



Collaboration et automatisation dans le transfert de connaissances : perception des agents logiciels par les contributeurs de Wikipédia

Mémoire

Maxime Clément

Maîtrise en médecine expérimentale
Maître ès sciences (M.Sc.)

Québec, Canada

© Maxime Clément, 2015

Résumé

Avec le développement considérable des médias sociaux, les projets collaboratifs de grande ampleur visant au partage de connaissance, tel Wikipédia, doivent gérer des quantités faramineuses d'informations. Cette croissance force ces communautés à fournir de nombreux efforts afin de maintenir la précision et la structure de la base de données. Ceci est particulièrement important pour les informations sensibles, telles que les informations de type médical. Des agents logiciels – appelés « bots » – ont été développés sur la plupart de ces nouveaux médias afin de faciliter la gestion de ces grandes quantités de données, notamment en assurant certaines tâches de maintenance et en effectuant des corrections sur les contributions d'utilisateurs. Toutefois, la manière dont des utilisateurs humains réagissent aux rétroactions réalisées par les bots est inconnue. En utilisant Wikipédia comme modèle, nous avons analysé les différents types de réaction des éditeurs selon les caractéristiques spécifiques liées aux tâches des bots, à partir d'un corpus de 6 528 interventions d'éditeurs humains, sur les pages de discussions de 50 bots actifs entre le 4 janvier 2012 et le 2 janvier 2013, sur les articles de la version anglaise de Wikipédia (soit 2 353 différentes discussions). L'activité des bots qui y est observée est très associée au fonctionnement de la communauté interne de Wikipédia. Les bots dont les activités étaient fortement liées au travail des autres éditeurs (c.à.d. avec un haut degré de contrainte ou de visibilité) déclenchent plus de réponses. En combinant différentes caractéristiques de bots, nous avons pu définir deux idéaux types opposés de bots avec des comportements distincts : les bots « servants » qui réalisent principalement des tâches laborieuses et répétitives à la place des éditeurs humains ; ainsi que les bots « policiers » qui appliquent systématiquement les normes et règles de conduite de Wikipédia. Ces derniers engendrent des réponses davantage polarisées chez les éditeurs. Publiés dans *Computer in Human Behavior*, nos résultats démontrent un étonnant niveau d'acceptation des bots, modulé par différentes réactions selon les types de tâches réalisées par ces derniers. Ceci questionne la capacité de la communauté de Wikipédia à atteindre l'équilibre entre la participation large à l'encyclopédie et sa standardisation.

Table des matières

Table des matières	v
Liste des tableaux	vii
Liste des figures	ix
Remerciements	xiii
Introduction	1
1 Wikipédia : structuration d'une communauté et appropriation d'un cadre technique	5
1.1 Wikipédia, une histoire courte	6
1.2 Wikipédia : entre individus et communauté	13
1.3 Participation à Wikipédia, des signaux contradictoires	24
2 Les bots dans Wikipédia	29
2.1 Apparition des bots	29
2.2 Entre contrôle et liberté	37
2.3 Démonstration par l'exemple : le cas de la protection contre le vandalisme	46
3 Hypothèses et objectifs	49
3.1 Question de recherche	49
3.2 Objectifs de recherche	49
3.3 Hypothèses de recherche	50
4 Méthodes	51
4.1 Protocole de recherche	51
4.2 Sélection des bots	53
4.3 Caractéristiques des bots	55
4.4 Traitement des commentaires et des discussions	55
4.5 Constitution d'idéaux types de bots	58
4.6 Traitement du « panneau d'affichage des propriétaires de bots »	58
4.7 Analyse statistique	59
5 Résultats	61
5.1 Caractéristiques de notre échantillon	61
5.2 Quantification des interactions au sein de l'échantillon	63

5.3	Variation des commentaires selon les variables dichotomiques caractéristiques des bots	64
5.4	Analyse des types de réaction des utilisateurs	66
5.5	Analyse des réactions face aux idéaux types de bots	69
5.6	Analyse des discussions sur le panneau d'affichage des propriétaires de bots	74
6	Discussion	79
6.1	Considération sur la validité générale de notre modèle	80
6.2	La présentation des bots face à la communauté	83
6.3	L'interaction entre les éditeurs et les bots	84
6.4	Les bots et la formation des normes	87
6.5	Les bots et la communauté; l'équilibre des pouvoirs	88
	Conclusion	94
	annexe	99
.1	100
.2	103
.3	104
	Bibliographie	115

Liste des tableaux

1.1	Les cinq piliers de Wikipédia.	9
1.2	Tableau récapitulatif en cinq points quant aux clivages entre les mondes académique et encyclopédique.	10
1.3	Hierarchie de droits et pouvoirs des contributeurs à Wikipédia.	23
2.1	Tableau résumé des premiers bots connus de l'encyclopédie.	31
5.1	Répartition des interventions selon le statut de l'utilisateur dans Wikipédia.	74
5.2	Répartition des discussions selon leur orientation.	75

Liste des figures

1.1	Historique de la croissance en articles et en participants de la version anglaise de Wikipédia.	12
1.2	Triangle dialectique.	19
1.3	Résolution des conflits.	20
1.4	Participation mensuelle pour diverses catégories de contributeurs à l'édition anglaise de Wikipédia.	25
1.5	Taux de rétention versus le nombre d'éditeurs actifs dans le Wikipédia anglais.	26
2.1	Ligne du temps des points marquants touchant l'implantation des bots au sein de Wikipédia	32
2.2	Schéma du processus d'autorisation des bots.	36
2.3	Participation à l'écriture des articles.	38
2.4	Proportion des éditions annulées chez diverses catégories de participants.	39
2.5	Schéma représentant le rôle des bots sur le travail individuel.	43
4.1	Une page personnelle d'un bot, ici Xqbot.	52
4.2	La page de discussion associée à la page personnelle.	53
4.3	Exemples de réactions d'utilisateurs pour les neuf types de réaction observés.	57
5.1	Activité de la population totale des bots de la version anglaise de Wikipédia.	61
5.2	Activité de l'échantillon de bots étudié.	62
5.3	Distributions des interactions en fonction des caractéristiques des bots.	65
5.4	Distribution des commentaires négatifs en fonction des caractéristiques des bots.	66
5.5	Les types de réaction des utilisateurs selon les caractéristiques des bots.	68
5.6	Nombre de discussions en fonction du nombre de critères atteint pour les deux idéaux types de bots identifiés.	70
5.7	Différences au niveau de l'intensité des interactions entre les utilisateurs vis-à-vis des bots «servants» contre les autres bots.	71
5.8	Distribution générale des types de réaction des utilisateurs selon l'idéal type des bots.	72
6.1	Processus dialectique de l'écriture d'un article.	90

« - Les machines font de leur mieux. »

« - Ouais, mais c'est tout c'quelles peuvent faire. Quand y a pépin, l'homme, lui, doit faire plus que son mieux, sans ça il est foutu. »

Nouvelle « Hérité »,
Isaac Asimov

Remerciements

Mes remerciements s'adressent tout d'abord à mon directeur de recherche, M. Matthieu Guitton, qui m'a été d'une aide indispensable au cours de cette maîtrise. Je tiens également à remercier le comité de correction de mon mémoire, pour leurs considérations.

Je voudrais aussi remercier les membres de mon laboratoire de recherche, soit Anna Lomnowska, Cécile Cristofari, Christine Marquilly et Catherine Lortie. Les discussions continues que nous avons eues ont alimenté de façon significative le contenu de ce mémoire. De plus, leur présence aura contribué à faire de mon passage un moment plus qu'agréable.

Enfin, je remercie chaleureusement Carole Dion, Josée Hurtubise, Lynn Lord et Olivia Roy-Malo pour le temps investi à m'aider à peaufiner ce mémoire.

Introduction

Parmi les communautés collaboratives en ligne, Wikipédia fait figure de modèle, autant par son statut d'encyclopédie la plus imposante de l'histoire (Loveland et Reagle, 2013) que par son incontournable présence sur le web, étant régulièrement classée dans les 10 sites web les plus fréquentés mondialement¹, et ce, plus de 10 ans après sa création. Wikipédia est le produit d'un processus enclenché il y a de cela de nombreux siècles, celui de l'accumulation de toutes les connaissances humaines d'une époque donnée. Mais, contrairement aux encyclopédies classiques, Wikipédia est souvent perçue comme une arène de libertés participatives auprès de bien des usagers. Toutefois, il s'agit avant tout d'une communauté d'individus agissant collectivement dans l'objectif commun de structurer et d'archiver les connaissances humaines qu'ils jugent pertinentes. Une question se pose alors : comment réussissent-ils une telle coordination lorsque des dizaines de milliers d'individus sont invités à y participer ?

Une particularité de cette communauté était de prendre racine dans un espace numérique. Ceci donna une occasion unique d'observer les individus s'associer ensemble et développer des objectifs communs. Cette capacité à spontanément s'organiser afin de se doter d'une orientation durable dans leurs actions est facilement observable dans les communautés virtuelles, telle qu'exposée au printemps 2014 avec l'événement très médiatisé de Twitch Play Pokemon². En effet, le 12 février a eu lieu une expérimentation bien inhabituelle sur le réseau social Twitch³, un site spécialisé dans la diffusion vidéo en direct de jeux vidéo. Cette journée a été marquée par la mise en ligne d'une version unique du jeu éponyme Pokémon Red⁴, permettant au public de diriger les actions du personnage principal en entrant les commandes directement dans la fenêtre de messagerie instantanée de Twitch. Malgré une latence moyenne de 20 secondes avant l'exécution de la commande, une foule démesurée y participa. Twitch affirma que l'événement a vu la participation (passive et active) de 1,16 million de personnes différentes avec une moyenne de 70 000 participants simultanés (et d'un taux de participation active de 10 %) ⁵. L'événement a gagné rapidement en popularité, parfois vivant des moments

1. <http://www.alexa.com/siteinfo/wikipedia.org>, consulté le 25 février 2014.

2. <http://www.twitch.tv/twitchplayspokemon>, consulté le 25 février 2014.

3. <http://www.twitch.tv/>, consulté le 3 juin 2014.

4. http://en.wikipedia.org/wiki/Pok%C3%A9mon_Red_and_Blue, consulté le 25 février 2014.

5. <http://arstechnica.com/gaming/2014/03/twitch-beats-pokemon-the-next-adventure-begins-sunday>, consulté le 21 mars 2014.

d'épiphanie ; parfois restant coincé à un passage précis de nombreuses heures, voire des jours. Un mécanisme de gouvernance a donc été implanté au septième jour par le créateur⁶. Ce nouveau mode, dénommé « Démocratie », compilait pour une vingtaine de secondes successives les commandes des utilisateurs avant d'appliquer la plus populaire. Cette modification donna l'occasion à la communauté émergente de s'organiser sur de nouveaux plans. En effet, la détermination du mode de jeu fut laissée au choix des participants, la sommation des commandes sur un certain temps pouvant à tout moment faire basculer le mode de gouvernance. Les canaux de communication facilitant la coordination des participants se multiplient alors rapidement, un script étant même utilisé afin de « cacher » les commandes de la messagerie instantanée pour n'y laisser que les messages à visée sociale, visant l'organisation de l'action collective. Le temps passant, de plus en plus de joueurs finirent par se concentrer sur ce jeu de deuxième niveau, la « Démocratie » étant de plus en plus acceptée face à ses succès sur le terrain. Toutefois, une méfiance constante s'y perpétua, beaucoup y voyant une perversion de l'objectif initial de Twitch Play Pokemon⁷.

Cet événement, qui dura seulement deux semaines, représente un microcosme des phénomènes cognitifs qui nous intéressent au sein de ce mémoire. La participation massive d'individus est une lame à double tranchant dans une communauté : elle permet l'accomplissement de tâches autrement impensables, mais sans une cohérence minimale, la majorité des efforts risque de disparaître dans la confusion sans laisser de traces. En effet, que l'encyclopédie offre à tous de s'engager sans contrainte a priori dans la forme de leurs contributions n'est pas une finalité pour Wikipédia. C'est au contraire la source et la possibilité permettant l'organisation d'une communauté, renforcée au fil des événements et des conflits. C'est ainsi que nous nous intéresserons à une forme particulière de contraintes que la communauté s'est dotée, soit l'utilisation massive de bots.

Les agents logiciels, appelés bots, sont généralement interprétés comme étant des programmes dont les tâches sont déterminées par algorithmes, incluant l'édition et la maintenance d'articles, et sont aussi décrits comme des outils ou protocoles managériaux touchant l'infrastructure d'un projet (Geiger et Ribes, 2010 ; Müller-Birn et collab., 2013 ; Niederer et van Dijck, 2010). À ce titre, l'utilisation des agents logiciels n'est pas une nouvelle donne, les utilisateurs d'IRC étant familiers avec cette pratique depuis de nombreuses années (IRC est le logiciel « éponyme » des messageries instantanées) (Golbeck et Mutton, 2004 ; Canavan, 2005 ; Geiger, 2014). Notons d'ailleurs que l'interface de Twitch Play Pokemon est permise

6. <http://gamerant.com/twitch-plays-pokemon-beats-red-future-streams/>, consulté le 25 février 2014.

7. Notons à ce sujet qu'à plusieurs reprises, des succès hors du commun furent réalisés en mode « anarchie », telle l'atteinte d'objectifs que le jeu ne laissait pourtant pas de deuxième chance à réaliser. <https://www.youtube.com/watch?v=ZI03x4NIB9E>, consulté le 15 septembre 2014.

justement grâce à l'utilisation d'un bot IRC⁸. La particularité de Wikipédia est d'élever cette interaction à une intensité et une ampleur qui ne se constate nulle part ailleurs dans une communauté collaborative en ligne. Mais au contraire de services en ligne dont la gestion reste fortement hermétique, tels YouTube, Twitter ou la majorité des sites d'information, Wikipédia a laissé la gestion interne du projet à sa communauté d'utilisateurs. Ceci permet aux participants de développer des normes et lignes de conduite tout en valorisant certaines formes de participation, en modifiant le terrain même qu'est Wikipédia. À terme, c'est toute une hiérarchie sociale de droits différenciés que se dote la communauté de l'encyclopédie. Face à cette situation, ce mémoire veut donc répondre à la question suivante : *comment les individus perçoivent l'action des agents logiciels sur leur propre travail et, dans un second temps, comment le travail des agents logiciels s'intègre-t-il dans le fonctionnement et la structuration de la communauté ainsi que dans le contenu de Wikipédia ?*

En premier lieu, ce mémoire aborde le renforcement progressif de la communauté de Wikipédia en observant l'importance que représentent les encyclopédies dans nos sociétés, sur les motivations menant des individus à participer à leur élaboration, sur les conflits inévitables qui en émergent ainsi que la réponse sociale qui en est donnée. En deuxième lieu, nous aborderons une des réponses particulièrement importantes de la communauté face à la montée drastique de la participation, soit celle de l'utilisation massive de agents logiciels. C'est donc l'interaction utilisateur-bot suivit de la réalisation du portrait de l'utilisation actuelle des agents logiciels dans la version anglaise de l'encyclopédie qui nous intéressera au début de cette seconde section. Ceci nous équipera des outils pertinents nous permettant d'observer la construction des normes sociales encadrant l'utilisation des agents logiciels, avant d'exposer la signification structurelle qu'ils représentent.

Munis des éléments conceptuels nécessaires afin de mener à bien cette recherche, nous avons analysé le contenu des discussions que les éditeurs et participants ont à propos des agents logiciels. À cette fin, nous avons analysé dans un premier temps les pages de discussions d'un échantillon de 50 agents logiciels, qui représente 3976 différents commentaires répartis en 2,353 discussions. Dans un deuxième temps, nous avons analysé l'ensemble des discussions ayant eu lieu en 2012 sur le panneau d'affichage de propriétaires des agents logiciels de Wikipédia, soit 79 discussions représentant 744 différentes interventions.

Publiés dans *Computer in Human Behavior* (Annexe .3), nos résultats montrent une relation aux agents logiciels différenciés selon le statut de l'éditeur au sein de l'encyclopédie. En effet, pour les utilisateurs occasionnels, les agents logiciels risquent bien souvent de constituer l'autorité régulant les actions qu'ils peuvent entreprendre. Ceci occasionne la majorité des conflits que nous avons observés. Dans le cas de la communauté interne, les agents logiciels

8. <http://www.gamespot.com/articles/over-60k-people-watching-twitch-play-pokemon-update/1100-6417762/>, consulté le 30 février 2014.

sont essentiellement perçus en tant qu'outils permettant d'alléger les éditeurs des tâches laborieuses et répétitives. Initié pour des tâches techniques de standardisation, l'application de normes adoptées par cette même communauté devient progressivement un des aspects les plus appréciés de l'utilisation des agents logiciels. Cette communauté interne montre principalement un fort désir d'avoir des agents logiciels fonctionnant efficacement et avec efficience. Enfin, pour la communauté restreinte des propriétaires d'agents logiciels, l'étude démontre chez ceux-ci un sentiment aigu d'être les dépositaires d'un nombre considérable de fonctions techniques de l'encyclopédie, en plus d'être les protecteurs de l'application des normes déjà évoquées. Durant la période observée, ceux-ci démontrent une forte inquiétude de rester dans le cadre des consensus établis par la communauté, n'hésitant pas à l'occasion à destituer certains des leurs.

Ce mémoire permet donc de jeter un regard approfondi sur le rôle social attribué aux agents logiciels dans l'encyclopédie, fenêtre particulièrement explicite de la formalisation de la communauté de Wikipédia. Utilisant à leur avantage le cadre sociotechnique fourni par le logiciel MédiaWiki, la communauté interne se dote progressivement de puissants outils afin de protéger sa vision de l'encyclopédie face à la multitude des contributeurs lambdas. Ceci au risque réel et souvent accepté par la communauté de réduire le caractère inclusif d'un tel projet.

Chapitre 1

Wikipédia : structuration d'une communauté et appropriation d'un cadre technique

Afin d'établir les concepts permettant de répondre à notre question de recherche ainsi qu'aux hypothèses qui lui sont associées, nous réalisons ici une revue de la littérature nous permettant de construire un cadre théorique en deux parties relativement hétérogènes.

Ce chapitre vise à discuter de l'émergence de Wikipédia en l'inscrivant dans un processus continu. Partant de l'histoire des encyclopédies jusqu'à l'avènement de leurs versions numériques contemporaines, nous tenterons de saisir la relation particulière que la communauté anglaise de Wikipédia entretient avec la construction institutionnalisée du savoir, particulièrement au sein des universités. Par la suite, allant au cœur du sujet de recherche, il sera question des processus et motivations amenant des individus à participer à ce projet collectif et de comment, au fil des conflits et de leur résolution, la communauté prend davantage conscience d'elle-même et ce à quoi elle désire tendre. Par la suite, nous exposerons comment la communauté se dote d'outils permettant de forcer les nouveaux contributeurs à suivre les buts et objectifs, qui sont progressivement formalisés dans Wikipédia en terme de normes et de règles de conduite. Finalement, nous abordons une conséquence de cette prise de conscience consistant en des problèmes de rétention des contributeurs affectant Wikipédia depuis quelques années.

Dans le deuxième chapitre, il sera particulièrement question de l'utilisation des bots au sein de cette encyclopédie. Consistant en une forme proactive d'implantation des normes adoptées par la communauté de Wikipédia, nous y observons, en premier lieu, ce qui est connu des processus cognitifs entourant l'interaction utilisateur-bot. Ensuite, nous réalisons un portrait de l'utilisation actuelle des bots dans l'encyclopédie, avant de retracer le fil des événements

ayant permis de construire progressivement un cadre normatif touchant l'utilisation et l'acceptation sociale des bots. Enfin, nous réalisons la description d'un cas particulier des bots dans Wikipédia : ceux dédiés à la lutte au vandalisme. Ceci nous permet d'observer, au travers de ce cas particulier, les divers éléments préalablement abordés.

1.1 Wikipédia, une histoire courte

* ENCYCLOPÉDIE [1755], s. f. (Philosoph.) Ce mot signifie enchaînement de connoissances ; il est composé de la préposition greque ξν, en, & des substantifs κύκλος, cercle, & παιδεία, connoissance. En effet, le but d'une Encyclopédie est de rassembler les connoissances éparses sur la surface de la terre ; d'en exposer le système général aux hommes avec qui nous vivons, & de le transmettre aux hommes qui viendront après nous ; afin que les travaux des siècles passés n'aient pas été des travaux inutiles pour les siècles qui succéderont ; que nos neveux, devenant plus instruits, deviennent en même tems plus vertueux & plus heureux, & que nous ne mourions pas sans avoir bien mérité du genre humain.

