, p. 52-53.

Christian, B. (2011), *The most Human Human. What Talking with Computers Teaches Us About What it Means to Be Alive*, New York, Doubleday.

Chu, Z., S. Gianvecchio, H. Wang et S. Jajodia (2010), « Who is tweeting on Twitter : Human, bot, or cyborg ? », *Annual Computer Security Applications Conference* (ACSAC), ACM, p. 21-30.

Cooren, F. (2010), « Ventriloquie, performativité et communication. Ou comment fait-on parler les choses », *Réseaux*, vol. 5, n° 163, p. 33-54.

Davis, C. A., O. Varol, E. Ferrara, A. Flammini et F. Menczer (2016), « BotOrNot : A system to evaluate social bots ». Actes de la 25th International Conference Companion on World Wide Web (p. 273-274). 2016 ArXiv : 1602.00975 [cs].

Dickerson, J. P., V. Kagan et V. S. Subrahmanian (2014), « Using sentiment to detect bots on Twitter : Are humans more opinionated than bots ? », *International Conference on Advances in Social Networks Analysis and Mining (*ASONAM), IEEE/ACM, août 2014, p. 620-627.

Ferrara, E., O. Varol, C. Davis, F. Menczer et A. Flammini (2016), « The Rise of Social Bots », *Communications of the ACM*, vol. 59, n° 7, p. 96-104.

Finley, K. (2013), « Twitter "joke bots" shame human sense of humor », *Wired*, https://www.wired.com/2013/08/humor-bots/, consulté le 3 mars 2018.

Floridi, L. (2014), *The Fourth Revolution : How the Infosphere is Reshaping Human Reality*, Oxford, Oxford University Press.

Ford, H., E. Dubois et C. Puschmann (2016), « Automation, Algorithms, and Politics. Keeping Ottawa Honest — One Tweet at a Time ? Politicians, Journalists, Wikipedians and Their Twitter Bots », *International Journal of Communication*, vol. 10, n° 24, p. 4891-4914.

Freitas, C., F. Benevenuto, S. Ghosh et A.Veloso (2015), « Reverse engineering socialbot infiltration strategies in twitter », *in Advances in Social Networks Analysis and Mining* (ASONAM), 2015 IEEE/ACM International Conference, IEEE, p. 25-32. arXiv : 1405.4927.

Furnas, A. et D. Gaffney (2012), « Statistical probability that Mitt Romney's new Twitter followers are just normal users : 0 % », *The Atlantic*, http://www.theatlantic.com/technology/archive/2012/07/statistical-probability-that-mitt-romneys-new-twitter-followers-are-just-normal-users-0/260539/, consulté le 13 juillet 2017.

Gehl, R. W. et M. Bakardjieva (2017) « Sociabots and Their Friends », *in* Gehl, R.W. et M. Bakardjieva (dir.), *Socialbots and Their Friends*, Londres, Routledge, p. 1-16.

Giddens, A. (1984), *La Constitution de la société*, trad. M. Audet, Paris, PUF.

Glassey, O. (2015), « Des communautés virtuelles au Big Data : les formes sociales des promesses des technologies de l'information », *in* Audétat, M. (dir.), *Sciences et technologies émergentes : pourquoi tant de promesses ?,* Paris, Éditions Hermann, p. 115-130.

Graham, T. et R. Ackland (2017), « Do socialbots dream of popping the filter bubble ? The role of socialbots in promoting deliberative democracy in social media », *in* Gehl, R. W. et M. Bakardjieva (dir.), *Socialbots and Their Friends*, Londres, Routledge, p. 187-206.

Granjon, F. (2011), « De quelques pathologies sociales de l'individualité numérique. Exposition de soi et autoréification sur les sites de réseaux sociaux », *Réseaux*, vol. 167, p. 75-103.

Guilbeault, D. (2016), « Automation, Algorithms, and Politics| Growing Bot Security : An Ecological View of Bot Agency », *International Journal of Communication*, vol. 10, n° 19, http://ijoc.org/index.php/ijoc/article/view/6135, consulté le 3 mars 2018.

Hern, A. (2015), « Randomly generated tweets by bot prompts investigation by Dutch police », *The Guardian*, 12 février, https://www.theguardian.com/technology/2015/feb/12/randomly-generated-tweet-by-bot-investigation-dutch-police, consulté le 13 juillet 2017.