« Encyclopédie », *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers*, Diderot, 1755, p.635

1.1.1 Formation et participation aux encyclopédies numériques

L'Histoire Naturelle de Plin L'Ancien, avec ses 37 volumes, est souvent considérée comme la première encyclopédie (Loveland et Reagle, 2013). Œuvre d'un seul homme, le premier acte de cette histoire se réalise aussi tôt qu'en l'an 79. Le mot « encyclopédie » lui-même n'apparaît qu'à partir du 16^e siècle (O'Sullivan, 2011). C'est un projet continuellement en gestation qui s'affirme au 18^e siècle avec Diderot et D'Alembert. C'est à cette époque qu'apparaissent les premières encyclopédies alphabétiques. Ces oeuvres se construisent grâce à des stratégies de synthèse des connaissances suivant trois principales formes : la collection compulsive, l'accumulation stigmergique, et la production corporative (Loveland et Reagle, 2013)¹. Dans le premier cas, nous retrouvons des individus réalisant des efforts hors du commun afin d'accumuler des connaissances souvent disparates. Dans le deuxième, il s'agit principalement du processus continu d'accumulation et d'ajouts successifs (stigmergie) aux anciens textes. Ce terme signifie ici la construction progressive d'un projet par stratification, signifiant l'apport successif de différents individus pour venir bonifier les apports précédents, le tout dans un objectif final commun. Enfin, en troisième, il s'agit de la présence d'un groupe cohérent qui réalise la production. Wikipédia, tel que nous l'exposons, est loin d'être un modèle unique d'organisation ; il regroupe des caractéristiques de ces trois formes (Loveland et Reagle, 2013).

1. La description que nous en réalisons est principalement basée sur cet article.

WikiWikiWeb est un logiciel créé par Ward Cunningham en 1995, dont l'objectif fut de créer une base de données la plus simple possible (Martin, 2010). Née de l'infrastructure technique offerte par ce logiciel, c'est en 2001 que Jimmy Wales et Larry Singer fondent Wikipédia (Colón-Aguirre et Fleming-May, 2012). Dû aux spécificités du projet, ceux-ci lancent donc en 2002, MediaWiki², qui est une version adaptée pour les besoins de l'encyclopédie du logiciel WikiWikiWeb. Ce n'est pas le premier projet de ces deux individus, car ceux-ci avaient fondé Nupédia le 9 mars 2000, une encyclopédie numérique libre d'accès et écrite par des spécialistes (Chen, 2010). Le changement de posture épistémologique est profond entre ces deux projets. Pour Nupedia, la publication est réalisée à la fin de son processus d'écriture, une fois que l'article atteint un niveau professionnel, dont la validité est jugée par les pairs (Niederer et van Dijck, 2010)³. Cependant, écrire à titre de spécialiste n'est pas facile ; c'est la réputation de l'auteur qui est en jeu. Celui-ci risque donc de mettre énormément de temps sur l'article avec le risque, s'il n'en est pas satisfait, de préférer ne jamais le publier. Wikipédia prend la posture inverse, autorisant quiconque à publier un article et à en modifier le contenu. La démarche offre des résultats radicalement différents, car Wikipédia atteint 20 000 articles au moment où Nupédia publie avec difficulté son centième (M. Livingstone, 2012). Il n'est, bien entendu, aucunement question de comparer la qualité moyenne des deux projets, mais la conclusion est déjà claire pour Wales qui concentre toutes ses énergies sur ce nouveau modèle coopératif (Chen, 2010). C'est en effet le succès que Wikipédia connaît depuis ses débuts. Source d'informations incontournable, elle est désormais régulièrement classée 6^e site internet en importance mondialement⁴. Ceci grâce à la visite mensuelle de 500 millions d'utilisateurs, qui consultent les 33 millions d'articles écrits à ce jour par 22 millions de contributeurs inscrits.

L'approche stigmergique est de toute évidence le mode principal de participation de Wikipédia, l'essentiel des contributions étant ajouté par stratification aux articles existants (Loveland et Reagle, 2013). D'articles ayant une seule phrase lors de leur publication initiale, une version mature possède souvent des milliers de contributions individuelles⁵. Sur ce point, notons qu'il est classique pour une encyclopédie de reprendre la nomenclature de ses prédécesseurs ; souvent, il est même question de copie pure et simple. Les droits de propriétés intellectuelles d'avant 1709, date de l'adoption des « *acts of Anne* » en Angleterre, y sont pour beaucoup (May et Sell, 2005). Ces actes ont formalisé ce qui ressemblera, pour la première fois, aux droits d'auteurs et brevets modernes. À cette époque, l'obtention de brevets et/ou droits d'auteurs est souvent autorisée pour la simple importation locale d'un livre ou d'une technologie provenant d'un autre pays. À une époque où la diffusion n'est pas chose évidente, certains pays réalisent beaucoup d'efforts afin de s'assurer d'être au sommet de l'art (May et

2. <http://en.wikipedia.org/wiki/MediaWiki>, consulté le 3 juin 2014.

3. <http://web.archive.org/web/20010331211742/www.nupedia.com/policy.shtml>, consulté le 28 février 2014.

4. Selon le classement d'Alexa : <http://www.alexa.com/siteinfo/wikipedia.org>

5. Exemple de l'article sur le Sida et ses plus de 5000 éditions : <http://en.wikipedia.org/w/index.php?title=HIV/AIDS&action=history>, consulté le 3 juin 2014.

Sell, 2005). C'est toutefois une tradition à part entière qui se développe au sein des encyclopédies quant à la réutilisation des écrits. Dans le cas de Diderot, par exemple, une majorité des figures retrouvée dans son encyclopédie provient de la *Description des arts et métiers* (Diderot, 1755), œuvre majeure qui lui est précurseur (Loveland et Reagle, 2013). Dans les faits, les encyclopédistes sont souvent considérés avant tout comme compilateurs. Ce n'est que bien plus tard que le terme d'éditeur va s'imposer avec la coordination nécessaire d'un collectif d'individus travaillant sur les mêmes articles (Loveland et Reagle, 2013). L'apparition de travaux originaux est donc l'exception, tout comme y retrouver la signature d'un « auteur ». L'anonymat y devient donc bien souvent la norme, certaines fois avec l'objectif direct de protéger les auteurs des autorités (Loveland et Reagle, 2013). Parallèlement, l'émergence des savoirs spécialisés sera accompagnée des encyclopédies corporatives. Il devient en effet rapidement très difficile pour un individu d'être apte à couvrir complètement et correctement un sujet. Diderot, déjà, se devait d'user plus de 140 spécialistes. Notons toutefois que l'apport de spécialistes de « renommée » constitue un jeu avant tout politique, autant pour les spécialistes que pour l'encyclopédie. Puisque les travaux originaux ne sont pas autorisés, chacun espère profiter de la renommée de l'autre (Loveland et Reagle, 2013). Dans certains cas, cette contribution distinctive sera obtenue en allant chercher le savoir localisé de populations locales. La contrainte géographique étant la motivation et la limitation de cette approche (Loveland et Reagle, 2013).

Comme il sera démontré, Wikipédia ne se démarque pas par son originalité en tant qu'encyclopédie. C'est plutôt sa capacité à utiliser conjointement ces diverses tendances, tout en enlevant certaines barrières traditionnelles à la participation, qui en fait un cas particulier qui a marqué le genre encyclopédique (O'Sullivan, 2011). Comme n'importe quel projet, Wikipédia existe sous une déclaration de principes. Principalement basée sur les politiques précédemment mises en place dans Nupédia⁶, cette déclaration est élaborée lors des premiers mois de l'encyclopédie (quelques discussions entourant l'acceptation de ces normes.⁷). Celle-ci se résume en cinq principes minimaux à respecter (table 1.1)⁸. Premièrement, Wikipédia est une encyclopédie libre, ce qui est distinct des journaux, dictionnaires, plateformes de publicité ou revues dédiées aux rumeurs. Deuxièmement, elle se doit d'être écrite d'un point de vue neutre, l'information doit donc aussi y être vérifiable et l'opinion individuelle y est proscrite au sein des articles. Troisièmement, elle est gratuite et libre d'être modifiée dans son ensemble. Quatrièmement, tous doivent être respectueux et civilisés. Enfin et non le moindre, Wikipédia autorise une interprétation créative des règles, pilier nommé avec la boutade « *ignore all rules* ». Malgré les politiques et lignes directrices qui construisent l'encyclopédie, il est toujours pos-

6. <http://web.archive.org/web/20001205000200/http://www.nupedia.com/policy.shtml#III>, consulté le 28 février 2014.

7. <http://web.archive.org/web/20010416035716/www.wikipedia.com/wiki/RulesToConsider>, consulté le 25 février 2014.

8. http://en.wikipedia.org/wiki/Wikipedia:Five_pillars, consulté le 25 février 2014.

Tableau 1.1 – Les cinq piliers de Wikipédia.

Est une encyclopédie	D'écriture neutre	Libre & gratuite	Respect & Civilité	Interprétation créative des règles
----------------------	-------------------	------------------	--------------------	------------------------------------

sible pour un participant d'y déroger si cela nuit à l'accomplissement de tâches particulières. Cela signifie que les principes y sont considérés prédominants sur la formulation des normes.

Notons enfin qu'il n'est pas surprenant de voir Wikipédia se fondre au modèle des logiciels libres (*Open Source*), car l'émergence des encyclopédies est considérée à la source de cette philosophie (Loveland et Reagle, 2013). L'idée principale derrière les logiciels libres repose sur le postulat que la multiplicité des regards permet de mieux d'identifier les erreurs existantes qu'un projet fermé. Ce principe est nommé *Loi de Linus* par Eric S. Raymond, en l'honneur de Linus Torvalds, fondateur de Linux⁹. De plus, « *En permettant aux individus de s'inclurent au sein de communautés dont ils acceptent d'être liés aux règles, ces espaces promettent un idéal libéral de gouvernance consensuelle et d'autonomie individuelle* » (Suzor et Woodford, 2013)¹⁰. Combiné à son statut d'organisme à but non lucratif, ces méthodes allègent Wikipédia de l'essentiel des problèmes liés à l'application des droits de propriété intellectuelle (Cassiman et collab., 2007). Cette section nous permet donc de comprendre différents éléments centraux au concept d'encyclopédie. Ces mêmes éléments ne sont toutefois pas acceptés de la même façon par tous, comme le témoigne l'accueil que reçoit Wikipédia, à ses débuts, des milieux académiques.

1.1.2 À contre-courant du savoir institutionnalisé

Au travers des technologies d'informations et de communications, c'est toute notre façon de rejoindre les individus et le savoir qui se transforment. Il est clair que la recherche d'informations académiques et scientifiques s'oriente progressivement autour des outils en ligne (Brossard, 2013). En abaissant considérablement les coûts de participation à des projets collectifs (Faraj et collab., 2011), Wikipédia représente un symbole de succès de ce phénomène contemporain. Malgré tout, le projet est souvent abordé avec grande méfiance par le monde universitaire, bien que ce phénomène soit considérablement en baisse (Konieczny, 2014). Des universités ont même explicitement interdit son utilisation en tant que référence (Waters, 2007).

Suite à la réalisation d'une étude quantitative sur l'utilisation de Wikipédia à l'Université Liverpool Hole, (Knight et Pryke, 2012) exposent cinq arguments qui pourraient expliquer cette prise de position (table 1.2). D'un, bien que l'encyclopédie ne soit pas antiacadémique,

9. <http://www.catb.org/~esr/writings/cathedral-bazaar/cathedral-bazaar/>, consulté le 28 février 2014.

10. Traduction libre.

Tableau 1.2 – **Tableau récapitulatif en cinq points quant aux clivages entre les mondes académique et encyclopédique.*** Particulièrement dans les arts et les sciences sociales

Académie	Wikipédia
Titre de maître	Antiélitiste
Droit d’auteur et publication	Culture libre
Cosigné	Anonymat
Création individuelle*	Création collective
Validation par les pairs	Validation par les membres

il y a une culture antiélitiste certaine. Celle-ci s’observe souvent lorsque des spécialistes d’un domaine s’attendent à la reconnaissance de leur expertise lors d’un débat touchant un sujet, reconnaissance qui n’est pas accordée dans bien des cas. De deux, la culture libre de Wikipédia va à l’encontre de la culture académique qui valorise le droit de l’auteur et de la publication. De trois, l’anonymat des articles Wikipédia est à contre-courant de la publication d’articles cosignés. De quatre, la collaboration créative met au défi la norme de création individuelle, souvent retrouvée dans les sciences sociales et les arts. Finalement, Wikipédia ne se base pas sur une validation par les pairs afin d’assurer la qualité des articles. Ces objections ne sont pas mineures, étant à l’origine du départ du fondateur et premier éditeur en chef de Wikipédia, Larry Sanger¹¹. Celui-ci fonde dans la foulée Citizendium¹². Ce nouveau projet, bien qu’acceptant les apports de tous, y refuse l’anonymat, accorde le statut d’expert à des spécialistes et est supervisé par un comité d’éditeurs (Soto, 2009). Sur ce dédain académique, notons que les encyclopédies n’ont jamais eu une grande importance dans les universités, Wikipédia n’étant donc pas l’exception ici (Waters, 2007).

Dans les faits, bien que Wikipédia possède différents problèmes, elle se compare positivement aux autres encyclopédies, telle l’*Encyclopédia Britannica*. Par exemple, (Giles, 2005) a demandé à des spécialistes de différents domaines de comparer 42 articles à caractère scientifique entre les deux encyclopédies. L’étude identifie finalement 8 erreurs majeures dans les deux cas, ainsi que 162 erreurs mineures pour Wikipédia contre 123 dans la *Britannica*. Dans une autre étude, 55 spécialistes questionnés sur la crédibilité des articles ont reconnu la relativement haute crédibilité de Wikipédia, bien que la qualité reste fort variable selon les domaines (Chesney, 2006 ; Mesgari et collab., 2014). Cette mesure de la qualité se base théoriquement sur quatre éléments : la précision des connaissances, leurs achèvements¹³, leurs actualités ainsi que leurs compréhensibilités (Fallis, 2008)¹⁴. Dans le cas de la compréhensibilité et de l’achèvement des articles, Wikipédia n’est pas en mesure de démontrer la présence systéma-

11. <http://features.slashdot.org/story/05/04/18/164213/the-early-history-of-nupedia-and-wikipedia-a-memoir>, consulté le 28 février 2014.

12. <http://en.citizendium.org/>, consulté le 2 septembre 2014.

13. Le fait qu’elles couvrent complètement les différents éléments d’intérêt d’un sujet.

14. Le fait qu’elles soient suffisamment bien formulées pour que le lecteur soit apte à les assimiler.

tique de ces éléments (Mesgari et collab., 2014). Toutefois, la qualité des articles, démontrée précédemment, y est aussi mesurée stable dans le temps (Ehmann et collab., 2008)¹⁵).

Cela peut paraître étonnant puisque 12% des discussions d'arrière-pages (onglet de discussion présent pour chaque article) est constitué de conflits (Schneider et collab., 2010). Toutefois, la recension des guerres d'éditions (*edits-wars*) touchant le contenu démontre que celles-ci ne représentent qu'environ 1 % des articles (Yasseri et collab., 2012). L'ampleur des conflits internes de ces articles empêche d'y retrouver un contenu de qualité, bien que cette proportion soit bien en dessous des estimations précédentes (Yasseri et collab., 2012). Logiquement, il est normal de retrouver le même type de conflit tant à l'extérieur de l'encyclopédie qu'à l'intérieur. Il n'est donc pas étonnant de voir que ces conflits touchent bien souvent la religion et la politique. Par exemple, *Israël*, *Adolf Hitler*, *l'Holocauste* et *Dieu* se retrouvent dans les 100 articles les plus conflictuels des 10 langages analysés de cet article (Yasseri et collab., 2012).

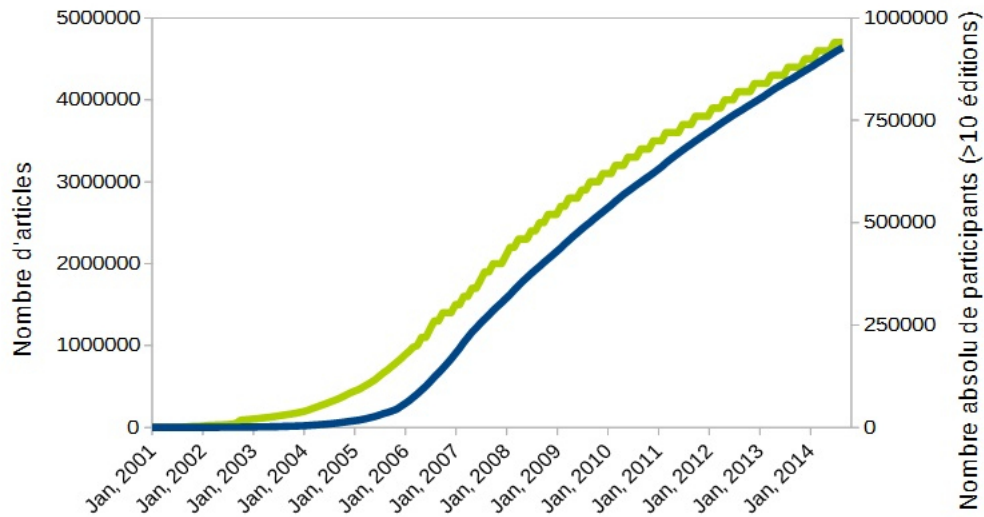
L'apport concret que Wikipédia réalise quand il est question de transmission des connaissances passe par sa prépondérance au sein des résultats des moteurs de recherche, augmentant la qualité générale de l'information exposée aux utilisateurs lambda d'Internet (Fallis, 2008). Plus largement, un avantage de ces nouveaux médias est d'exposer les individus à des informations qu'ils n'auraient probablement pas rencontrées autrement (Brossard, 2013). Ceci est d'autant plus important qu'on observe qu'à la Liverpool Hope University, 75 % des étudiants et universitaires parcourent Wikipédia pour commencer leurs travaux scolaires et 70 % l'utilisent pour avoir des informations complémentaires en classe (Knight et Pryke, 2012). Ces constatations sont aussi retrouvées chez les étudiants en médecine (Judd et Kennedy, 2011). Ceci ne s'arrête toutefois pas aux étudiants : les patients se documentent et s'informent de plus en plus sur leurs conditions médicales ou sur différents traitements à partir de ces nouveaux médias (Laurent et Vickers, 2009).

1.1.3 Des particularités des versions spécifiques de Wikipédia

La prépondérance de l'encyclopédie en ligne Wikipédia en fait un modèle de choix pour notre étude. Aux fins de cette recherche, nous nous sommes concentrés sur une version spécifique de Wikipédia, c.-à-d. la version anglaise. En effet, tous les projets collaboratifs ne connaissent pas le même succès. Il est observé que ceux-ci connaissent souvent une première phase de lente croissance, suivie d'une seconde à croissance rapide si le projet trouve un écho auprès d'un public significatif (Alluvatti et collab., 2011). Cette affirmation fait écho à la loi de Metcalfe, affirmant que les possibilités de collaborations et d'échanges augmentent au carré du nombre de participants (Metcalfe, 1995). Lorsque ce momentum n'est pas atteint, le projet

15. Certains cas existent démontrant toutefois la légitimation d'une information pourtant sans fondement : <http://www.newyorker.com/online/blogs/elements/2014/05/how-a-raccoon-became-an-aardvark.html>, consulté le 28 février 2014.

FIGURE 1.1 – **Historique de la croissance en articles et en participants de la version anglaise de Wikipédia.** Progression du nombre d'articles (vert pâle et l'axe Y de gauche) et pour le nombre de contributeurs ayant réalisés au moins 10 éditions depuis leur inscription (bleu foncé et l'axe Y de droite).



est souvent voué à mourir à petit feu, devenant souvent la base pour un projet succédant (tel Nupédia dans le cas de Wikipédia). Dans le cas de Wikipédia, plusieurs projets connaissent de toute évidence une forte activité. Malgré tout, de par le statut particulier de la langue anglaise sur Internet, cette version de l'encyclopédie forte de 4,46 millions d'articles, connaît un succès et une popularité peu comparable aux autres versions linguistiques (s'en suit la version néerlandaise (1,76 M), allemande (1,69 M) et suédoise (1,61 M))¹⁶.

Cette dominance de la version anglaise n'est pas étonnante au vu de la distribution des données générales de l'encyclopédie. Elle connaît en effet une croissance exponentielle (*scale-free*) de plusieurs de ses statistiques, autant dans la grosseur de la base de données, le nombre total de mots, de liens, du nombre d'articles, du nombre d'éditeurs actifs ainsi que les très actifs (Voss, 2005). La phase exponentielle de la version anglaise s'observe aussitôt qu'en 2004 (Halfaker et collab., 2012), avant de fléchir vers une croissance linéaire aux environs de 2007, autant dans le nombre d'articles que dans le nombre d'éditeurs (Suh et collab., 2009). Il s'agit désormais de 4,6 millions d'articles pour près de 1 million de contributeurs ayant réalisé au moins 10 éditions depuis 2001 (voir la figure 1.1). De nombreux paramètres reflétant l'activité de Wikipédia s'avèrent être distribués selon des lois de puissance (Ciampaglia, 2011). Cette donnée s'observe diachroniquement au travers du nombre d'éditions des utilisateurs dans un article tout comme les communications qu'ils ont entre eux (Muchnik et collab., 2013), le nombre d'articles édités par un utilisateur (Nazir et Takeda, 2008) ainsi que la popularité des

16. http://en.wikipedia.org/wiki/List_of_Wikipedias, consulté le 27 février 2014.

articles¹⁷. Encore, cela s’observe pour le nombre d’auteurs par article, le nombre d’hyperliens entrants et sortants, le nombre de liens brisés par article (Voss, 2005) ainsi que pour la récurrence des guerres d’éditions et le temps entre deux éditions dans un article hautement contesté (Yasseri et collab., 2012), qui suivent la même distribution.

Malgré la forte similitude quant à ces progressions statistiques, il est nécessaire de différencier les versions de l’encyclopédie. Étant gérée par une communauté à chaque fois distincte, la structure de pouvoir et de normes n’a pas évolué selon les mêmes dynamiques. Puisque chacune des politiques et règles de conduite est adoptée suite à l’atteinte d’un consensus, les communautés indépendantes approfondissent le processus de différenciation à chacun des débats qui prend place. De plus, étant donné l’objet principal de notre étude, soit l’utilisation des bots dans Wikipédia, le choix de la version anglaise s’impose. Il y est en effet constaté de fortes divergences dans leur utilisation entre les différents projets. Par exemple, les bots antivandalismes, si importants et populaires dans le cadre de Wikipédia anglais, sont pourtant interdits dans la version allemande de l’encyclopédie (Müller-Birn et collab., 2013). Sans trop s’avancer pour l’instant, notons aussi que l’importance des bots est fortement corrélée au statut de la langue (Niederer et van Dijck, 2010). La proportion des éditions appartenant aux bots, toutes versions de Wikipédia confondues, est de 21.5 % en 2008, mais grimpe à 39 % si la version anglaise n’est pas considérée. Cette importance relative des bots, en fonction du nombre d’éditions, devient prédominante lorsqu’il est question de langues en danger (Niederer et van Dijck, 2010). Pour ceux-ci, les bots réalisent beaucoup de travail de redirection par hyperliens des articles afin de s’assurer de briser l’isolement vécu par les projets dont le langage est peu répandu. La diversité des tâches réalisées par les bots sera d’autant plus grande au sein des plus gros projets. Enfin, au même titre que le classement par langue de Wikipédia, les participants d’une version particulière vont avoir eux-mêmes tendance à se regrouper en sous-communautés d’affinités, telle celle se dédiant à la gestion des images dans l’encyclopédie (Viégas, 2007), ou le projet Médecine¹⁸. Selon les éléments précédemment mentionnés, il est donc toujours question de la version anglaise de Wikipédia dans ce mémoire, sauf s’il est explicitement mentionné le contraire.

1.2 Wikipédia : entre individus et communauté

Au travers de la plateforme qu’est Wikipédia, un grand nombre d’individus cherchent à acquérir et à échanger des connaissances. Ceci dit, loin d’être libérés de leurs soucis et habitudes de vie par l’utilisation d’Internet, les individus ont tendance à reproduire dans les communautés virtuelles beaucoup des comportements, normes et règles qu’ils ont en société (IRL selon le jargon ; *in real life*) (Pastinelli, 2007 ; Guitton, 2012). Une particularité des échanges

17. http://en.wikipedia.org/wiki/Wikipedia:Does_Wikipedia_traffic_obey_Zipf's_law?, consulté le 27 février 2014.

18. https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine, consulté le 3 septembre 2014.

numériques est la permanence des écrits. Les participants ont donc tendance, comme partout ailleurs, à donner une cohérence d'ensemble aux interventions qu'ils réalisent, qui se fondent en une nouvelle culture communément partagée à différents niveaux par une même communauté (Pastinelli, 2009). Cette section est l'occasion d'observer les motivations d'un individu à participer à Wikipédia, soit en contribuant à quelques occasions, soit en joignant de plain-pied la communauté sous-jacente. Une communauté qui, suite à de nombreux questionnements et conflits, finit par s'imposer à elle-même de nombreuses règles de conduite. Enfin, nous observons comment cette organisation interne change rétroactivement l'accueil qui est réservé aux nouveaux contributeurs.

1.2.1 Structuration de collaborations individuelles

La porte d'entrée évidente de la participation à Wikipédia est par la lecture de ses articles, généralement évoquée en terme de participation périphérique légitime (Antin et Cheshire, 2010). Avant de se sentir apte à agir ou à participer, un individu cherche toujours à avoir un minimum d'information sur une situation donnée. Par la nature ouverte des articles ainsi que par l'accessibilité de leurs archives et des discussions d'arrière-page, il peut être relativement aisé d'atteindre cette masse minimale d'informations nécessaires avant de se sentir confiant pour y contribuer (Shao, 2009). Toutefois, un sondage réalisé auprès des lecteurs de Wikipédia démontre que les connaissances qu'ils possèdent du fonctionnement de l'encyclopédie décroissent proportionnellement avec la teneur administrative et d'arrière-plan des règles de gestion et de participation (Antin et Cheshire, 2010). Confirmant à prime abord que la lecture est bel et bien la porte d'entrée de l'encyclopédie, cette constatation expose aussi que certains mécanismes de gestion interne de l'encyclopédie deviennent rapidement obscurs pour la majorité des lecteurs. Il faut dire que Wikipédia n'explique pas cette nouvelle réalité normative dans sa présentation au public, s'affichant toujours en tant que portail ouvert à toutes les contributions.

Plusieurs éléments peuvent toutefois être identifiés afin d'expliquer ce qui porte les individus à s'impliquer plus fortement dans Wikipédia. Dans les théories organisationnelles classiques, il est observé que c'est grâce à des motivations intrinsèques et extrinsèques que les individus seront portés à partager les connaissances qu'ils possèdent (Bock et collab., 2005). Dans le cas des motivations extrinsèques, souvent invoquées en terme de capital social dans le cas de Wikipédia, plusieurs études touchant d'autres projets collaboratifs affirment toutefois que cette forme de motivation est notablement mitigées dans leurs résultats (Hsu et collab., 2007 ; Chang et Chuang, 2011 ; Chiu et collab., 2006). Pour les participants, il est loin d'être évident qu'ils recevront une récompense pour le travail qu'ils effectuent. En effet, même dans le cas d'une entreprise et des avantages salariaux qu'elle offre par exemple, cette forme de motivation à la participation peut avoir un effet négatif sur la motivation à partager un savoir (Yang et Lai, 2010). Le cas des salaires liés à la performance y est questionné, puisque cela encourage

un employé à conserver pour lui les connaissances qu'il possède, ce qui constitue son avantage compétitif. La deuxième forme souvent évoquée touche les motivations intrinsèques. Cette motivation se retrouve chez les individus obtenant leur satisfaction de l'acte même qu'ils réalisent (Yang et Lai, 2010). Ces deux propositions ne semblent pas satisfaisantes afin d'expliquer la participation à Wikipédia (Yang et Lai, 2010). D'un côté, la participation ne se traduit que très rarement en gain de réputation (Reagle, 2010a). De plus, l'incitation monétaire ne peut être considérée vu la très grande désapprobation constatée envers quiconque pouvant toucher un salaire pour sa participation dans l'encyclopédie (il est connu que certaines compagnies engagent des éditeurs afin de bien *entretenir* certains articles les touchant¹⁹). De l'autre, la survie d'une édition est peu assurée, voire très basse chez les nouveaux utilisateurs (Halfaker et collab., 2012). Le plaisir de partager ses connaissances peut enclencher une participation, mais cela risque bien de ne pas être suffisant afin de la maintenir. En effet, il est généralement reconnu que cette satisfaction intrinsèque est obtenue au travers du lien social qui est créé, lien social qui est sporadique et difficile à instaurer quand il est question du travail sur les articles (Yang et Lai, 2010). Les formes classiques de motivations intrinsèques et extrinsèques ne semblent donc pas être particulièrement efficaces pour décrire les incitatifs à la participation dans une communauté collaborative en ligne.

Toutefois, Yang et Lai (2010) considèrent deux types supplémentaires de motivations à la participation dans le cadre de Wikipédia. Nous retrouvons premièrement la motivation externe conçue en soi-même (*external self-concept motivation*), qui consiste en l'adoption d'un comportement que l'individu croit valorisé par un groupe de référence. Puis, nous retrouvons aussi la motivation interne conceptualisée en soi-même (*internal self-concept motivation*) consistant en l'adoption d'un comportement jugé concordant avec nos propres standards internes. Ces deux types de motivations sont aussi approximativement retrouvés chez Kuznetsov2006, Wagner2007. C'est cette dernière forme qui est retrouvée avec la plus grande influence sur la participation individuelle dans l'étude de Yang et Lai (2010). Il semble, en effet, que les individus obtiennent beaucoup de satisfaction et de confiance lorsqu'ils se voient confirmer la justesse de leurs connaissances et expertise sur un sujet. Ce n'est donc pas sans raison qu'il existe un sentiment d'appropriation des articles par leurs principaux auteurs, même si cela interfère avec les normes de l'encyclopédie (Halfaker et collab., 2009 ; Thom-santelli et collab., 2009). En effet, en interdisant les travaux originaux, malgré cette propension à s'approprier les articles, c'est la notion même de propriété qui est formellement interdite dans Wikipédia. Ceci interdit autant les travaux originaux que de s'affirmer responsable d'un article.

Cependant, l'article de Yang et Lai (2010) n'explicite pas l'importance quant au statut interne et au niveau général de participation chez les individus au sein de son échantillon. Bien que le contributeur lambda risque de participer avec l'objectif que l'encyclopédie atteigne le

19. <http://arstechnica.com/tech-policy/2014/01/wikimedia-foundation-employee-ousted-over-paid-editing/>, consulté le 27 février 2014.

standard de qualité recherché par celui-ci, un utilisateur plus régulier risque de correspondre davantage avec la motivation extrinsèque conçue en soi-même. Là-dessus, Restivo et van de Rijt (2012) démontrent de façon expérimentale que l'utilisation de « *barnstars* »²⁰, récompenses symboliques octroyées par les pairs, augmente la participation chez les récipiendaires. D'ailleurs, une autre étude conclut que « *ces résultats suggèrent que la participation des utilisateurs dans les systèmes de production collaborative peut être conçue principalement comme un processus d'influence sociale médiatisée* » (Ciampaglia, 2011)²¹. D'autres études tirent des conclusions semblables (Kuznetsov, 2006 ; Wagner et Prasarnphanich, 2007). Ce n'est donc pas pour rien que la participation à des sous-communautés de l'encyclopédie, tels les *Wiki*projects, semble être un facteur déterminant dans la rétention des contributeurs (Forte et collab., 2012). La seule étude que nous avons identifiée allant à l'encontre de cette approche possède comme prémisse que contrairement aux organisations réelles, les communautés virtuelles ne possèdent pas de règles formelles ni de structure de pouvoir, ce qui s'avère fortement discutable (Hsu et collab., 2007). En effet, tel qu'il sera exposé en détail au cours des deux prochaines sections, ce n'est évidemment pas le cas pour Wikipédia qui forme progressivement un complexe système social au cours de son développement.

Une des manières de conceptualiser les motivations externes conçues en soi-même qui se révèle pertinente pour notre étude est celle de Benedict Anderson dont les travaux ont largement porté sur les imaginaires sociaux (Anderson, 2006). Celui-ci décrit l'émergence du sentiment d'appartenance dans une collectivité par l'établissement d'un imaginaire commun. Un individu attribuera le partage de cet imaginaire, que cela soit en termes de croyances ou d'attitudes, à tous ceux qu'il juge appartenant à ce groupe, et ce, même si l'individu ne rencontrera jamais l'immense majorité des autres membres le composant. Par exemple, bien que les pages de discussions soient explicitement réservées à l'amélioration des articles, il est possible d'observer des activités de commémoration s'installer sur celles-ci, lorsqu'elles touchent des événements traumatisants ou catastrophiques (Ferron et Massa, 2013). De plus, cette appartenance, malgré les jeux de pouvoir réels entre les individus, met en place une communauté qui est foncièrement perçue horizontalement par les participants (Anderson, 2006). Par définition, cette conception se limite donc dans un espace particulier ainsi que pour un nombre de participants limité. Cette définition devient particulièrement explicite si elle est mise en face d'un second groupe qui s'oppose au premier, ce qui expose clairement la frontière construite (Anderson, 2006). Dans le cas de Wikipédia, nous pouvons constater le discours créé par la communauté interne mettant vis-à-vis d'elle l'anarchie des utilisateurs lambdas (Halfaker et collab., 2012). Chez les contributeurs ayant beaucoup de motivations à participer à la communauté, il est possible d'observer une certaine crainte devant la posture ultra

20. « *est coutume de récompenser un contributeur de Wikipédia pour son dur travail et sa diligence en lui accordant une barnstar.* » Cette récompense symbolique, sous forme d'images laissées en commentaire sur la page de discussion d'un éditeur, peut être décernée par n'importe quel participant pour n'importe quel autre participant. Source : <http://en.wikipedia.org/wiki/Wikipedia:Barnstars>.

21. Traduction libre.

inclusive de Wikipédia. Enfin, de par les sondages, recensement, registre, etc., cette communauté imaginaire se matérialise progressivement en inscrivant dans l'espace les démarcations fournies par sa grille de lecture (Anderson, 2006).

Approfondissant les conceptions entourant la relation des participants d'avec les autres membres de Wikipédia, quatre types de réactions²² entre ceux-ci ont été observés : positive, négative, directive et sociale (Zhu et collab., 2013). Alors que les commentaires négatifs et directifs augmentent les efforts individuels sur une tâche particulière (effort focal), les commentaires positifs et sociaux augmentent la productivité générale. Chez les nouveaux utilisateurs, l'impact est non-négligeable, particulièrement lorsque l'on considère l'annulation d'une édition comme une forme radicale des commentaires négatifs (Restivo et van de Rijt, 2012 ; Halfaker et collab., 2011 ; Geiger et collab., 2012). En effet, en tenant compte de l'importance des motivations intrinsèques conçues en soi-même (Yang et Lai, 2010), les commentaires négatifs sont souvent considérés comme une remise en cause des connaissances et de l'expertise chez la personne visée (Zhu et collab., 2012). Les participants développent alors souvent des tactiques afin de réduire leur exposition à de tels commentaires, notamment par l'arrêt de la participation. Zhu et collab. (2012) indiquent toutefois que ces résultats sont peu valides quand il est question des utilisateurs expérimentés. Considérant maintenant l'importance des motivations externes conçues en soi-même chez les utilisateurs expérimentés, nous émettons l'hypothèse qu'il s'agit d'un impact mineur sur leurs éditions futures, ceux-ci étant plus aptes à accepter la critique des autres membres.

À ce point, la transformation des éditeurs occasionnels en éditeurs actifs semble suivre plusieurs tendances différentes. Il est constaté que, dès les premières éditions, certains éléments distinguent les futurs éditeurs actifs des autres (Panciera et collab., 2009). Dès les premiers jours, ceux-ci réalisent beaucoup plus d'éditions que ceux dont la participation ne sera pas maintenue dans le temps. Notons toutefois que toutes les catégories d'utilisateurs commencent leurs implications par une rafale d'éditions qu'ils ne réussiront pas à maintenir dans le temps (en ce sens qu'ils diminueront la fréquence à laquelle ils éditeront) (Panciera et collab., 2009). Présentant ici encore des lois de puissance, il est constaté que 10 % de la communauté réalisent 90 % des contributions totales à l'encyclopédie, dénotant d'autant plus la présence d'un groupe particulier au cœur des participants à Wikipédia (Ortega et collab., 2008). La participation active se traduit souvent par la spécialisation du participant, avec notamment les *experts substantifs*, les *éditeurs techniques*, les *éditeurs de contre-vandalisme* ainsi que les *éditeurs sociaux de réseautages* (Welser et collab., 2011). Selon ces rôles non exclusifs, les efforts de chacun sont répartis en différentes sphères de l'encyclopédie. Comme il sera possible pour le lecteur de constater à la prochaine section, les utilisateurs actifs font davantage référence aux normes de la communauté afin de légitimer leurs contributions face

22. L'article utilise le terme de « *Feedback* ».

aux nouveaux utilisateurs (Panciera et collab., 2009), indiquant l'importance première de l'appartenance à la communauté pour ce groupe d'utilisateurs.

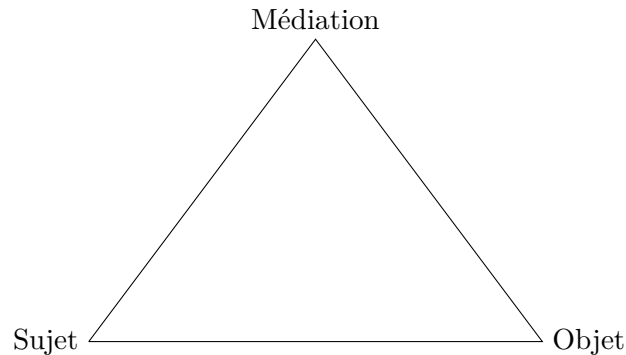
1.2.2 Contradiction interne et émergence de normes communes

Parmi les 5 piliers de Wikipédia, nul n'a été autant utilisé et maltraité que celui concernant l'interprétation créative des règles. Nommée non sans raison « *Ignore all rules* »²³, cette norme est en place afin de rappeler à tous qu'aucune règle ne devrait venir limiter le développement du projet, les participants étant donc invités à ignorer certaines règles lorsque la nécessité se présente. D'ailleurs, il est noté à quel point ce règlement a servi de mécanisme permettant de relâcher les tensions accumulées par la communauté (Joyce et collab., 2012). Nous pouvons comprendre que l'objectif de ce pilier est d'instaurer un terrain spécifiquement neutre, afin de laisser la communauté gérer ses décisions et orientations selon la nécessité. Ce scepticisme à l'égard des règles se traduit justement par la nécessité de l'atteinte d'un consensus avant de procéder à l'implantation d'une nouvelle règle, politique ou guide de conduite. Toutefois, la hausse drastique de la participation est venue modifier l'équilibre interne de l'encyclopédie. En effet, une communauté collaborative est tirée entre deux tendances difficiles à accorder, soit entre la liberté d'action et l'organisation managériale (Yang et collab., 2013). L'article démontre que la taille d'un groupe influence fortement l'efficacité du travail réalisé. Chez un groupe limité en participation, beaucoup de liberté et d'indépendance peuvent être conservées chez les individus dans la forme de leurs contributions. L'augmentation en taille force, dans le but de conserver une certaine efficacité, à délaissier cette liberté au profit de l'application d'une structure managériale pour la répartition des tâches (Yang et collab., 2013).

L'auteur Michel Freitag, au travers de son œuvre majeure *Dialectique et Société* (Freitag, 2011)(réédition de son livre publié en 1986), offre un éclairage particulièrement pertinent afin de comprendre les processus d'instauration des normes au sein d'une communauté. Tel que présenté dans la figure 1.2 à la page 19, ce processus s'expose en trois pôles, soit : le sujet (ici, les contributeurs), un objet (leur travail sur les articles), ainsi qu'une médiation, qui est le détour par lequel le sujet matérialise son action sur l'objet (soit les représentations de Wikipédia que ceux-ci possèdent). Dans ses premiers moments, une communauté fonctionne généralement de manière consensuelle (à ne pas confondre avec les décisions « consensuelles » prises par la communauté de Wikipédia lors d'un débat). Ceci est possible par l'existence d'un discours communément partagé quant à l'origine de la communauté. En tant que communauté libérale, les participants de Wikipédia sont généralement au fait des cinq piliers, déclaration d'intentions faite par les fondateurs de l'encyclopédie. Les premiers conflits se règlent alors essentiellement par un rappel de ces mêmes fondateurs des principes de base, souvent par leur réinterprétation face à des conditions changeantes. Leur autorité provient donc de leur titre

23. http://en.wikipedia.org/wiki/Wikipedia:Ignore_all_rules

FIGURE 1.2 – Triangle dialectique.

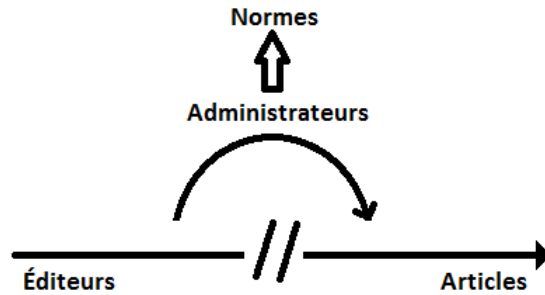


d'interprètes privilégiés de la nature de Wikipédia (Stvilia et collab., 2008)²⁴.

Cependant, cela ne s'avère pas suffisant, car la proportion des occasions menant à la mésentente augmentent parallèlement à l'ampleur de la participation (Kittur et collab., 2007b). Trois types de conflits principaux peuvent être identifiés dans les communautés collaboratives : ceux reliés aux tâches (différents points de vue sur une tâche) ; ceux de type affectives (causés par les relations interpersonnelles) ; et ceux reliés aux processus (sur la forme que doit prendre la tâche pour l'accomplir) (Arazy et collab., 2013). À l'exception des conflits reliés aux tâches, les conflits causent généralement une baisse de la productivité. La méthode la plus efficace proposée jusqu'à maintenant pour exposer les conflits dans Wikipédia utilise comme postulat qu'ils sont essentiellement composés de l'affrontement de points de vue entre deux individus particuliers. Ces conflits sont donc mesurés selon la propension des individus à inter-annuler leurs éditions respectives (Gerardo et collab., 2012 ; Yasseri et collab., 2012, 2014)). Il s'agit donc de l'affrontement de deux visions opposées dont le lieu de friction est un article particulier. Par exemple, les conflits sur Wikipédia correspondent souvent à ce qui est retrouvé sur le plan géopolitique (Apic et collab., 2011) tout en étant modulé selon l'origine culturo-langagière des participants (Yasseri et collab., 2014). Dans ce cadre, un débat qui dégénère se transpose bien souvent en conflit interpersonnel (Arazy et collab., 2013). Les conflits se produisant dans l'espace de Wikipédia sont d'ailleurs reconnus pour être particulièrement agressifs (Reagle, 2010b), ce qui participe activement à réduire la participation des femmes dans l'encyclopédie (Päivärinta et Lanamäki, 2010 ; Laniado et collab., 2012 ; Ridge, 2013). La réversion « *revert* » est souvent utilisée comme une arme à part entière au sein de ces débats (Yasseri et collab., 2012). De toute évidence, certains débats sont donc causés par des événements externes à l'encyclopédie, causant des conflits temporaires, voire sporadiques (telle une guerre, un conflit social, etc). En fait, la reproduction des conflits profonds présents dans nos sociétés est souvent observée sur les pages de l'encyclopédie. Dans les cas extrêmes,

24. Par exemple, un poste honorifique existe dans la version anglaise de Wikipédia pour Jimmy Wales, fondateur de Wikipédia (Arazy et collab., 2014)

FIGURE 1.3 – **Résolution des conflits.** Schéma représentant l’incapacité de la communauté à atteindre un but commun et comment cette contradiction est résolue.



il peut être question de guerre perpétuelle « *never-ending war* », rendant l’article peu apte à être utilisé (Yasseri et collab., 2012).