Hingston, P. F. (dir.) (2012), *Believable Bots, Natural Computing Series*, Berlin, Springer.

Hofstadter, R. et R. Douglas (dir.) (1996), « The Ineradicable Eliza Effect and Its Dangers, Epilogue », *in Fluid Concepts and Creative Analogies : Computer Models of the Fundamental Mechanisms of Thought*, New York, Basic Books, p. 157.

Howard, P. N. et M. R. Parks (2012), « Social Media and Political Change : Capacity, Constraint, and Consequence », *Journal of Communication*, vol. 62, n° 2, p. 359-362.

Hughes, T. P. (1986), « The Seamless Web : Technology, Science, Etcetera, Etcetera », *Social Studies of Science*, vol. 16, n° 2, p. 281-292.

Hwang, T., I. Pearce et M. Nanis (2012), « Socialbots : Voices from the fronts », *Interactions*, vol. 19, n° 2, p. 38-45.

Jackson, R. R. et S. D. Pollard (2007), « Bugs with backpacks deter vision-guided predation by jumping spiders », *Journal of Zoology*, vol. 4, n° 273, p. 358-363.

Jones, S. (2015), « How I learned to stop worrying and love the bots », *Social Media + Society*, vol. 1, n° 1, p. 1-2.

Kaufmann, L., R. Luque et O. Glassey (2016), « Faire être « Anonymous » : figuration et dé-figuration d'un collectif impropre », *Raison publique*, vol. 20, p. 145-174.

Laird, B. et O. Schwartz (n.d.), *Bot or not*. http://botpoet.com, consulté le 3 mars 2018.

Latour, B. et S. Woolgar (1988), *La Vie de laboratoire. La production des faits scientifiques*, Paris, La Découverte.

Latzko-Toth, G. (2017), « The Socialization of Early Internet Bots : IRC and the Ecology of Human-Robot Interactions Online », *in* Gehl, R. W. et M. Bakardjieva (dir.), *Socialbots and Their Friends*, Londres, Routledge, p. 47-68.

Leonard, A. (1997), *Bots : The Origin of New Species*, San Francisco, Wired Books.

Li, S. (2015), « For spambots, flattery gets you everywhere », *The Atlantic*, 15 janvier. https://www.theatlantic.com/technology/archive/2015/01/for-spambots-flattery-gets-you-everywhere/384665/, consulté le 3 mars 2018.

Licoppe, C. (2010), « Les apparitions médiatisées et leurs effets performatifs. Le cas des sonneries téléphoniques et la « crise de la sommation », *Réseaux*, vol. 5, nº 163, p. 131-162.

Madrigal, A. (2014), « That time 2 bots were talking, and Bank of America butted in », *The Atlantic*, 7 juillet.

Maréchal, N. (2016), « Automation, Algorithms, and Politics| When Bots Tweet : Toward a Normative Framework for Bots on Social Networking Sites (Feature) », *International Journal of Communication*, vol. 10, nº 10, p. 5022-5031.

Marra, G. (2011), « Geppetto's army : Creating international incidents with twitter Bots », *SXSW*, http://lanyrd.com/2011/sxsw/scryb/, consulté le 9 février 2015.

Martosko, D. (2013), « Barack Obama is political king of the fake Twitter followers, with more than 19.5 MILLION online fans who don't really exist », *Mail Online*, 24 septembre.

Massanari, A. (2017), « Contested Play : The Culture and Politics of Reddit Bots », *in* Gehl, R. W. et M. Bakardjieva (dir.), *Socialbots and Their Friends*, Londres, Routledge, p. 111-127.

Mcmillan, R. (2012), « Twitter bots fight it out to see who's the most human », *Wired* [Magazine], 7 novembre. https://www.wired.com/2012/11/twitterbots/, consulté le 3 mars 2018.

Metz, R. (2016), « Why Microsoft Accidentally Unleashed a Neo-Nazi Sexbot », *MIT Technology Review*, 24 mars, https://www.technologyreview.com/s/601111/why-microsoft-accidentally-unleashed-a-neo-nazi-sexbot/, consulté le 10 juillet 2017.