Toujours selon Freitag (2011), l’augmentation de la participation multiplie les points de vue opposés sur ce à quoi doit ressembler l’encyclopédie, autant dans sa forme que dans sa nature. Gagnant en autonomie, les éditeurs récurrents se dotent donc collectivement de mécanismes leur permettant de régler ces différends, toujours dans l’esprit initial de l’encyclopédie. Ils instaurent ainsi des mécanismes de gouvernance permettant de justifier le choix qui sera fait entre les différentes lectures individuelles (représenté dans le schéma 1.3). C’est ainsi que la décision consensuelle s’impose comme mode de fonctionnement, bien que le projet original de Wikipédia ne compte rien de tel. Aussi, il est important de préciser qu’il n’est pas nécessairement question d’avoir l’accord explicite de tous pour qu’une décision soit qualifiée de consensuelle (Kriplean et collab., 2007). Effectivement, il est plutôt question de consensus au niveau du débat : que l’argumentation mobilisée soit apte à répondre de façon satisfaisante aux principales critiques de l’un des deux camps. Pour rendre effectif ce mode de décision, l’honnêteté et l’intégrité des individus est une condition non négligeable, particulièrement quand il est question de l’utilisation frauduleuse de « *sock puppets* »²⁵, se traduisant littéralement comme « marionnette à main ». Ce terme réfère aux cas de création de faux comptes permettant d’augmenter artificiellement le point de vue d’un parti pris particulier lors de discussions. Sans une participation honnête et intègre, c’est la validité même du mode de décision qui est remise en cause (Bar et Maheswaran, 2014). Afin de juger de cette entente, la communauté s’est dotée d’administrateurs élus par les contributeurs. Le mandat de ces administrateurs est révocable en tout temps par cette même communauté²⁶. En tant qu’administrateurs, ceux-ci sont responsables de juger de l’atteinte d’une entente consensuelle. Par tradition, l’administrateur qui prend cette décision doit s’être abstenu d’être participant au débat sur lequel il tranchera.

25. http://en.wikipedia.org/wiki/Wikipedia:Sock_puppetry, consulté le 22 juin 2014.

26. <https://en.wikipedia.org/wiki/Wikipedia:Administrators>, consulté le 22 juin 2014.

Le cas de EVE Online s'avère être un très bon exemple des modalités de gestion d'une telle communauté. Dans les projets collaboratifs, c'est par les participants que s'acquière la légitimité des règles (Suzor et Woodford, 2013). Ceci est particulièrement vrai lorsque ces organisations se présentent en tant que communautés libérales d'individus ayant accepté des principes communs minimaux, incluant bien souvent une gouvernance consensuelle et une protection de l'autonomie individuelle. Alors qu'en démocratie représentative, le consentement est mesuré par les élections, Wikipédia obtient ce même consentement par l'atteinte du consensus lors de débats, qui se traduit en la création d'une nouvelle politique ou règle de conduite. Toutefois, et cela n'est pas négligeable pour notre sujet, le pouvoir d'établir ces conditions de participation (*terms of use*) n'est pas égal entre tous. L'avantage de la formation des normes réside largement du côté de l'organisation. Dans le cas de Wikipédia, cette fonction est progressivement remplie par la communauté interne du projet, celle qui participe activement aux discussions normatives. La survie de ces normes est fortement reliée à leur acceptation puisque c'est de par l'assiduité des individus à les reproduire, en conformant leurs actions et interactions à celles-ci, que ces dernières se maintiennent dans le temps (Freitag, 2011). Cette communauté interne n'a pas chômé, en 2006, les articles concernant les règles de conduite possédaient la plus forte croissance en volume de toute l'encyclopédie (Butler et collab., 2008) de sorte qu'il existe, en 2012, 383 articles concernant les politiques en plus de 449 règles de conduite (Müller-Birn et collab., 2013). Malgré le maintien de l'« *Ignore all rules* », nous constatons que la communauté de Wikipédia a une idée de plus en plus précise de ce en quoi doivent consister la participation et le contenu légitime de l'encyclopédie.

1.2.3 La formalisation d'une communauté

Ces considérations deviennent omniprésentes face à une communauté qui croît exponentiellement en participation au cours de ses premières années. Alors qu'il est question de quelques centaines d'individus en 2001, elle comprend quelques milliers en 2004 pour atteindre un sommet de 56 000 contributeurs actifs en 2007 (Halfaker et collab., 2012). Sur le plan des éditions, une moyenne de 180 éditions par minute au sein de l'encyclopédie en 2007 s'observe (Halfaker et Riedl, 2012)²⁷. Montrant bien le succès d'ouverture qu'est Wikipédia, il est impossible à ce point, pour quiconque, de garder un regard d'ensemble sur l'encyclopédie, même par une lecture simple (Almeida et collab., 2007). Ce qui est différent de la croissance initialement constatée, c'est qu'elle prend, ici, origine dans la participation des petits éditeurs ayant réalisé moins de cent éditions (Kittur et collab., 2007a). Ceci est apparemment en contradiction avec les affirmations de Jimmy Wales stipulant que Wikipédia repose sur un noyau stable de 500 éditeurs²⁸. Cette croissance de la participation, justifiée par une hausse de la popularité et de l'implication, est donc considérée comme la principale raison menant à l'instauration d'un

27. Bien qu'il ne soit pas spécifié s'il s'agit de l'ensemble de Wikipédia ou de seulement la version anglaise.

28. <http://www.businessinsider.com/2009/1/who-the-hell-writes-wikipedia-anyway>, consulté le 27 février 2014.

système managérial sociotechnique. Un système sociotechnique est constitué de la codépendance entre des structures sociales et des structures techniques. La communauté de Wikipédia l'utilise afin de protéger une vision minimale de l'encyclopédie (Halfaker et collab., 2012). Tel que décrit par Jahnke : « *D'une main, les communautés en ligne consistent en des acteurs qui utilisent un système technique pour communiquer et partager des connaissances. De l'autre, le système technique influence l'interaction entre les membres de la communauté* » (Jahnke, 2010)²⁹. Particulièrement, autant les aspects sociaux instaurent des artefacts techniques, autant ces artefacts influencent et limitent à leur tour les futurs comportements et structures sociales (M. Livingstone, 2012).

Pour la communauté en place, Wikipédia n'est, dès la fin de 2006, déjà plus le projet anarchique des débuts, les *manières de faire* informelles étant déjà autrement raffinées dans plusieurs domaines (Viégas, 2007 ; Viégas et collab., 2007). Une montée de l'importance des administrateurs se constate donc au même moment que la proportion d'éditations qu'ils réalisent est en baisse (relativement au nombre total d'éditations). D'ailleurs, les administrateurs constituent des acteurs privilégiés dans le maintien de la vision de Wikipédia, participant activement à communiquer avec les utilisateurs moins expérimentés (Laniado et Tasso, 2011). Comme beaucoup de contributeurs expérimentés, ils agissent dans l'intérêt de protéger une vision de Wikipédia, bien entendu dynamique, mais qui n'est pas nécessairement partagée par les nouveaux utilisateurs. En effet, puisqu'une coordination est nécessaire dans l'écriture des articles, l'ajout perpétuel de nouveaux participants est coûteux en temps et en efforts puisque les problèmes de coordination se multiplient dans les lieux et le temps (Kittur et collab., 2007b ; Kittur et Kraut, 2008). Parce que la barrière à la participation est drastiquement abaissée comparée aux autres formes d'encyclopédie, une très haute proportion de contributions et de nouveaux articles risque de ne pas être pertinente dans le cadre d'une encyclopédie telle que Wikipédia (Gelley et Suel, 2013) (ceci est entre autres dû à une norme imposant un caractère de notoriété minimale afin d'autoriser la présence d'un article, critère qui n'est pas protégé des interprétations subjectives (Taraborelli et Ciampaglia, 2010)). Comme l'écrit Joyce, « *Wikipédia est notable parce qu'elle a réussi à créer un système sociotechnique dans lequel l'apparente contradiction logique entre la participation ouverte et la contrainte des rôles et des décisions basées sur des règles, peuvent coexister ; il s'agit d'une bureaucratie délibérative* » (Joyce et collab., 2012)³⁰. Toutefois, les auteurs ne s'entendent pas à savoir s'il s'agit bien d'un processus de bureaucratisation de Wikipédia (Butler et collab., 2008), ou au contraire, l'affirmation de son caractère collaboratif désormais transcrit dans une structure sociotechnique (Aaltonen et Kallinikos, 2013).

L'une des possibilités offertes à la communauté par le cadre sociotechnique (Jahnke, 2010 ; Zelenkauskaitė et Massa, 2012) est de permettre de considérer les différentes options dispo-

29. Traduction libre.

30. Traduction libre.

nibles aux participants afin d’instaurer une structure bureaucratique. Leur but y est d’assurer que les objectifs collectivement déterminés se maintiennent comme principaux motivateurs de l’action collective (Joyce et collab., 2012). Ceci est d’autant plus important que la perception commune que l’utilisateur lambda possède de Wikipédia tourne encore énormément autour de la flexibilité et l’ouverture offertes par le logiciel. Chez les nouveaux contributeurs, la supposition reste donc souvent que la participation à l’encyclopédie y est égalitaire, libre et non structurée (Butler et collab., 2008). « *L’idée que beaucoup d’individus ont de Wikipédia est que celle-ci représente un phénomène émergent – la sagesse de la foule, l’intelligence de l’essaim – des milliers et des milliers d’individus ajoutant chacun un peu de contenu et de cela émerge un corps de travail cohérent* » (Niederer et van Dijck, 2010)³¹. Après tout, il est indiqué sur la page principale « *Bienvenue à Wikipédia ; le projet d’encyclopédie libre que vous pouvez améliorer* »³². Comme il est constaté jusqu’à maintenant, c’est toutefois dans sa nature dynamique que la communauté aborde le cadre sociotechnique puisqu’il offre particulièrement un grand éventail de possibilités pour organiser et structurer la participation. Cette liberté effective de la communauté sur l’infrastructure lui donne les coudées franches lorsqu’elle discute d’une possible nouvelle règle ou norme. Bien que la communauté ne le formule pas en ces termes, ces mesures se retrouvent effectivement à limiter la capacité d’agir des éditeurs en dehors des objectifs de coordination adoptés par la communauté (Yang et collab., 2013). Elle donne la base pour l’établissement d’un système de coordination large autorisant l’instauration des mécanismes permettant d’appliquer les normes méritocratiques (Cassiman et collab., 2007).

Statut interne	Effectif
Conseil d’administration	10
Administrateurs système & Développeurs	44
Intendants	38
Vérificateur d’utilisateurs	13-39
Masqueurs de modifications	38
Bureaucrates	35
Administrateurs	1409
Bots	738
Utilisateurs enregistrés (confirmation du courriel)	524
Utilisateurs enregistrés (autoconfirmation)	4 902 643
Utilisateurs anonymes	environ 30%
Utilisateurs bloqués	+ de 2500

Tableau 1.3 – **Hiérarchie de droits et pouvoirs des contributeurs à Wikipédia.** Classement selon les utilisateurs ayant le plus de droits et de pouvoirs (haut) jusqu’à ceux en ayant le moins (bas).

31. Traduction libre.

32. https://en.wikipedia.org/wiki/Main_Page

À l'aboutissement des 5 premières années, c'est donc l'instauration de protocoles précis concernant la distribution des permissions qui est observée, accordant des droits différenciés allant de l'utilisateur bloqué jusqu'aux développeurs logiciels (Niederer et van Dijck, 2010). Le tableau 1.3 ci-présent est assemblé par nos soins à partir de données principalement fournies par Wikipédia³³. Toutefois, cette liste est loin d'être exhaustive. Dans Arazy et collab. (2014) par exemple, nous retrouvons une liste différente ne mettant pas l'emphase sur les mêmes éléments. Wikipédia se divise en un nombre de sous-communautés qui est, en effet, difficilement référençable. Quatre raisons peuvent expliquer l'élaboration d'autant de rôles : l'organisation des processus et la répartition des plus exigeants aux membres fiables et pertinents ; la gestion même de ces tâches et l'établissement de leurs règles ; la supervision dans l'établissement de nouvelles décisions, ainsi qu'en dernier lieu, leur application (Arazy et collab., 2014). De toute évidence, une forte habilité technique est souvent retrouvée chez les éditeurs en position de leadership. Le renforcement, à la fois de la démocratie et de la bureaucratie, nécessite l'apport d'individus ayant les capacités techniques de transformer les décisions en capacités organisationnelles (Müller-Birn et collab., 2013). Ces pouvoirs sont distribués selon la confiance de la communauté envers un éditeur, celui-ci ayant vraisemblablement démontré qu'il partage une vision interne de la communauté concomitante avec les notions précédemment abordées quant aux motivations (*motivation externe conçue en soi-même*). La communauté témoignera de cette confiance en leur autorisant la prise en charge de tâches toujours plus critiques pour l'encyclopédie.

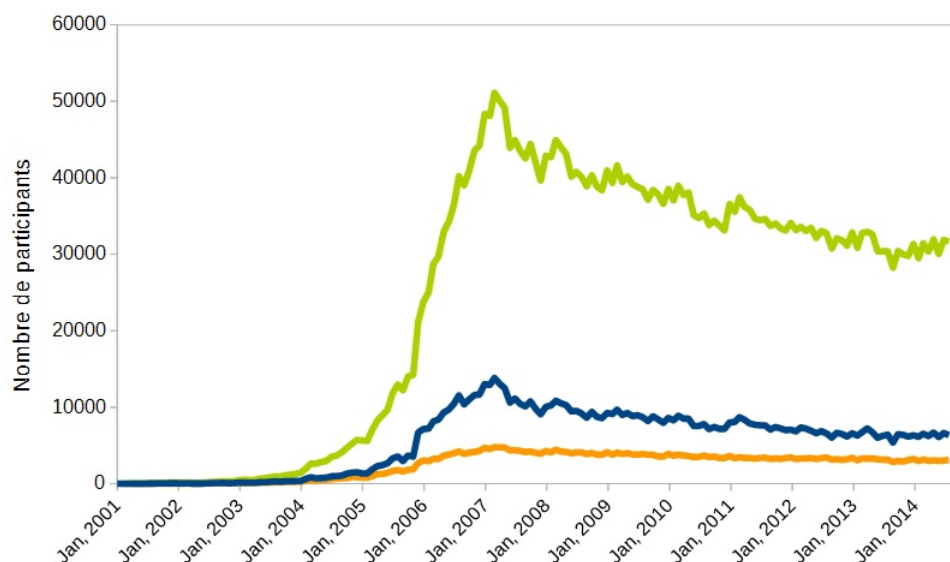
Dans la foulée de l'instauration d'un système de normes formelles, la communauté de contributeurs intensifie le contrôle qu'elle possède sur Wikipédia. En effet, l'implémentation de ces normes est à chaque fois le résultat de consensus et de processus de décisions communautaires (Cunningham et Leuf, 2001 ; Hsu et collab., 2007 ; Jahnke, 2010). Il s'agit d'autant d'artéfacts qui stabilisent progressivement une définition précise de l'encyclopédie. À terme, ce processus se traduit alors dans la concentration du pouvoir vers la communauté active, qui dispose, à la fois, des leviers contrôlant les aspects techniques de Wikipédia que des canaux légitimes de prise de décisions. Sans en être un objectif verbalisé, les utilisateurs lambdas et occasionnels s'en retrouvent toutefois relativement dépossédés. Notre prochaine section apportera un éclairage supplémentaire sur ces considérations.

1.3 Participation à Wikipédia, des signaux contradictoires

En tant que communauté libérale et volontaire, il est aisé pour la communauté interne de supposer que les nouveaux contributeurs acceptent les règlements fondamentaux de l'encyclopédie. Ceci est sans compter que la forte présence de normes et règles de conduite n'est que peu explicitée aux nouveaux venus.

33. https://en.wikipedia.org/wiki/Wikipedia:User_access_levels, consulté le 6 juin 2014.

FIGURE 1.4 – **Participation mensuelle pour diverses catégories de contributeurs à l'édition anglaise de Wikipédia.** Représentés dans le graphique, les nouveaux éditeurs ayant réalisé leur dixième édition au cours du mois (bleu), les éditeurs ayant réalisé au moins 5 éditions (vert) et au moins 100 éditions (orange) pendant le mois. Données obtenues ici : <https://stats.wikimedia.org/EN/TablesWikipediaEN.htm>.



Quand des individus contribuent pour la première fois à l'encyclopédie, leurs participations observent deux tendances : celles qui ne participent qu'à quelques occasions sans réussir à s'intégrer, ainsi que celles devenant membres à long terme (Ciampaglia et Vancheri, 2010). Toutefois, autant les contributeurs récurrents que les occasionnels sont désormais proportionnellement en diminution dans l'encyclopédie, en comparaison du nombre de participants total (Halfaker et collab., 2012). En effet, la hausse exponentielle du nombre d'articles et d'utilisateurs est constatée avoir pris fin au cours de l'année 2007 (Ransbotham et Kane, 2011). Il se constate aussi une diminution de la « classe moyenne » de l'encyclopédie depuis 2008 (Suh et collab., 2009), tous les groupes d'éditeurs sauf ceux réalisant plus de 1000 éditions par mois voyant leur effectif stagner ou même diminuer (voir figure 1.4). En fait, la durée moyenne de participation d'un contributeur se situe entre 200 et 400 jours (Soto, 2009). Ces constatations sont aussi retrouvées au sein d'une étude supportée par la Fondation Wikimedia³⁴, tel que représenté dans la figure 1.5. Quatre raisons sont évoquées pour expliquer les contraintes à la croissance de Wikipédia (Suh et collab., 2009) : le nombre en diminution des opportunités de participation, le coût à la hausse de la supervision bureaucratique et de la coordination, le nombre en baisse de sujets pouvant facilement fonder un article ainsi que la hausse de la qualité des outils utilisés par les éditeurs et administrateurs. Particulièrement, le manque d'opportunité est censé favoriser les conflits et la domination par les membres les plus impli-

34. <https://wikimediafoundation.org/wiki/Home>

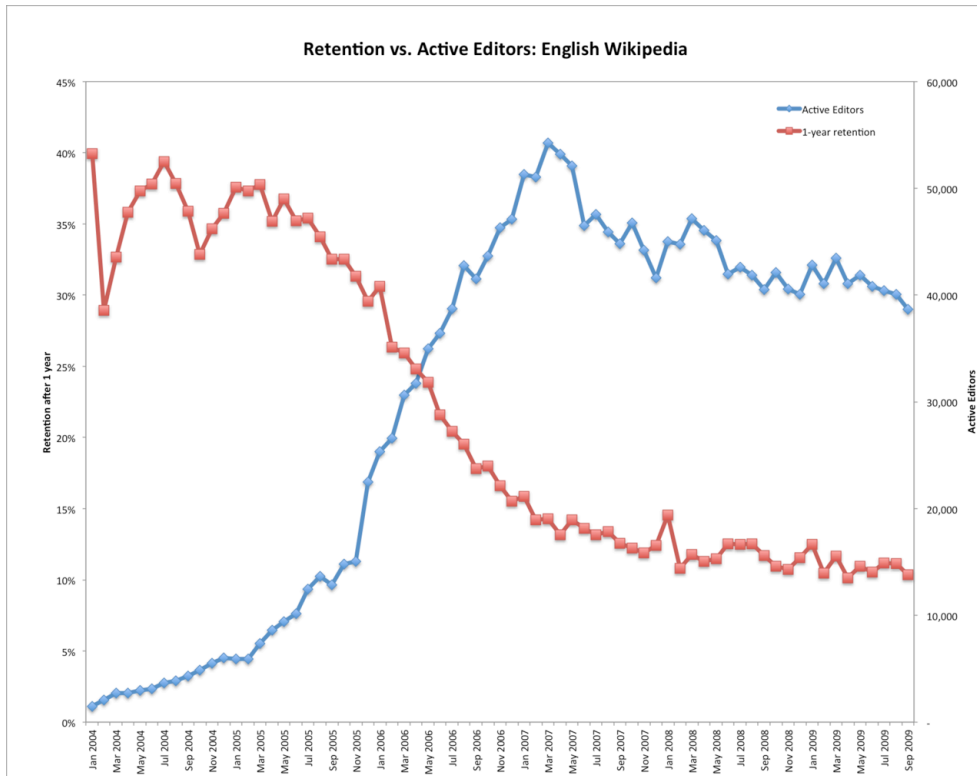


FIGURE 1.5 – **Taux de rétention versus le nombre d’éditeurs actifs dans le Wikipédia anglais.** Graphique provenant des résultats d’une étude interne publiée à l’adresse suivante : https://strategy.wikimedia.org/wiki/Editor_Trends_Study.

qués (Suh et collab., 2009). L’annulation des premières éditions d’un individu est d’ailleurs fortement corrélée à l’abandon de sa participation (Halfaker et collab., 2011, 2012). Il faut dire que les nouveaux éditeurs voient une augmentation de leurs éditions refusées (Aaltonen et Kallinikos, 2013), avec 60 % des utilisateurs enregistrés qui vont cesser leurs contributions après seulement 24 premières heures (Panciera et collab., 2009). Entre autres, la politique agressive touchant les nouveaux articles semble participer directement à ce « *turnover* », puisqu’il est constaté 6 fois plus élevé chez ceux voyant leur premier article supprimé (Gelley et Suel, 2013).

Cette difficulté à retenir les contributeurs constitue une inquiétude pour beaucoup de chercheurs (Aaltonen et Kallinikos, 2013 ; Butler et collab., 2008 ; Faraj et collab., 2011 ; Ransbotham et Kane, 2011). Bien qu’inquiétante, cette difficulté à enrôler et conserver les contributeurs doit toutefois être replacée dans son contexte : contrairement à beaucoup de projets collaboratifs, Wikipédia est pris avec des problèmes de surpopulation. Il s’agit d’un contrecoup du succès incontestable de l’encyclopédie. Beaucoup d’hypothèses sont émises afin d’expliquer ce phénomène, tel que représenté dans ce document³⁵ de Wikimedia, sans toutefois réussir à

35. https://upload.wikimedia.org/wikipedia/commons/4/4e/Decline_Theories_And_Support.pdf

identifier une cause principale. Pour certains auteurs (Butler et collab., 2008 ; Halfaker et collab., 2012), la mise en place de cette structure managériale sociotechnique a de toute évidence changé la relation de l'encyclopédie aux nouveaux usagers. La formalisation des normes de la communauté peut donc y être conçue comme une stratégie de l'encyclopédie afin de se protéger des contributeurs ne partageant pas les mêmes objectifs que la communauté interne (que cela soit consciemment ou non) (Gelley et Suel, 2013). La communauté, tout en réalisant de plus en plus d'efforts afin de « conditionner » les nouveaux contributeurs (Geiger et collab., 2012), semble donc prête à accepter cette baisse de participation (Priedhorsky et collab., 2007 ; Stvilia et collab., 2005).

La littérature concernant la participation à Wikipédia reconnaît l'importance de la communication avec les nouveaux utilisateurs quand il est question de leur rétention (Zhu et collab., 2013 ; Ransbotham et Kane, 2011 ; Halfaker et collab., 2011). Malgré tout, le premier message que reçoivent ces derniers est dans moins d'un tiers des cas, en 2011, réalisé par un autre éditeur. Un tiers consiste en des messages semi-automatisés provenant de logiciels d'assistance à l'édition et un dernier est envoyé à l'aide d'agents logiciels appelés bots (Geiger, 2014). Sachant que les bots ont l'interdiction formelle d'envoyer des messages de bienvenue, considérés trop froids et impersonnels, il s'agit donc essentiellement de messages d'avertissements (spam, vandalisme, erreur dans la saisie d'une adresse, etc.). Il est difficile de dire que cela soit accidentel ; dans le cas de ClueBot NG, principal bot anti-vandalisme de Wikipédia anglais, son algorithme est conçu pour discriminer directement les nouveaux utilisateurs ainsi que les utilisateurs anonymes (Velasco, 2010 ; M. Livingstone, 2012 ; Okoli et collab., 2012). Notre prochaine section sera donc dédiée à l'étude de l'utilisation des bots au sein de Wikipédia, outils privilégiés par la communauté dans un nombre croissant de situations au sein de l'encyclopédie.

Chapitre 2

Les bots dans Wikipédia

Dans ce chapitre, l'utilisation des bots au sein de la version anglaise de l'encyclopédie est au centre de notre intérêt. Tout d'abord, nous analyserons la littérature touchant les phénomènes cognitifs inscrits dans la perception des contributeurs quant au travail réalisé par les bots. Par la suite, nous retraçons un historique de l'émergence des bots dans Wikipédia, en dénotant particulièrement l'instauration de normes qui accompagnent leur acceptation. Ceci nous mènera jusqu'à la création d'un portrait identifiant l'utilisation contemporaine des bots. Pour conclure l'explication des concepts utilisés durant cette étude, nous observerons l'utilisation des bots selon un cas de figure particulier : la lutte contre le vandalisme dans l'encyclopédie.

2.1 Apparition des bots

Avec la montée en puissance des communautés de production collaborative, l'utilisation des bots atteint une ampleur inégalée, dont les conséquences n'ont pas encore été complètement évaluées. Les agents logiciels, appelés bots, sont généralement interprétés comme étant des programmes dont les tâches sont déterminées par algorithmes, incluant l'édition et la maintenance d'articles, et sont aussi décrits comme des outils ou protocoles managériaux touchants l'infrastructure d'un projet (Geiger et Ribes, 2010 ; Müller-Birn et collab., 2013 ; Niederer et van Dijck, 2010). Utilisés par différentes catégories de contributeurs de l'encyclopédie, les agents logiciels vont aussi parfois rajouter directement du contenu à des articles. Toutefois, ce n'est pas seulement Wikipedia qui porte un intérêt quant à la création d'agents logiciels. Parmi les doyens des communications numériques, les utilisateurs d'IRC « *Internet Relay Chat* »¹ sont depuis longtemps exposés au travail de bots s'occupant de tâches administratives ou fournissant des données particulières aux usagers (Golbeck et Mutton, 2004 ; Canavan, 2005 ; Geiger, 2014)².

1. <http://www.irg.org>, consulté le 3 octobre 2014.

2. Le plus ancien IRC bot connu date de 1991 : <https://groups.google.com/forum/#!topic/alt.irc/DaOSPwtIDPY>, consulté le 25 février 2014.

Si nous pouvons désormais constater l'influence incontestable que représentent les agents logiciels (bots), il ne faut pas oublier que leur utilisation est facilitée par l'émergence des APIs « *Application Programming Interface* », tel que souligné par Geiger (2014). Les APIs consistent en des protocoles permettant à n'importe quel utilisateur de réaliser des requêtes sur les bases de données d'un projet/site internet (ici, les serveurs de MediaWiki), autant pour obtenir des informations que lui en communiquer (Geiger, 2014). Elles permettent, entre autres, de faciliter le travail des universitaires quand ils questionnent les bases de données dans le cadre de leurs recherches touchant Wikipédia. Ces mêmes dispositions ont rendu possible l'utilisation de logiciels d'assistance à l'édition par les contributeurs (Halfaker et collab., 2012). Ceux-ci consistent en des algorithmes utilisés grâce à une interface graphique assistant les contributeurs dans leurs décisions, ce qui augmente radicalement leur capacité d'actions. Ces logiciels agissent essentiellement comme des sélecteurs de contenus répondant à une série de conditions (telle la détermination d'éditations considérées comme du vandalisme). En proposant souvent une action à prendre, le logiciel va présenter une sélection de données directement à l'utilisateur, qui n'a habituellement qu'à valider la décision prise par le logiciel. Ceci est d'autant plus vrai que désormais, MediWiki fournit une structure de base à tous ceux qui voudraient développer des bots³. Ces éléments participent à la dynamique qui fait des agents logiciels un outil privilégié et désormais incontournable dans la gestion de l'encyclopédie.

2.1.1 Historique de l'utilisation des bots dans Wikipédia

C'est dès les débuts de l'encyclopédie que certains membres de la communauté font usage de bots, aussi tôt qu'en 2002 (Steiner, 2014). Ces agents logiciels réalisent principalement des migrations de contenu libre de droits provenant de sources déjà bien établies, grâce aux précédents projets touchant le domaine public. Ces ajouts sont réalisés dans l'objectif d'accélérer l'implémentation de contenu dans l'encyclopédie. L'article de Wikipédia retraçant cette histoire^{4 5} constitue un regard des premiers intéressés sur ce phénomène. Ce « rapport interne », écrit par des contributeurs de l'encyclopédie, est résumé dans le tableau 2.1 et montre bien l'utilisation sporadique qui est faite de ces premiers bots, souvent programmés pour la réalisation d'une tâche unique.

L'utilisation des bots est rapidement devenue un sujet d'engouement et d'inquiétudes chez les contributeurs. En fait, l'un des risques négatifs liés à l'utilisation des bots est exposé dès le premier cas connu de leur utilisation. Ce cas consiste en un utilisateur décidant unilatéralement d'importer massivement du contenu en provenance d'un dictionnaire biblique

3. <http://sourceforge.net/projects/pywikipediaibot/>, consulté le 25 février 2014.

4. http://en.wikipedia.org/wiki/Wikipedia%3AHistory_of_Wikipedia_bots, consulté le 25 février 2014.

5. La lecture des discussions entourant l'écriture de cet article expose une construction essentiellement consensuelle sur l'historique des bots, exposant toutefois quelques interrogations quant à la valeur concrète de l'apport concernant la *Easton's Bible Dictionary* et la version 1911 de l'*Encyclopédia Britannica*.

6. <http://www.gutenberg.org/>, consulté le 3 juin 2014.

Tableau 2.1 – **Tableau résumé des premiers bots connus de l’encyclopédie.**

Date	Sujet	Spécificité
Octobre 2001	Plus de 300 articles créés à partir de la <i>Easton’s Bible Dictionary</i> . Le contenu est accusé d’être biaisé dans sa nature et dans son écriture.	Le propriétaire a usé d’un compte anonyme même s’il est un contributeur régulier. Il négocie après coup avec les autres participants.
Février 2002	Création de centaines d’articles à partir de la <i>Federal Standard 1037C</i> , définissant des termes en télécommunication.	La base de données a été pré-nettoyée avant de servir à la création d’articles.
Août 2002	Première importation massive d’articles de l’ <i>Encyclopédia Britannica 1911</i> à partir du projet Gutenberg ⁶ .	Il s’agit du même contributeur que pour le cas de la <i>Easton’s Bible Dictionary</i> .
Octobre 2002	Création de 30 000 articles touchant des villes américaines, provenant des données de recensements américains.	L’action de Ram-Man et de son bot <i>rambot</i> entraîne la création de 2000 articles corrompus, résultant directement en l’instauration progressive d’un cadre régissant les bots.

du domaine public⁷. Dû à la nature biaisée et datée du contenu, mais encore plus parce que l’événement se réalise grâce à un compte d’utilisateur anonyme, le bot est immédiatement bloqué de sa capacité d’édition au sein de l’encyclopédie. Il s’avère toutefois qu’il s’agit d’un des membres de la communauté interne, ce qui mène les principaux acteurs à rapidement trouver un terrain d’entente quant à son utilisation. Cela se traduit par l’établissement de quelques règles informelles (tel une limite d’une édition par 20 minutes, de s’assurer de ne pas remplacer des articles déjà existants, etc.).

En fait, le véritable choc amenant les bots au centre de l’intérêt des membres de la communauté interne prend place l’année suivante. À ce moment, le contributeur Ram-Man et son bot *rambot* crée très rapidement 30 000 articles sur des villes et comptés américains. L’événement est particulièrement notable parce qu’à l’époque, le contenu de Wikipédia en anglais est composé de 50 000 articles. Cet apport fait donc croître l’encyclopédie de 60 % sur un très court laps de temps. C’est par cette action d’importation massive que se déclenche un véritable débat mettant pour la première fois les bots au centre de l’intérêt des participants (M. Livingstone, 2012). Lors de ce débat, certains participants se demandent, à juste titre, la nécessité d’un tel apport en contenu. Alors que la plupart des articles existants se trouvent encore au stade de brouillon, l’encyclopédie se dirige, avec ce genre de contributions, vers la création d’un article pour chaque « *one-horse town* » de la planète (M. Livingstone, 2012). C’est ainsi donc que commence le débat sur l’inclusion officielle des bots au sein de Wikipédia,

7. Le débat est archivé ici : <http://lists.wikimedia.org/pipermail/wikipedia-l/2002-August/>, consulté le 3 juin 2014. Pour le premier contact réalisé par le propriétaire du script, voir le message 003587.

rambot constituant informellement le premier bot autorisé sur Wikipédia (Halfaker et Riedl, 2012). Ceci pousse la communauté à instaurer un cadre formel entourant l'utilisation des bots.

2.1.2 Instauration de normes

Par son action non concertée, *rambot* fait émerger le premier débat normatif sur la menace d'un bot pouvant agir sans concertation, que cela soit avec ou sans intentions malveillantes. Le potentiel d'action massif des bots expose à tous la possibilité d'engendrer rapidement de lourds dommages au contenu de l'encyclopédie ou d'y manipuler de l'information à grande échelle. Cette inquiétude est amplifiée par le fait que Wikipédia se présente en tant que site généré entièrement par les utilisateurs (Halfaker et Riedl, 2012). Donc, autant dans la promotion que dans les inquiétudes touchant les bots, le discours dominant concerne essentiellement l'aspect multiplicateur de force que représentent les bots. Toutefois, puisque les premiers cas de leur utilisation se réalisent sans problèmes majeurs, les utilisateurs vont être autorisés à continuer leurs expérimentations d'utilisation des bots. La discussion enclenchée par les actions de *rambot* va se conclure par la création du *Bot Approvals Group*. Bien qu'abordé plus tard, la création de ce groupe peut-être situé au sein de la ligne du temps de la figure 2.1.

Ce pouvoir multiplicateur de force, utilisé dans la promotion des bots, alimente parallèlement la principale inquiétude retrouvée sur Wikipédia à leur égard. Inquiétude exposée par le cas de *Willy on Wheel*. C'est entre 2004 et 2005 que ce bot malicieux fait une première apparition marquante sur l'encyclopédie. Ce bot particulier, dont l'histoire peut être retracée dans

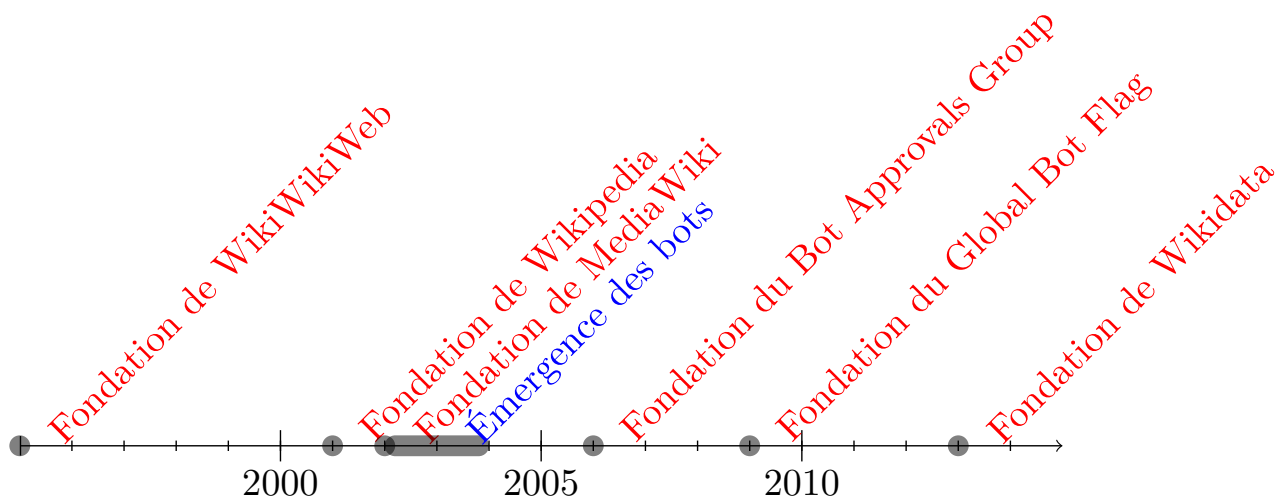


FIGURE 2.1 – Ligne du temps des points marquants touchant l'implantation des bots au sein de Wikipédia

les annales de l'Internet⁸, a causé beaucoup de tracas à la communauté. Il s'agit d'un bot réalisant du vandalisme de masse (la nature de ses gestes lui a donné le qualificatif de « *troll* »), modifiant chacune des pages passant sous sa main. Spécifiquement, il agit en transformant le titre des articles pour y ajouter « *on Wheel* » à la fin, brisant *de facto* la classification interne de l'encyclopédie. Pour ceux qui observent le déroulement des événements, il est clair que les auteurs de ce bot ont une connaissance fine du fonctionnement de l'encyclopédie ainsi que des individus ayant le pouvoir de les arrêter, démontrant une forte capacité d'ingénierie sociale (M. Livingstone, 2012). Comme dans le cas de la *Easton Bible Dictionary*, précédemment mentionné, les auteurs semblent donc provenir de la communauté interne de l'encyclopédie. La réponse qui y est donnée est autant sociale que technique. Sur le plan social, il est dès lors décidé que les contributeurs anonymes de Wikipédia n'auront plus le droit de renommer, et donc de déplacer, des articles par eux-mêmes⁹. Le déplacement de page devient un privilège des administrateurs à cet instant. Ces anonymes peuvent toutefois continuer de faire partie du processus décisionnel pouvant mener au déplacement des dits articles. Sur le plan technique, il s'observe une légitimation des bots pour de nouvelles tâches. En effet, certains bots instaurés d'urgence sont mis à contribution pour protéger dynamiquement l'encyclopédie contre ce type de vandalisme (M. Livingstone, 2012). C'est donc aussi par la réaction de la communauté à cet épisode qu'est abordée pour la première fois la possibilité d'autoriser des pouvoirs d'administrateurs aux bots. Cet événement marque l'esprit de certains membres de la communauté, puisque la conférence de Kriplean et collab. (2008) donne l'occasion d'observer que des membres utilisent le terme « *Willy suckpuppets* » afin de catégoriser ce type particulier de vandales.

Une norme supplémentaire s'implante après le déroulement de ce qui est considéré comme étant le plus gros conflit concernant l'utilisation des bots dans la jeune histoire de Wikipédia. Cette norme, présentée en tant que « *opt-out* », est décrite en détail par Geiger et collab. (2011). Celle-ci est instaurée avec l'arrivée d'*HagermanBot* en 2006, un bot dédié à signer les messages publiés sur les pages de discussions en arrière-plan des articles. La signature des messages est depuis longtemps conseillée aux membres, mais son utilisation se veut, tel la majorité des normes et règles de conduite, optionnelle. L'arrivée d'*HagermanBot* systématise son application. À cette étape, les bots demeurent principalement perçus en tant que multiplicateurs de force par la communauté, tel que rapporté par Stvilia et collab. (2008). Ce débat est important, car il cristallise la représentation d'une deuxième forme de bots chez les utilisateurs : ceux venant imposer le respect d'une norme adoptée par la communauté. Mais à la différence de systématiser l'application de normes et de gabarits techniques, il est ici question de normes régulant les formes de contributions légitimes des participants. Cet événement est

8. http://www.thinkdigit.com/Internet/Top-10-Trolls-in-Internet-History_3261.html, consulté le 26 février 2014.

9. http://en.wikipedia.org/wiki/User:Dcoetzee/Willy_on_Wheels:A_Case_Study, consulté le 25 février 2014.

l'occasion d'observer le rejet d'une règle qui n'a pourtant causé que peu de remous jusqu'à ce moment dans la communauté. Le passage d'une application volontaire jusqu'à une application systématique illustre bien l'impact problématique que peut posséder une gouvernance logarithmique. C'est donc à la suite de ce débat fort en émotions et à l'impossibilité d'une entente sur l'utilisation actuelle de ce bot qu'est adoptée la norme « *opt-out* ». Elle permet à un contributeur de signifier son refus de voir ses éditions modifiées ou supprimées par des tâches réalisées par les bots. Cette exception « *opt-out* », appliquée grâce à un gabarit nommé « *no-bot* », peut aussi s'appliquer sur un article à part entier (Halfaker et collab., 2012). Afin de bénéficier de cette possibilité, il faut toutefois être un utilisateur enregistré. Cette norme permet de créer une porte de sortie pour les utilisateurs les plus critiques des bots. À défaut d'une entente stricte entre tous, cette solution va permettre aux promoteurs de l'utilisation des bots de poursuivre le transfert des tâches laborieuses vers des agents logiciels (Geiger et collab., 2011). Cela nous permet aussi d'exposer ce que la communauté entend par « consensus » lors de sa prise de décisions. Il ne s'agit pas d'obtenir l'unanimité lors d'un débat, mais bien de répondre à toutes les critiques valables par une réponse qui, sans nécessairement invalider la critique, la contient sous le seuil du tolérable pour tous. Sans régler l'ensemble des débats, l'ajout de ce gabarit « *no-bot* » permet effectivement d'atteindre ce seuil du tolérable pour la communauté. D'ailleurs, cette norme semble souvent usée par des utilisateurs récurrents réalisant des tâches limites ou complexes, déclenchant souvent les drapeaux rouges des bots. La mise en place de ce garde-fou qu'est la norme « *opt-out* » va avoir beaucoup d'influence sur l'utilisation subséquente des bots, désormais épargnés de la reproduction de ce conflit.