Middlebrook, S. T. et J. Muller (2000), « Thoughts on Bots : The Emerging Law of Electronic Agents », *The Business Lawyer*, vol. 56, nº 1, p. 341-373.

Mitter, S., C. Wagner et M. Strohmaier (2014), *Understanding the impact of socialbot attacks in online social networks*. arXiv Preprint arXiv : 1402.6289.

Morozov, E. (2012), « Muzzled by the bots », *Slate*, 26 octobre, http://www.slate.com/articles/technology/future_tense/2012/10/disintermediation_we_aren_t_seeing_fewer_gatekeepers_we_re_seeing_more.html, consulté le 3 mars 2018.

Muhle, F. (2017), « Embodied Conversational Agents as Social Actors ? », *in* Gehl, R. W. et M. Bakardjieva (dir.), *Socialbots and Their Friends*, Londres, Routledge, p. 86-109.

Murthy, D., A. Powell, R. Tinati, N. Anstead, L. Carr, S. Halford et M. Weal (2016), « Automation, Algorithms, and Politics. Bots and Political Influence : A Sociotechnical Investigation of Social Network Capital », *International Journal of Communication*, vol. 10, nº 20, p. 4882-4890.

Neff, G. et P. Nagy (2016), « Automation, Algorithms, and Politics| Talking to Bots : Symbiotic Agency and the Case of Tay », *International Journal of Communication*, vol. 10, nº 17, p. 4915-4931.

Qiu, X., D. F. M. Oliveira, A. S. Shirazi, A. Flammini et F. Menczer (2017), « Limited individual attention and online virality of low-quality information », *Nature Human Behaviour*, vol. 1, nº 132, p. 1-7.

Searle, J. R. (1980), « Minds, Brains and programs », *The Behavioral and Brain Sciences*, vol. 3, Cambridge University Press, 1980, traduction française : Searle, J. R. (1987) « Esprits, cerveaux et programmes », *in* D. Hofstadter et D. Dennett, *Vues de l'esprit*, Paris, Interéditions, p. 354-373.

Seward, Z. (2014), « Twitter admits that as many as 23 millions of its active users are automated », *Quartz*, 11 août.

Stein, T., E. Chen et K. Mangla (2011), « Facebook immune system », *in Workshop on Social Network Systems* (SNS), ACM.

Subrahmanian, V. S. *et al.* (2016), « The DARPA Twitter Bot Challenge », *Computer* [IEEE], vol. 49, nº 6, p. 38-46.

Tassin, J. et C. A. Kull (2012), « Pour une autre représentation métaphorique des invasions biologiques », *in Natures Sciences Sociétés*, vol. 20, p. 404-414.

Tsvetkova, M., R. García-Gavilanes, L. Floridi et T. Yasseri (2017), « Even good bots fight : the case of Wikipedia », *PLoS ONE*, vol. 12, nº 2, p. 1-13.

Turing, A. (1950), « Computing Machinery and Intelligence », *Mind*, vol. 59, nº 236, p. 433-460.

Varol, O., E. Ferrara, C. A. Davis, F. Menczer et A. Flammini (2017), « Online Human-Bot Interactions : Detection, Estimation, and Characterization Computer Science, *Eleventh International AAAI Conference on Web and Social Media*, Montréal, 15-18 mai.

Voirol, O. (2008), « Pluralité culturelle et démocratie chez John Dewey », *Hermès, La Revue*, 51(2), p. 23-28.

Wang, A. H. (2010), « Detecting spam bots in online social networking sites : A machine learning approach », *Conference on Dta and Applications Security and Privacy*, ACM, p. 335-342.

Weil, P. (2017), « The Blurring Test », *in* Gehl, R. W. et M. Bakardjieva (dir.), *Socialbots and Their Friends*, Londres, Routledge, p. 19-46.

Weizenbaum, J. (1966), « Eliza – a computer program for the study of natural language communication between man and machine », *in* Communication of the ACM, vol. 9, nº 1, p. 36-45.

Woolley, S. C. et P. N. Howard (2017), « Computational Propaganda Worldwide : Executive Summary », Working Paper n° 2017.11, University of Oxford, http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Casestudies-ExecutiveSummary.pdf, consulté le 3 mars 2018.