Avant ce conflit, la communauté se trouve dans un paradigme incluant peu de déterminismes technologiques. Les bots appliquent des décisions représentant l'image de Wikipédia, souvent en participant à l'apport massif de nouveaux contenus. Cette relation change après le débat concernant *HagermanBot*. Les bots deviennent plus forts et importants, mais seront d'autant plus contrôlés par la communauté. Avant d'aborder plus en détail le rôle sociostructurel que représentent désormais les bots, prenons un moment pour réaliser un portrait des processus et étapes désormais nécessaire qu'un contributeur doit effectuer avant d'être autorisé à faire usage d'un bot dans l'encyclopédie.

2.1.3 Autorisation des bots

C'est donc en 2006 qu'est officiellement instauré le *Bot Approvals Group* (BAG)¹⁰, responsable de l'autorisation et de la supervision des bots pour la version anglaise de Wikipédia. De par la puissance d'action accessible aux bots, ceux-ci doivent désormais obtenir un « *bot flag* » afin d'être autorisés à réaliser des modifications automatisées à même l'encyclopédie. Certains critères se devront d'être rencontrés afin d'obtenir ce statut : « *Selon la politique, un bot doit démontrer qu'il est : sans danger, est utile, ne consomme pas de ressources sans*

10. http://en.wikipedia.org/wiki/Wikipedia:Bot_Approvals_Group, consulté le 25 février 2014.

nécessité, réalise des tâches pour lequel il y a consensus, respecte assidûment les politiques et règles de conduite, utilise des messages informatifs, correctement formulés, pour chacun des sommaires d'éditions ou pour les messages laisser aux éditeurs » (M. Livingstone, 2012)¹¹. Le non-respect de ces conditions, ainsi que le fait de ne pas s'enregistrer avec le BAG, peut résulter en le bannissement du bot ainsi que de son propriétaire (Geiger, 2014). Rappelons qu'à la différence des bots, les utilisateurs humains n'ont à aucun moment l'obligation de s'inscrire auprès de Wikipédia afin de proposer une édition. Le processus que nous décrivons dans les prochains paragraphes quant à la version anglaise de Wikipédia est reconnu comme étant le plus bureaucratique des différentes versions de l'encyclopédie (Geiger, 2011).

La première étape dans l'obtention d'une autorisation est de compléter une demande auprès du Bot Approvals Group (BAG)¹². Ce groupe est composé principalement par des membres ayant un statut de bureaucrates et d'administrateurs. Selon les critères déjà énoncés¹³, le BAG évalue individuellement les différentes tâches soumises par le demandeur. Notons que certains types de tâches y sont spécifiquement refusés : la correction de l'orthographe (puisque les mots ont un contexte spécifique qui ne peut pas être validé sans jugement humain) ; la correction des liens interwiki (depuis 2013, puisque cette tâche est désormais remplacée par le projet WikiData¹⁴) ; la création massive d'articles (tel que réalisé par *rambot* en 2006 (Kennedy, 2009)) ; ainsi que les changements dits cosmétiques (tel le remplacement de caractères invisibles). Il est donc constaté que les tâches des premiers bots réalisant des importations massives ou des corrections orthographiques (Geiger et Ribes, 2010), sont désormais interdites dans l'encyclopédie. Malgré le fait d'avoir participé à la promotion de l'utilisation des bots, ces premières tentatives ont été jugées maladroites et malvenues. Après cette acceptation initiale des tâches, une période d'essais habituellement de courte durée (entre 20 et 100 éditions) va être accordée afin de valider la bonne capacité du bot à réaliser la tâche particulière (tel un faible taux de faux-positifs), ainsi que de la présence du propriétaire. Dénotant l'importance de l'acceptation par la communauté, cette période de validation veut s'assurer que le propriétaire soit proactif afin de répondre aux problèmes potentiels liés à l'utilisation du bot. Le propriétaire se doit donc de répondre aux questionnements de la communauté et des erreurs commises par son bot, par exemple en réalisant des mises à jour de son code. Étant donné cette force perturbatrice que peut représenter un bot, les autorisations ne sont habituellement accordées qu'à des individus ayant réalisé plus de 500 éditions à l'encyclopédie. Une fois la période d'essai terminée, le BAG peut autoriser l'utilisation du « *bot flag* », souvent en y assortissant certaines conditions, telle la fréquence maximale d'éditions. De plus, depuis 2009, il existe un « *global bot flag* » qui permet d'être autorisé à éditer dans la majorité des versions de l'encyclopédie. Le processus d'acceptation devra être initié à nouveau pour toutes tâches

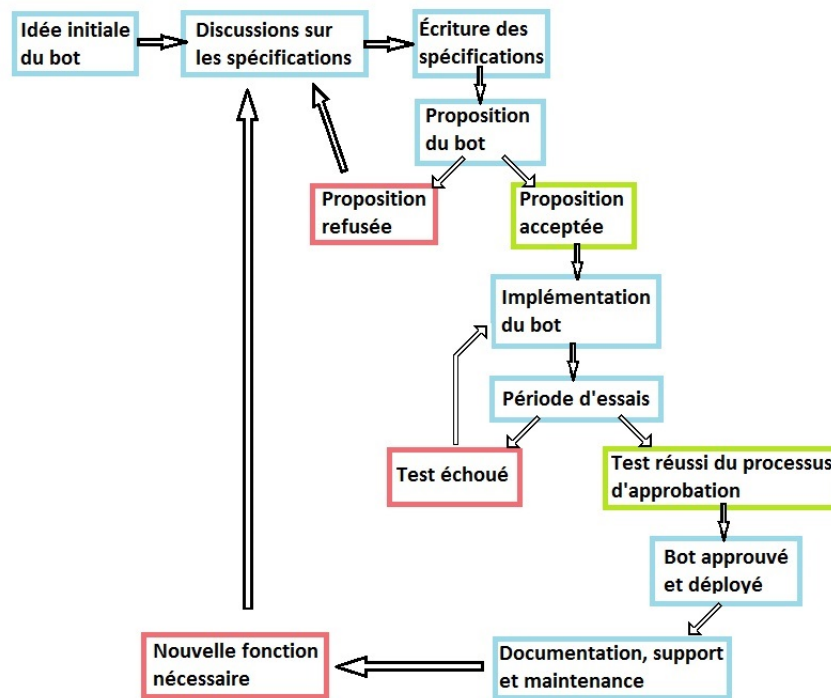
11. Traduction libre.

12. http://en.wikipedia.org/wiki/Wikipedia:Bots/Requests_for_approval, consulté le 3 juin 2014.

13. http://en.wikipedia.org/wiki/Wikipedia:Bot_policy, consulté le 3 juin 2014.

14. https://www.wikidata.org/wiki/Wikidata:Main_Page, consulté le 3 juin 2014.

FIGURE 2.2 – Schéma du processus d'autorisation des bots. Reproduction et traduction d'un schéma retrouvé chez Ciampaglia (2011).



supplémentaires soumises. Voici une schématisation des étapes en figure 2.2. De plus, nous retrouvons généralement un bouton d'urgence permettant à n'importe quel administrateur de désactiver le bot si ce dernier en vient à agir chaotiquement. À défaut d'être obligatoire, l'absence de cette option peut obliger l'administrateur à bloquer le bot ¹⁵.

Outre l'automatisation d'une tâche, nous retrouvons divers autres avantages autorisés par l'utilisation d'un bot. Il y a d'abord la possibilité d'éditer à une vitesse plus élevée que ce qui est permis à un utilisateur humain sans déclencher les drapeaux rouges utilisés par les équipes de la lutte au vandalisme (équipes composées à la fois de bots et d'humains). Les bots ont aussi le droit d'agir sur les pages semi-protégées (Halfaker et collab., 2012), sans compter que si la tâche en question le nécessite, il sera possible d'obtenir les droits d'administrateurs (seulement si le propriétaire est lui-même un administrateur). De plus, les éditions qu'ils réalisent vont être cachées de la liste des différentes versions des articles, puisque leurs hautes fréquences risquent de créer de la confusion pour les utilisateurs consultant l'historique des pages. Encore davantage, notons que les éditeurs réguliers utilisant la « *watch list* », une liste permettant de suivre les modifications récentes apportées aux articles auxquels ils sont abonnés, peuvent y cacher les contributions des bots en y cochant une option particulière.

15. http://en.wikipedia.org/wiki/Wikipedia:Bot_policy, consulté le 3 juin 2014.

Enfin, les éditeurs ne voulant pas interagir avec les bots peuvent désormais insérer un gabarit « no-bot », évoqué précédemment.

Toutefois, les bots se doivent d'éviter toute contre-annulation d'une édition. En ce sens, si l'édition d'un bot est annulée par un humain, ce premier n'est pas autorisé à la contre-annuler. Il s'agit d'une version plus stricte de la norme du 3RR (*3 reverts rule*), limitant à 3 fois le nombre autorisé de contre-réversions entre deux éditeurs, politique elle-même appliquée grâce au bot 3RRbot (Geiger et collab., 2011).

2.2 Entre contrôle et liberté

Si le processus entourant l'utilisation des bots est si détaillé, c'est bien parce que leur utilisation est fréquente et récurrente. Cette section établit un portrait de l'utilisation contemporaine des bots, avant de détailler la façon dont le travail des bots s'insère dans la structure sociale et technique de l'encyclopédie.

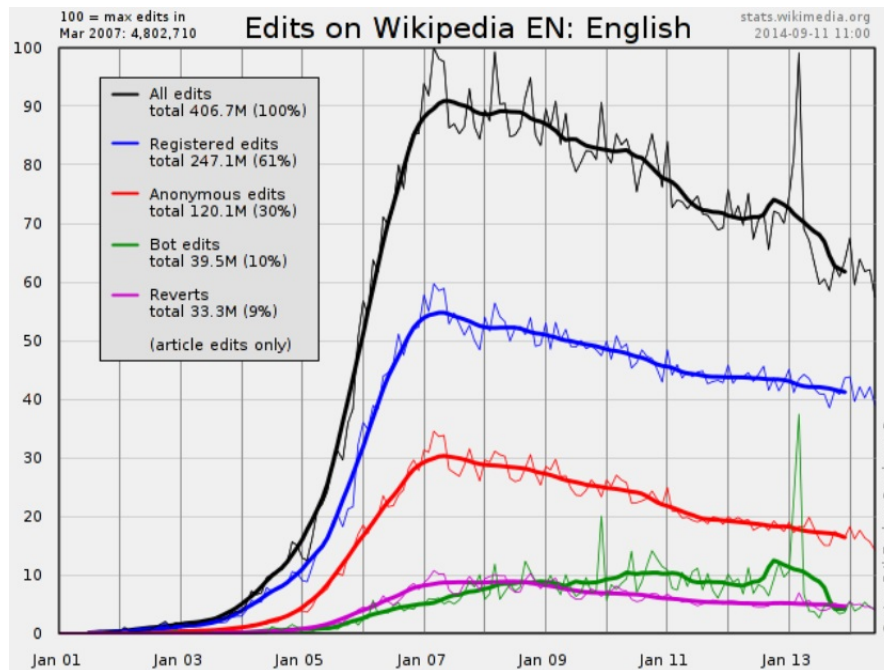
2.2.1 Portrait général de l'utilisation des bots dans la version anglaise de Wikipédia

Peu d'articles et de communications concernant l'utilisation des bots dans Wikipédia sont disponibles avant 2009. Ceci est dû à leur rôle essentiellement anecdotique sur le terrain, ne représentant que 3-5 % des éditions totales de la version anglaise de l'encyclopédie en 2005-06 (Geiger, 2009), proportion qui monte à 16,33 % au cours de l'année 2009. En 2012, il est possible d'estimer que les éditions réalisées par les bots représentent approximativement 22 % des éditions totales de l'encyclopédie¹⁶. Depuis la création de l'encyclopédie jusqu'à 2014, les bots réalisent une moyenne globale de 10 % (39,6 millions sur les 412 millions) de l'ensemble des éditions de l'encyclopédie¹⁷. Cette forte participation est corollaire à l'augmentation du nombre des bots et des tâches qu'ils réalisent. Ceux-ci passent de 400 bots autorisés en 2010 (Niederer et van Dijk, 2010) jusqu'à 700 dès l'année suivante, représentant 1554 tâches (M. Livingstone, 2012). À la fin de 2013, il s'agit de 872 différents bots. Ces deux décomptes sont constitués des tâches et bots autorisés dans la version anglaise de Wikipédia, ceux-ci peuvent donc, sans avoir le « *global bot flag* », être principalement utilisés dans une autre version de l'encyclopédie. Cette proportion baisse si nous ne comptons que les éditions faites au sein des articles, représentées en figure 2.3.

16. Ratio de données provenant de deux sources : https://en.wikipedia.org/wiki/Wikipedia:List_of_bots_by_number_of_edits et <http://stats.wikimedia.org/EN/TablesDatabaseEdits.htm>, consulté respectivement le 18 novembre 2014.

17. <http://stats.wikimedia.org/EN/BotActivityMatrixEdits.htm>, consulté le 30 juin 2014.

FIGURE 2.3 – **Participation à l’écriture des articles.** Ce graphique est calculé à partir de statistiques internes à Wikipédia. Il peut être retrouvé à l’adresse suivante : <https://stats.wikimedia.org/EN/EditsRevertsEN.htm>.



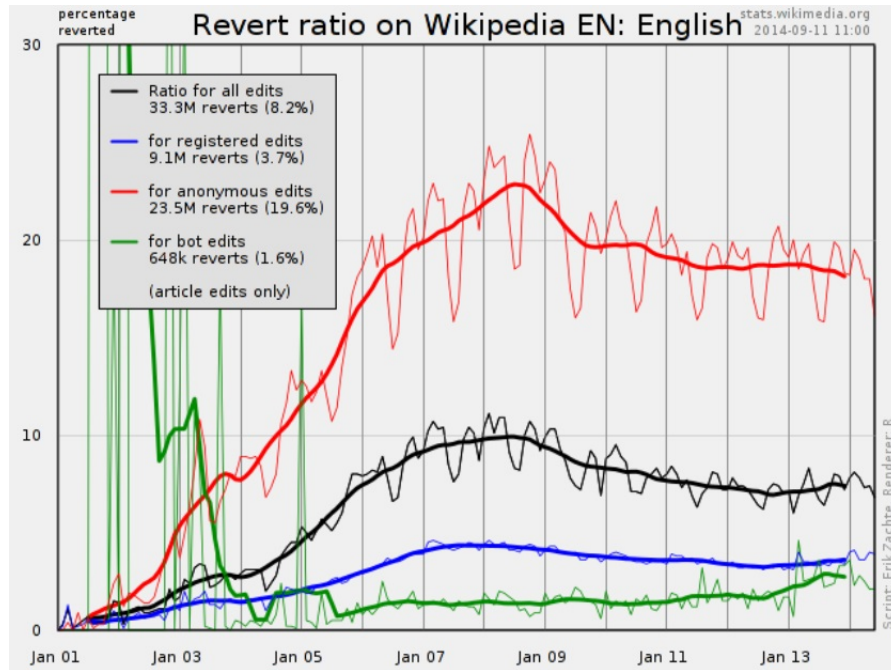
La baisse dans la fréquence d’édérations des bots s’explique par un changement significatif ayant touché Wikipédia à partir de 2013. Il s’agit de la création de Wikidata¹⁸. Wikidata est un projet cherchant à centraliser et structurer l’ensemble des relations sémantiques entre différents concepts et éléments composant Wikipédia, autant pour ce qui est des hyperliens intra-wiki qu’inter-wiki. Ces tâches étant précédemment réalisées par de nombreux bots, on constate après coup une baisse significative dans le nombre de bots actifs (Steiner, 2014). Celui-ci constate qu’environ 85 % des éditions de Wikidata sont réalisé par des bots¹⁹. De façon générale, c’est désormais en moyenne 42 % des éditions de l’ensemble des projets de MediaWiki qui est réalisé par des bots. Étonnamment, celui-ci calcule que seulement 5 % des éditions totales de l’encyclopédie seraient encore réalisées par des bots. Ces résultats sont passablement différents des statistiques internes de l’encyclopédie que nous avons utilisées précédemment. Bien que la méthodologie utilisée par Steiner (2014) semble appropriée, il semble peu probable que la proportion des éditions réalisées par les bots soit désormais aussi faible. Les statistiques de Wikipédia nous permettent d’observer qu’en 2012, il y a une moyenne de 116 bots ayant réalisé au moins 5 éditions en un mois sur différents articles, proportion réduite à environ 56 bots pour 2014²⁰. Dans tous les cas, notons que cette modification du

18. <http://wikidata.org>, consulté le 14 octobre 2014.

19. Application réalisée par Steiner permettant de calculer l’apport des bots : <http://wikipedia-edits.herokuapp.com/>, consulté le 3 juin 2014.

20. https://stats.wikimedia.org/EN/TablesWikipediaEN.htm#editor_activity_levels, consulté le 5 novembre 2014.

FIGURE 2.4 – **Proportion des éditions annulées chez diverses catégories de participants.** Ce graphique est calculé à partir de statistiques internes à Wikipédia. Il peut être retrouvé à l'adresse suivante : <https://stats.wikimedia.org/EN/EditsRevertsEN.htm>.



rôle des bots s'est réalisée après la période qui est à l'étude de ce mémoire.

Bien que l'utilisation des bots ait pris plusieurs années avant d'atteindre une utilisation quasi-systématique et structurelle, ceux-ci ont malgré tout laissé des marques de leur passage dès les premières années de Wikipédia. Déjà en 2006, 17 des 20 plus prolifiques éditeurs sont des bots (Geiger et Ribes, 2010). Cette utilisation est aussi renforcée par une très grande efficacité du travail des bots, notamment par un très bas taux d'éditions annulées (Adler et collab., 2008). Ceci peut être constaté dans la figure 2.4. De plus, 57 % des annulations d'éditions des bots sont réalisées par d'autres bots. Il n'est peut-être pas étonnant alors de retrouver que 12 des 50 des contributeurs annulant le plus d'éditions ainsi que 32 des 50 plus annulées sont des bots²¹.

2.2.2 Interaction avec les agents logiciels

Cette utilisation massive des bots par les contributeurs de l'encyclopédie est passablement originale. En effet, dans bien des médias sociaux, l'utilisation de bots est souvent très mal vue par les communautés particulières. Il n'est donc pas anodin que la majorité des écrits recensés porte sur des méthodes d'identification de l'usage de ces bots ainsi que diverses stratégies pour s'en prémunir. Du à cette perception négative, leur utilisation est principalement réalisée

21. <https://stats.wikimedia.org/EN/EditsRevertsEN.htm>, consulté le 5 novembre 2014

par des individus ayant de mauvaises intentions. Notamment, les bots y sont utilisés afin de s'attaquer à l'organisation de groupes sociaux, à en perturber leurs communications, rendre inutilisable des comptes utilisateurs ou tout simplement briser le réseau social, tel l'utilisation d'attaques de type *D-DOS*²² (Mitter et collab., 2013). L'utilisation la plus fréquente restera sans contredit dans le cadre d'activités de spam. D'autant plus que nous retrouvons désormais des applications offrant des services « clé en main » pour l'utilisation de spambots. Plus récemment, il est même possible d'observer des individus détourner ces infrastructures de spam afin de massivement perturber l'organisation sociale et politique de groupes d'oppositions, comme lors des élections en Russie (Thomas et collab., 2012). Les spambots peuvent même être utilisés pour polariser artificiellement des discussions afin d'identifier les participants à des réseaux sociaux autrement latent (Aiello et collab., 2012).

Sur un autre aspect, il faut aussi noter que l'utilisation des bots dans Wikipédia n'a pas comme objectif la création d'intelligences artificielles. Dans le cas de Wikipédia, il est plutôt question de développer des agents d'assistance algorithmique (Halfaker et collab., 2012; Müller-Birn et collab., 2013; Geiger, 2011). L'objectif ici est de prendre des décisions pondérées par l'acquisition d'informations plus ou moins facilement accessibles à des utilisateurs humains, dans le but de réaliser une tâche dûment autorisée par la communauté. Ceci éloigne donc les bots de Wikipédia de l'utilisation classique que la littérature touchant les interactions homme-machine envisage. Depuis deux décennies en effet, de nombreux chercheurs tentent de mesurer la propension des humains à travailler avec les agents logiciels ainsi que des robots physiques, particulièrement dans le cadre d'une conception orientée vers la communication avec des utilisateurs (Goodrich et Schultz, 2007). Toutefois, certaines de leurs considérations peuvent nous être pertinentes. Il est à noter que l'attente préalable que se fait un individu à propos du bot influence l'impression et l'appréciation qu'il aura de celui-ci (Rutjes, 2013). Dans le cas d'un individu au fait d'être en relation avec un bot, les interactions qu'il maintient ne risquent pas d'être influencées outre mesure (sous condition que le bot respecte des standards minimaux de communication) (Holtgraves et collab., 2007). De plus, avoir un rôle en interdépendance avec celui d'un bot, même si cela n'est qu'appréhendé, renforce la volonté des individus de travailler en équipe avec lui (Nass et collab., 1996), ce qui accroît d'autant plus le lien de reconnaissance (Nass et collab., 1996). En fait, il est constaté qu'un individu a tendance à appliquer les mêmes règles sociales et attentes envers les bots qu'envers les humains. Ces agents logiciels risquent donc d'être abordés avec les mêmes catégories sociales et règles de politesse (Nass et Moon, 2000). Ceci correspond donc à ce qui est mesuré des interactions humaines dans les communautés virtuelles; les individus ont tendance à y reproduire les mêmes normes que celles qu'ils appliquent en société (Pastinelli, 2014).

22. , consulté le 23 octobre 2014.

C'est davantage dans le cadre de la théorie de l'acteur-réseaux de Bruno Latour (Latour, 1991) que peut se comprendre l'interaction que la communauté de l'encyclopédie possède avec les bots. Cette approche établit une distinction entre les objets (physiques, limités et positifs ; qui peuvent donc être circonscrits précisément) et les choses (qui s'imposent à nous, telle une structure, une loi). Selon Latour (1991), il nous est nécessaire d'aborder ces choses comme acteurs à part entière lors de nos analyses ; ces bots médiatisent nos interactions et sont acteurs de l'organisation sociale (Latour, 1991). En plus, ces acteurs non humains sont les dépositaires de la stabilité de nos institutions et organisations. Cette façon de concevoir est parfois retrouvée sous l'appellation de CASA (« *Computers Are Socials Actors* ») (Nass et collab., 1996). En ce sens, les bots sont donc considérés comme des « agents sociaux » (Holtgraves et collab., 2007 ; Breazeal, 2002).

Quoi qu'il en soit, trois postures face aux bots vont dominer lors des premières années chez les contributeurs (Geiger, 2011). Premièrement, ils sont souvent représentés comme des outils recueillant des données ; deuxièmement, comme des acteurs participants à s'assurer que la qualité de l'information présente dans l'encyclopédie ; enfin, certains ne se les représentent qu'en tant que nuisance pour la collecte de données puisque MediaWiki n'est pas apte à faire la différence entre bots et utilisateurs. Ces deux premières facettes se regroupent dans la vision classique de l'assistant récupérant des tâches laborieuses et répétitives (Geiger, 2011). Comment cette interaction se perpétue à ce jour reste encore peu explicite.

2.2.3 Rôle structurel

Analyser Wikipédia en mobilisant un paradigme sociotechnique nous permet de bien comprendre l'ampleur et l'importance des rôles occupés par le bots au sein de l'infrastructure technique de l'encyclopédie. Lors de la fondation de Wikipédia, le logiciel MediaWiki y représente l'image initiale de l'encyclopédie. Elle consiste en la traduction technique de sa déclaration de principes, c'est-à-dire les cinq piliers. Cette ouverture du système, combinée au peu d'initiative volontairement affichée par MediaWiki d'ajouter de nouvelles fonctions, favorise l'organisation collective de la communauté. Celle-ci peut alors ajouter des fonctions *ad hoc* au logiciel MediaWiki afin d'augmenter sa capacité d'action.

Cette propension à continuellement ajouter des fonctions au logiciel s'explique simplement. En effet, WikiWikiWeb, bien qu'adapté pour Wikipédia en le logiciel qu'est MediaWiki, n'a pas été initialement développé dans l'objectif spécifique d'être utilisé dans le cadre d'une encyclopédie²³. L'auteur ne se doutait donc pas de ce qui serait nécessaire comme fonctions minimales pour cette utilisation particulière. Structurellement, les bots permettent donc d'ajouter ces fonctions, ou inversement, des limitations *ad hoc* au logiciel MediaWiki. Ces ajouts au code de MediaWiki se conceptualisent sous l'appellation de *bespoke code* (Geiger, 2014). Ce

23. <http://c2.com/cgi/wiki?FrontPage>, consulté le 15 novembre 2014.

concept, qui regroupe l'ensemble du code extérieur au logiciel même, expose la mise en place de nouveaux éléments techniques permettant de pallier aux manques du logiciel mère. Spécifiquement, il s'agit de code n'étant pas partie intégrale au projet, c'est-à-dire, aux serveurs de MediaWiki. C'est au contraire à la suite, ou parallèlement, à MediaWiki que ce code est utilisé par des tiers partis. Selon les estimations de Geiger, ceci n'est pas un phénomène négligeable, le nombre de lignes de codes considéré comme *bespoke* étant estimé dans un ordre de grandeur supérieur (x10) au logiciel même de MediaWiki (Geiger, 2014). En ce sens, l'utilisation de gabarits et de balises constitue une forme primaire de *bespoke* code, les bots en formant une deuxième forme. Enfin, les outils semi-automatiques, tel que les logiciels d'assistance à l'édition *Twinkle*²⁴, *Huggle*²⁵ ou *Auto-Wiki Browser*²⁶, en représentent une troisième. De même que l'auteur de l'article, nous remarquons que l'estimation de son importance ne prend pas en compte l'ensemble des logiciels et codes qui ne sont pas directement reliés à Wikipédia, tel le système d'exploitation, le fureteur internet, etc.

Cette façon de faire n'est pas sans risque, tel que le relève MZMcBride. Éditeur proéminent réalisant de nombreuses tâches grâce à l'utilisation de bots, celui-ci prévient que la forte utilisation des bots peut devenir un problème à long terme puisque leur travail vient cacher les faiblesses du logiciel MediaWiki (MZMcBride dans (M. Livingstone, 2012)). Considérant que ce logiciel n'a pas de service de messagerie intégré par défaut, cet avertissement doit être pris avec ambivalence. Le risque principal identifié ici est de tenir pour acquis des fonctions dont la perpétuité n'est pas associée à celle de Wikipédia. Geiger (2014) réalise à ce titre une petite histoire ethnographique des bots programmés par lui-même afin de démontrer à quel point ceux-ci sont dépendants de conditions extérieures à Wikipédia. Il y affirme que le premier bot créé pour ses recherches est passé hors d'usage indéfiniment lorsque l'auteur a déménagé d'appartement.

Enfin, plusieurs auteurs ont déjà tenté de décrire l'utilisation des bots grâce à l'usage de typologies. Dans l'article de Niederer et van Dijck (2010) par exemple, nous retrouvons une dualité entre les bots avec éditions « co-écriture » et sans éditions « administratif ». Selon cette conceptualisation, les bots « administratifs » agissent dans le but de faire respecter certaines normes, autant dans la mise en application des bannissements que du vandalisme. Selon cette catégorisation, ils s'occupent aussi d'extraire des données de l'encyclopédie, autant pour la recherche que pour mesurer la validité de la propriété publique des éléments de l'encyclopédie. Dans le cas des bots de « co-écriture », il s'agit de bots éditant à même les articles de Wikipédia. Différemment, Halfaker et collab. (2012) décrivent quatre types de bots : un, ceux transférant du contenu provenant de base de données publique ; deux, ceux surveillant et peaufinant les articles ; trois, ceux ajoutant des fonctions à MediaWiki ; ainsi

24. <https://github.com/azatoth/twinkle>, consulté le 3 juin 2014.

25. <https://github.com/huggle>, consulté le 3 juin 2014.

26. <http://sourceforge.net/projects/autowikibrowser/>, consulté le 3 juin 2014.

que quatre, ceux s’occupant de la protection de l’intégrité des articles. Enfin, une troisième classification est retrouvée chez Müller-Birn et collab. (2013). Il s’agit d’une triple division entre les « *Content focus* » bots, les « *Task focus* » bots ainsi que les « *Community focus* » bots. La première catégorie regroupe les bots agissant sur les articles. La deuxième permet d’aider les éditeurs dans leurs tâches en compilant des données pour eux ou en les informant de changement survenu pour des variables spécifiques. Enfin, la troisième regroupe surtout des tâches ne touchant pas directement les articles, principalement en appliquant des normes de participation.

2.2.4 Impact des bots sur la communauté

Considérer les bots comme des agents sociaux nous permet d’observer la mise en place de certaines des normes adoptées par la communauté. Habituellement, les normes de Wikipédia sont d’applications volontaires et souvent au cas par cas. Exécutées par les bots, il s’agit d’une application systématique des politiques et des règles de conduite (Geiger, 2011). Nous avons pu constater ceci avec *HagermanBot* et la signature systématique des commentaires. Depuis cet évènement particulier, un des aspects les plus évidents de l’utilisation des bots semble donc le « conditionnement » des nouveaux usagers, que cela soit fait directement (tel l’algorithme de *ClueBot NG* discriminant les nouveaux utilisateurs) ou indirectement par l’annulation d’éditations ne respectant pas les standards de l’encyclopédie. Cette conceptualisation de la construction des normes se résume dans le tableau 2.5.

Ce changement d’application de la mission des bots, à tout le moins dans leur portée, est bien décrit chez Müller-Birn et collab. (2013). Ceux-ci analysent le cas d’un bot en particulier, soit *Xqbot*. Dans un premier temps, ce bot s’est occupé de la suppression de pages obsolètes dans l’encyclopédie. Toutefois, les tâches subséquentes qu’il réalise s’en éloignent fortement. L’une des tâches initiales de ce bot est de retirer les votes surannés lors de la ré-élection d’un

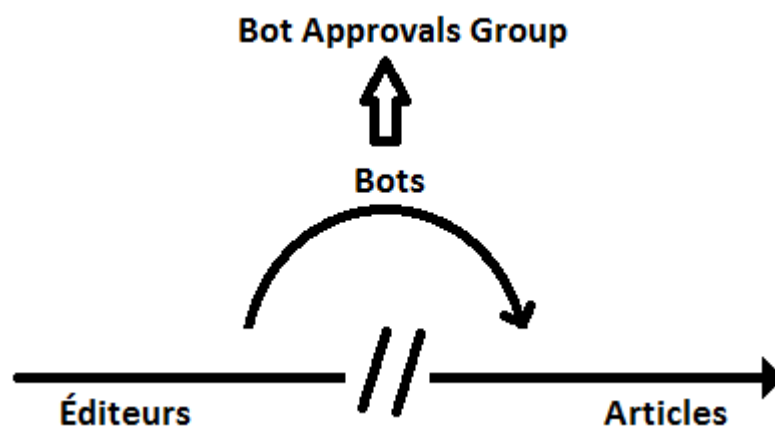


FIGURE 2.5 – Schéma représentant le rôle des bots sur le travail individuel.

administrateur. Toutefois, cette tâche s'est rapidement transférée vers la gestion d'autres types de votes et de consultations. L'utilisation de ce bot a induit à la fois un changement notable des types de tâches qu'il réalise ainsi qu'un élargissement de l'application de certaines de ses autres tâches (Müller-Birn et collab., 2013). Ces mêmes auteurs constatent que près de 40 % des éditions réalisées par les bots le sont en dehors de l'espace « article », ce qui est contre la politique officielle. Le suivi de ces bots est d'autant plus difficile qu'il est peu évident de savoir dans quelle mesure les propriétaires sont autorisés (ou contraints) dans la réécriture du code de leur bot après avoir reçu leur acceptation par le BAG (M. Livingstone, 2012).

Considérant le changement du degré ainsi que du lieu d'application parfois constaté chez un bot, un questionnement se pose : est-ce qu'un consensus obtenu dans une sphère de Wikipédia peut s'appliquer tel quel dans une autre ? L'établissement des normes étant déjà un processus laborieux, il est bien plus facile d'obtenir l'autorisation d'utilisation d'un bot appliquant cette même norme que d'établir un nouveau consensus changeant la capacité d'action de ce bot. Dans le domaine juridique, l'importance de « l'esprit de la loi » est souvent notée, qui est le contexte social relié à son adoption, mais est-ce vraiment le cas pour Wikipédia ? Il semble évident à ce point que l'utilisation des bots soit bien souvent à l'encontre de l'interprétation créative des règles (voir tableau 1.1 à la page 9), car elle ne laisse aucun espace (*a priori*) pour les acteurs d'agir autrement. Ces questions apparaissent essentielles de par l'importance structurelle accordée aux bots. En effet, puisqu'il existe une hiérarchie de droits dans Wikipédia (voir tableau 1.3 à la page 23), les différents contributeurs possèdent différents droits et privilèges selon leur échelon. Les bots, bien qu'en dessous du niveau d'administrateurs (dans certains cas, si le propriétaire est lui-même un administrateur, le bot pourra aussi avoir les droits d'administrateurs), se retrouvent au-dessus des utilisateurs enregistrés. Ceci leur donne donc une prédominance sur l'essentiel des autres utilisateurs. Bien entendu, il existe le gabarit « *no-bot* », bien qu'il soit difficile de mesurer son utilisation concrète.

Dû à ce cadre particulier et au désir de mieux coordonner les contributeurs, il est observé une valorisation des capacités techniques entre ceux-ci (Cassiman et collab., 2007). Il est donc fréquent d'y observer que les individus ayant un certain « leadership » jumellent habiletés techniques et organisationnelles (Cassiman et collab., 2007). La prééminence de ces caractéristiques chez les éditeurs de premier plan vient sans doute expliquer la propension de la communauté à vouloir résoudre ses problèmes et défis grâce à l'implémentation et l'utilisation de bots. Cassiman observe donc que les individus intéressés par l'efficacité et l'augmentation de la capacité d'action des participants « *empowerment* » proposent souvent l'utilisation de la programmation informatique. C'est que concrètement, l'utilisation des bots se traduit par une augmentation de la capacité d'action et donc, la sphère d'influence de son propriétaire.

Ce processus peut se décrire comme une gouvernance dite « algorithmique » (Müller-Birn et collab., 2013). Ceci sous-entend que la stabilité ainsi que la prise de décisions sont auto-

risées grâce à des artifices techniques. L'augmentation du temps mis par la communauté sur des aspects managériaux influencera de beaucoup le transfert de tâches vers des agents logiciels (Niederer et van Dijck, 2010 ; Geiger, 2013). Toutefois, cela a comme résultat d'obscurcir le processus menant à l'adoption et l'application de diverses tâches structurantes. Il s'agit d'une quête interne à Wikipédia, toujours à la recherche d'un équilibre entre son immense communauté et l'automatisation de sa complexe gestion du contenu (Niederer et van Dijck, 2010). Malgré ce renforcement continu du contrôle algorithmique, la critique d'une conception sociotechnique stricte de la communauté, notamment du « *code is law* » développé par Lessig (1999), reste possible. L'observation de l'échec temporaire de certains bots, tels ceux occupant des rôles de lutte au vandalisme (tel qu'il est observé dans notre dernière section), met en lumière les processus sociaux sous-jacents à leur utilisation. En effet, au début de 2011, *ClueBot NG* tombe hors service à quelques reprises, ce qui permet aux chercheurs de constater comment cette sous-communauté se réorganise spontanément Geiger (2014). « *Un tel moment de brisure révèle que le pouvoir algorithmique de Wikipédia est hautement distribué et décentralisé, en fort contraste d'un système de gestion plus intégré et institutionnalisé tel la YouTube's ContentID.* » (Geiger, 2014)²⁷. Pour certains, il est clair que les éditeurs n'auraient pu maintenir le rythme toutes ces années sans l'assistance de ses bots (Niederer et van Dijck, 2010), p10). Malgré tout, cela n'empêche pas certains auteurs de comparer Wikipédia à un « *Mechanical Turk qui tourne les actions et jugements des individus en des fonctions d'un programme informatique.* » (Carr, 2008)²⁸. Cette confiance accordée aux bots se traduit en l'abaissement de deux niveaux de supervision humaine : (1) de par leur droit d'éditer sans supervision, (2) tout en étant caché de l'historique (M. Livingstone, 2012). Les éditions réalisées par les bots y sont considérées par défaut « sur », incitant donc à abaisser les gardes fous en place dans l'encyclopédie.

Pour conclure, l'utilisation des bots permet d'autoriser l'ouverture de nouveaux espaces discursifs, autrement impossible à envisager, que cela soit par l'augmentation du temps de travail libre des individus qu'en créant des capacités d'actions dans de nouveaux domaines d'activités. Leur capacité d'action en sera d'autant plus élargie, réalisant désormais des tâches hier encore de trop grande ampleur. Toutefois, ceci va de pair avec une autre dynamique au niveau de la communauté : « *Parce que les bots sont construits pour reproduire algorithmiquement une vision particulière de ce qu'est ou devrait être les articles de l'encyclopédie ou des collaborations basées sur un wiki, ils ont une profonde influence sur ce qu'est Wikipédia et de comment elle fonctionne, autant en tant qu'encyclopédie qu'en tant que communauté.* » (Geiger, 2014)²⁹. Un des résultats évidents est donc la restructuration des jeux de pouvoir à l'interne du projet alors qu'aucun changement majeur du terrain initial (MediaWiki) n'est constaté. La redistribution du travail entre humains et bots vient aussi transformer l'organi-

27. Traduction libre.

28. Traduction libre.

29. Traduction libre.

sation de Wikipédia, puisqu'une couche logicielle valide automatiquement et semi automatiquement les contributions humaines (Geiger et Ribes, 2010). Avec l'apport des algorithmes de classement d'éditions, la distribution du travail s'opère selon une série de règles qui n'est pas exposée à la majorité des participants (Geiger et Ribes, 2010). Le cas de *ClueBot NG*, supervedette antivandalisme de l'encyclopédie anglaise, permet d'observer ces règles par l'algorithme que le bot utilise. Celui-ci est programmé pour discriminer les utilisateurs anonymes ainsi que les nouveaux utilisateurs. C'est pourquoi notre dernière section va s'intéresser à la lutte au vandalisme.

2.3 Démonstration par l'exemple : le cas de la protection contre le vandalisme

Afin d'explicitier les concepts venant d'être exposées, nous abordons ici la lutte contre le vandalisme, sujet particulièrement exploré par divers contributeurs et chercheurs de différents horizons. De prime abord, le vandalisme n'est pas un concept pouvant être facilement traduit en décisions algorithmiques. Ceci-dit, la recherche continue d'indices permettant d'identifier ce type de contributions malvenues nous expose la variabilité de formes que peut prendre le vandalisme (Velasco, 2010). Cela peut consister en des contributions vulgaires ou écrites en langage populaire, écrites à la première personne, une opinion biaisée, etc. Cela peut aussi être l'effacement complet du texte d'un article (*blancking*) ou de son remplacement par des déclarations violentes, sexuelles ou par du pur non-sens, etc. (Priedhorsky et collab., 2007; Velasco, 2010). Selon Wikipédia, le vandalisme est constitué de « *toute addition, suppression, ou changement fait au contenu dans une tentative délibérée de compromettre l'intégrité de Wikipédia* » (Gorgeon et Swanson, 2011)³⁰. À titre d'exemple, Gorgeon et Swanson (2011) analysent l'article « Web 2.0 ». Bien qu'il y soit retrouvé des débats d'interprétations sur ce terme, cet article risque peu de vivre de grands conflits. Une analyse des 5 970 éditions réalisées sur une période totalisant un peu moins de cinq ans identifie que le vandalisme et le spam y représentent pourtant 15,7 % des éditions totales (Gorgeon et Swanson, 2011). Malgré le relativement faible nombre de conflits au sein des articles (Yasseri et collab., 2012), cette omniprésence du petit vandalisme, c'est-à-dire le plus flagrant et facile à annuler, a forcé beaucoup de contributeurs à développer des techniques fines afin de l'enrayer.

La communauté anti-vandale est l'une des sous-communautés les mieux organisées de l'encyclopédie³¹. Grâce à la revue de littérature réalisée jusqu'à maintenant dans ce mémoire, il n'est pas étonnant de constater que cette communauté utilise fortement des bots et logiciels d'assistance à l'édition. La forme de ces outils demeure cependant particulière à chaque sous-communauté. « *Dans le contexte de la gouvernance algorithmique, les sondages, c'est-à-dire*

30. Traduction libre.

31. http://en.wikipedia.org/wiki/Wikipedia:Counter-Vandalism_Unit, consulté le 25 février 2014.

« *l'opinion de la communauté* », sont particulièrement informatif comme façon de déterminer le consensus en fonction de (1) l'implémentation de mécanismes algorithmiques ou (2) définir les critères pour les évaluer. » (Müller-Birn et collab., 2013)³². Cette construction sociale des normes est explicite entre les différentes versions de Wikipédia. Dans la communauté allemande, par exemple, l'utilisation de bots afin d'automatiser la lutte au vandalisme y est interdite puisque cette tâche est perçue comme de la gestion d'exceptions. En ce sens, le jugement humain ne peut pas y être mis de côté (Müller-Birn et collab., 2013). La conception différenciée de la catégorie « vandalisme » entre les versions impose donc une réponse différente selon les communautés.

On peut donc comprendre la lutte au vandalisme au travers du concept de cognition partagé développé par Latour (1992). En ce sens, la protection contre le vandalisme est un processus complexe impliquant un nombre étonnant d'acteurs humains et non-humains (Geiger et collab., 2013). Les bots, de par leur rapidité, forment la première barrière de protection, annulant les plus évidentes de ces éditions en quelques secondes. Par exemple, les auteurs de ClueBot NG affirment que celui-ci est apte à identifier 55 % du vandalisme total de l'encyclopédie³³, et ce, pratiquement en temps réel (Geiger et collab., 2013). En fait, c'est même 13,7 % du nombre total d'annulations dans l'encyclopédie qui est réalisé par ce bot au cours de janvier 2011 (Geiger, 2014). C'est donc « *beaucoup du travail de patrouille des nouvelles contributions [...] qui est performé non par un humain lisant des articles complets de l'encyclopédie, mais au contraire par des bots automatiques hautement sophistiqués, qui annule le vandalisme, les pourriels, et autres contributions malicieuses.* » (Geiger, 2009)³⁴. Suite à ce premier tri, nous retrouvons les éditeurs assistés par logiciels. Les deux premiers outils utilisés par ce groupe ont été *AntiVandalBot* et *VandalProof* (Halfaker et Riedl, 2012). *AntivandalBot*, par l'application d'une série de règles et heuristiques, permet l'identification des cas de vandalisme les plus flagrants afin de restaurer les articles touchés. *AntiVandalproof* agit différemment, s'agissant d'une interface graphique facilitant la prise de décision. Le logiciel d'assistance à l'édition *Huggle* en est la version contemporaine, contribuant à la lutte au vandalisme en facilitant énormément l'utilisation des « *roll-back* ». Cette action consiste en l'annulation simultanée de multiples éditions afin de revenir à une version « stable » d'un article (Halfaker et Riedl, 2012). Bien qu'il n'est pas directement question de ces logiciels dans ce mémoire, leur utilisation n'est pas anecdotique. Par exemple, en 2009, 12 % des éditions de l'encyclopédie sont réalisés grâce à ces logiciels (Geiger et Ribes, 2010). L'utilisation des logiciels d'assistance à l'édition permet d'alléger les utilisateurs de la plupart des calculs, étant face à un classement visuel des changements soupçonnés de vandalismes, suggérant déjà un jugement que l'éditeur valide manuellement. L'aide de la communauté est d'autant plus aisée qu'avec des logiciels comme *Huggle* et *Twinkle*, des individus avec peu d'expériences ou de connaissances dans le

32. Traduction libre.

33. https://en.wikipedia.org/wiki/User:ClueBot_NG, consulté le 3 septembre 2014.

34. Traduction libre.

domaine visé peuvent facilement participer à la lutte (abaissant *de facto* le coût de participation). Cette facilité à modifier massivement l'encyclopédie porte Gelley et Suel (2013) à affirmer que l'utilisation de ces outils peut mener à des suppressions exagérées (*over-zealous*) de contributions. La résolution de ce qui reste de vandalisme, pour ce qui est effectivement identifié, vient du travail éditorial direct ainsi que de discussions menant à un accord entre les partis. Dans un article subséquent, Geiger et Ribes (2010) ajoutent une quatrième catégorie, soit les « *batches scripts* ». Ce dernier acteur ajoute une couche logicielle supplémentaire après coup afin de formaliser et standardiser des modifications dont la finalité n'est pas directement la lutte au vandalisme.

Lors de la même étude, Geiger et Ribes (2010) observent trois périodes de temps durant lesquelles *ClueBot NG* est tombé inactif. N'étant pas maintenu directement dans les serveurs de MediaWiki, il est difficile pour les utilisateurs de connaître le moment où celui-ci redeviendra fonctionnel. S'il est encore possible pour les participants de reprendre à part entière le travail des bots, il est évident qu'ils ne sont pas capables de le faire avec la même efficacité et rapidité. Alors que plusieurs auteurs affirment que l'ajout de ces bots permet de rendre résilient ce projet qui ne le serait pas autrement (Niederer et van Dijck, 2010), Geiger et Ribes (2010) affirment constater l'inverse puisqu'il y a maintient de ces fonctions lors de l'échec de l'un de ses plus importants acteurs. Toutefois, relativisant leur déclaration, ils comparent la chute de *ClueBot NG* à l'avènement d'une catastrophe naturelle : il est peut-être possible de maintenir la cadence pendant un certain temps après l'événement, mais est-ce envisageable à long terme? (Geiger et Ribes, 2010). Nous supposons que cela est possible tant que cette crise reste limitée dans l'espace et dans le temps.

Chapitre 3

Hypothèses et objectifs

3.1 Question de recherche

Ce projet de recherche vise principalement à décrire les interactions entre les utilisateurs de Wikipédia et les agents logiciels agissant en son sein, en prenant comme modèle la version anglaise de l'encyclopédie collaborative. Ce projet utilise les concepts développés dans notre rappel de la littérature, incluant les communautés de pratique et collaboratives, la structuration sociotechnique des normes ainsi que l'interaction éditeurs / agents logiciels. Notre recherche, à la fois quantitative et qualitative, vise à comprendre les représentations que les utilisateurs ont du rôle des agents logiciels au sein de l'encyclopédie. Notre question de recherche se présente ainsi : *Comment les contributeurs perçoivent-ils l'action des agents logiciels sur leur propre travail et, dans un second temps, comment le travail des agents logiciels s'intègre-t-il dans le fonctionnement et la structuration de la communauté ainsi que dans le contenu de Wikipédia?*

3.2 Objectifs de recherche

Afin de répondre à notre question de recherche, nous aborderons les objectifs spécifiques suivants :

Objectif spécifique 1 Quantifier le travail des agents logiciels et l'identification de leur distribution en différentes classifications catégorielles selon les fonctions que les agents logiciels réalisent.

Objectif spécifique 2 Quantifier les interactions qu'ont les contributeurs de Wikipédia avec les agents logiciels et leurs propriétaires sur les pages de discussions associées aux bots de l'encyclopédie.

Objectif spécifique 3 Identifier si les interactions qu'ont les utilisateurs avec les agents logiciels sont différenciées selon certaines caractéristiques dichotomiques des agents logiciels.

Objectif spécifique 4 Au sein des discussions enclenchées par les utilisateurs sur les pages de discussions associées aux agents logiciels, identifier si on y retrouve des types de réaction différentes selon les caractéristiques dichotomiques des agents logiciels.

Objectif spécifique 5 Identifier le rôle perçu des agents logiciels, tant chez les éditeurs de Wikipédia que chez les propriétaires des agents logiciels.

3.3 Hypothèses de recherche

En ce qui concerne nos hypothèses de recherche, nous avons prévu au départ de l'étude que :

1. Les agents logiciels continuent d'augmenter leur importance au sein de l'encyclopédie.
2. Les agents logiciels représentent un outil important pour leur propriétaire leur permettant d'alléger le poids des tâches nécessaires qu'ils ont à réaliser pour l'encyclopédie.
3. Parmi les différentes classifications catégorielles, le mode d'application des agents logiciels, allant de l'application automatique jusqu'aux agents logiciels fonctionnant uniquement par abonnement, va influencer l'appréciation que les éditeurs auront de ceux-ci.
4. Les agents logiciels occupant des tâches touchant l'imposition et le respect des normes de l'encyclopédie seront perçus négativement par la communauté, puisqu'ils participent activement à la stabilisation dans le contenu de Wikipédia et donc, à une perte d'autonomie des utilisateurs en général.
5. Dans Wikipédia, les agents logiciels représentent la formalisation de sa communauté originellement informelle, en traduisant les normes sociales qu'elle adopte en politiques et règles de conduite, qui sont ensuite appliquées à même l'infrastructure de Wikipédia.

Chapitre 4

Méthodes

4.1 Protocole de recherche

Nous réalisons cette étude dans l’objectif d’éclaircir les phénomènes cognitifs entourant l’interaction entre une partie des contributeurs à Wikipédia et les agents logiciels. Bien que nous réalisons cette recherche dans le but d’obtenir des résultats généralisables à l’ensemble des contributeurs de l’encyclopédie, nos analyses seront de fait limitées aux contributeurs ayant interagi avec les agents logiciels sur les pages de discussion associées à ces derniers. Cette considération restera présente lors de notre discussion. Plutôt de valider une théorie préalablement identifiée par la littérature, nous avons réalisé une série d’itérations dans notre traitement des données disponibles pour nos recherches. La méthode de recherche que nous utiliserons correspond donc sur bien des points à la théorisation ancrée. Notamment, cette méthode de recherche procède d’un aller-retour itératif entre approche inductive et déductive (Glaser et Strauss, 1967). Ceci permet de s’assurer que les interprétations que nous réaliserons à partir de nos données soient bien ancrées avec la situation concrète de notre terrain. Cette méthode nous semblait particulièrement à propos pour le champ de recherche peu exploré à ce jour qu’est l’interaction *in situ* de certains des contributeurs de Wikipédia avec les agents logiciels. Nos résultats contribuent à mieux définir les représentations que possèdent les contributeurs à Wikipédia face aux agents logiciels et les tâches qu’ils réalisent dans l’encyclopédie.

Les agents logiciels, appelés bots, sont généralement décrits comme étant des programmes réalisant des tâches déterminées par algorithmes, incluant l’édition et la maintenance d’articles, et sont aussi interprétés comme des outils ou protocoles managériaux touchant l’infrastructure d’un projet (Geiger et Ribes, 2010 ; Müller-Birn et collab., 2013 ; Niederer et van Dijk, 2010). Le propriétaire d’un agent logiciel particulier est l’utilisateur enregistré de Wikipédia qui est responsable des actions de l’agent logiciel. Chacun des agents logiciels possède une page de discussion, qui est une page d’arrière-plan spécialement dédiée à des discussions avec ou à propos du contributeur, qui est ici un agent logiciel (voir figure 4.1 et 4.2 à la

page 52). Dans certains cas, la page de discussion des agents logiciels est fusionnée avec celle du propriétaire (souvent par redirection du lien). Nous appelons "discussions" l'ensemble des commentaires regroupés sous un même titre, que cela soit un seul ou de nombreux commentaires. Une édition consiste en une modification entreprise par un utilisateur ou un agent logiciel dans l'espace principal de Wikipédia, c'est-à-dire essentiellement sur les articles, mais cela inclut également n'importe quelle modification entreprise par un agent logiciel dans le cadre de ses tâches. Un commentaire, pour sa part, fait référence à une édition réalisée sur une page de discussion par les utilisateurs de Wikipédia. Ces discussions sont divisées par des marqueurs visibles sur les pages de discussion. La succession directe de commentaires au sein de ces marqueurs forme une discussion, même si celle-ci peut n'être qu'un unique commentaire sans réponse. La communauté interne de Wikipédia regroupe les utilisateurs qui éditent avec une certaine régularité et qui sont, au moins partiellement, à l'affût des débats internes ayant lieu dans le projet collaboratif, y participant eux-mêmes à l'occasion. Enfin, un veilleur de pages est un membre de la communauté interne de Wikipédia qui surveille des pages de discussion particulières¹. Les veilleurs de pages s'abonnent à des articles et à leur page de discussion et reçoivent des notifications lorsque des activités ont lieu sur des sujets qui leur sont d'intérêt. Dans le cadre de notre étude, il s'agit de pages de discussion reliées aux agents logiciels. Ces veilleurs de pages agissent en répondant ou en alimentant de faits particuliers une discussion sur ces pages, discussion qu'ils intègrent même s'il n'y sont pas initialement impliqués.

FIGURE 4.1 – Une page personnelle d'un bot, ici Xqbot.

The screenshot shows the Wikipedia user page for 'User:Xqbot'. At the top, there are links for 'User page', 'Talk', 'Read', 'Edit', and 'View history', along with a search bar. The main heading is 'User:Xqbot' with the subtext 'From Wikipedia, the free encyclopedia'. A central box contains the following text: 'This user account is a bot that uses pywikipedia framework, operated by Xqt (talk). It is a legitimate alternative account, used to make repetitive automated or semi-automated edits that would be extremely tedious to do manually. The bot is active, and, by virtue of the fact that it is flagged globally, is permitted to maintain interwiki links per the global rights policy. In addition, its local approval can be viewed here.' To the right of this box is a table titled 'Xqbot' with the following data:

Xqbot	
This user is a bot (talk · contribs)	
Operator	Xqt
Approved?	Yes
Flagged?	Yes
Task(s)	<ul style="list-style-type: none"> • solve double redirects • fix interwiki • add missing references tag • maintaining featured articles
Edit rate	up to 12 edits p.m.
Edit period(s)	daily
Automatic or manual?	Automatic
Programming language(s)	Python (pywikipedia framework)
Exclusion compliant?	Yes
Source code published?	Yes [1]

1. Dans la version anglaise de Wikipédia, le veilleur de pages est nommé un « *talk page stalker* ».

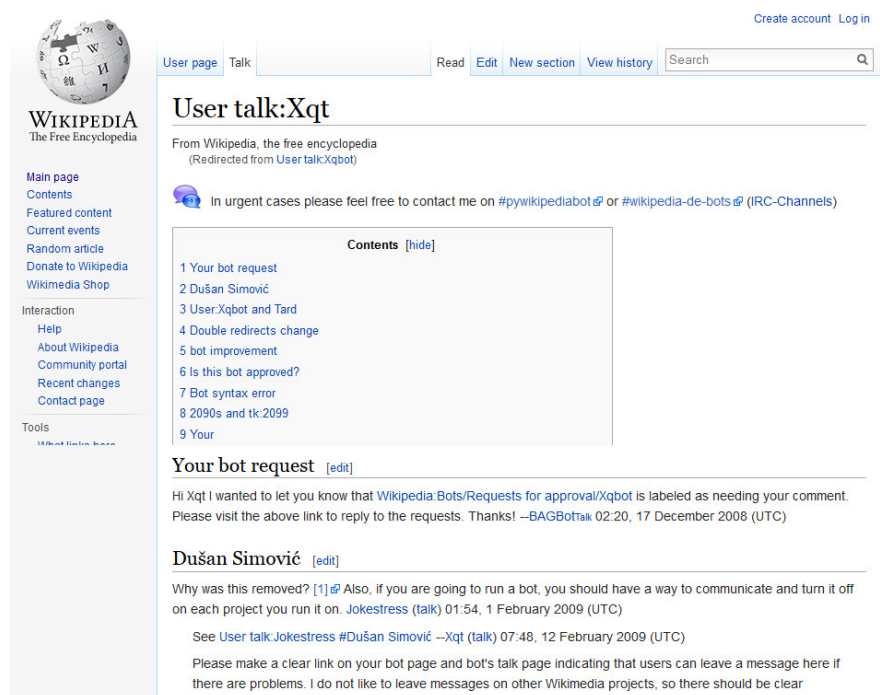


FIGURE 4.2 – La page de discussion associée à la page personnelle.

4.2 Sélection des bots

Afin de répondre à notre premier objectif de recherche, soit la quantification du travail des agents logiciels et l'identification de leur distribution en différentes classifications catégorielles selon les fonctions que les agents logiciels réalisent, nous avons évalué les agents logiciels actifs de la version anglaise de l'encyclopédie. Ceci est possible grâce aux messages des contributeurs laissés sur les pages de discussion des agents logiciels. Puisque chaque langue spécifique correspond à une version unique de Wikipédia, nous retrouvons une communauté différente pour chacune d'entre elles. Les normes exactes ainsi que les méthodes utilisées par la communauté sont sensiblement différentes d'une langue à une autre. Nous avons donc concentré nos efforts sur la principale et plus imposante communauté de Wikipédia, celle de langue anglaise. Seuls les bots actifs sur les pages du Wikipédia anglais entre le 4 janvier 2012 et le 2 janvier 2013 ont été considérés. Les données d'activités des bots ont été accumulées à partir de la page « bots selon le nombre d'édits » du Wikipédia anglais², qui est maintenu automatiquement à jour grâce à l'agent logiciel BernsteinBot. Celui-ci utilise un décompte d'édits extrait de Toolserver (une suite logicielle spécifiquement utilisée par les éditeurs de Wikipédia)³. L'activité des bots a été mesurée en calculant la différence en éditions d'un bot particulier entre le 2 janvier 2013 et le 4 janvier 2012. Ce total comprend les bots sans « *bot flag* » ainsi que les éditions supprimées. Un total de 170 bots actifs a ainsi été identifié. À

2. http://en.wikipedia.org/wiki/Wikipedia:List_of_bots_by_number_of_edits, consulté le 4 novembre 2014.

3. <http://toolserver.org/>, consulté le 4 novembre 2014.

des fins d'analyse, nous avons extrait l'ensemble des discussions des pages étant directement reliées à la page de discussion de ces bots. Ce fut le cas pour les pages dont le but est de permettre aux utilisateurs de laisser un commentaire, un avis, etc. touchant un bot, incluant la propre page de discussion du propriétaire, lorsque pertinent. Ensuite, nous avons calculé l'âge d'activité du bot en soustrayant le jour de création de sa page personnelle à la dernière journée de notre étude, soit le 2 janvier 2013. De plus, nous avons aussi calculé le nombre de tâches réalisées par le bot en additionnant le nombre de tâches acceptées par le *Bot Approval Group* au cours de l'histoire du bot. De ce nombre, nous avons soustrait les tâches obsolètes et retirées, telles qu'identifiées sur la page personnelle du bot, afin d'avoir le nombre de tâches qui étaient effectives durant la période de collecte.

Seulement les bots ayant réalisé au minimum 1000 éditions et ayant au minimum 12 commentaires d'utilisateurs sur leur page de discussion au cours de la période à l'étude ont été considérés. Le premier seuil nous permet d'éviter d'analyser des bots n'ayant pas un niveau suffisant d'activité pour déclencher une réaction minimale des contributeurs (seuil utilisé pour les statistiques internes de l'encyclopédie⁴). Le deuxième seuil représente simplement la présence d'au moins une interaction de contributeur par mois pour un bot en particulier. Les bots dont la plupart des contributions ont été réalisées dans une autre version que l'encyclopédie anglaise, ou dont les contributions et interactions ont été accomplies dans une autre langue que l'anglais, ont été exclus. Quand plusieurs bots partagent la même page de discussion, seul le bot avec le plus haut nombre d'édits a été considéré, ce qui totalise de 51 bots. Toutefois, l'un d'entre eux (Yobot, classé premier en terme d'édits pendant notre période d'observation) s'est retrouvé au centre de conflits répétés, bloquant temporairement à au moins deux reprises au cours de la période d'observation. À cause d'un problème d'interprétation des tâches du bot, les activités de ce dernier ont résulté en un nombre très important d'édits fortement, très souvent et largement contestées par la communauté en termes de pertinence et de légitimité. Pour cette raison, la page de discussion de ce bot ne peut être considérée comme un modèle standard des réactions des utilisateurs quant à l'usage des bots. Ce bot a donc été exclu du traitement de données quantitatives de cette étude. Toutefois, lors de l'interprétation de nos résultats, le cas Yobot nous permettra d'exposer certaines attentes des contributeurs de l'encyclopédie envers les agents logiciels. Au final, un corpus de 50 bots compose cette étude.

Après avoir comptabilisé le nombre d'agents logiciels dans le cadre de notre premier objectif spécifique, nous avons mesuré le nombre d'édits réalisé par ceux-ci ainsi que le nombre de tâches accomplies (incluant la moyenne \pm ESM, la médiane, le maximum, le minimum), autant dans notre échantillon que dans la population totale d'agents logiciels. Grâce aux mêmes statistiques, nous avons comptabilisé le nombre global de commentaires et de discussions,

4. <https://stats.wikipedia.org/EN/TablesWikipediaEN.htm>, consulté le 5 novembre 2014.

tant pour les commentaires d'utilisateurs, que pour les propriétaires d'agents logiciels et les veilleurs de page.

4.3 Caractéristiques des bots

Toujours dans le but de répondre à notre premier objectif de recherche, les bots ont été classés selon plusieurs catégories de caractéristiques dichotomiques. Nous avons identifié ces catégories par le biais des principaux éléments comptabilisés lors de l'acceptation des bots⁵. La première caractéristique d'un bot est la nature de son propriétaire (soit un administrateur de Wikipédia « admin » ou un utilisateur enregistré sans pouvoirs d'administrateur). La seconde caractéristique concerne le fonctionnement du bot, soit en étant principalement contraignant dans l'action des autres utilisateurs (protection contre le vandalisme, surveillance de contenu sous droits d'auteur, avertissement des utilisateurs) ou en facilitant leur travail (substitutions de gabarits, correction de redirection de liens ou de liens inactifs). La troisième caractéristique est la localisation des éditions du bot, qui peut être soit en avant-plan (dans un article ou une page d'utilisateur, incluant aussi leurs pages de discussion respectives) ou en arrière-plan (tels les catégories, les portails, les pages d'aide, les gabarits). La quatrième caractéristique consiste en un bot agissant soit automatiquement (fonctionnant selon une politique « opt-out »), soit seulement par inscription « opt-in ». Finalement, une cinquième caractéristique consiste en une estimation globale du degré de complexité des informations que le bot doit analyser avant de réaliser une édition, qualifié soit de haute (protection du vandalisme, validation des droits d'auteurs, suggestion de pages à éditer, etc.) ou de basse (édition purement factuelle ou managériale, telles la correction de faux liens ou la distribution des courriels parmi les utilisateurs). Cette dernière classification a été réalisée grâce à la lecture des demandes d'autorisations auprès du BAG⁶.

Pour déterminer l'appartenance des agents logiciels à une catégorie particulière, nous avons observé pour chacun la page retraçant leur processus d'acceptation. Quand les bots réalisaient de multiples éditions de natures différentes, la sélection des caractéristiques a été définie par le type d'édition principalement commenté par les utilisateurs. Nous répondons à cet objectif spécifique en réalisant un décompte des bots au sein de leurs différentes catégories respectives.

4.4 Traitement des commentaires et des discussions

Une première phase de traitement des données a été réalisée sur le nombre total de commentaires au sein de la page de discussion d'un bot. Notre deuxième objectif spécifique est de quantifier les interactions qu'ont les contributeurs de Wikipédia avec les agents logiciels et leurs propriétaires sur les pages de discussion associées aux bots de l'encyclopédie. Lorsque

5. https://en.wikipedia.org/wiki/Wikipedia:Bots/Requests_for_approval, consulté le 3 juin 2014.

6. https://en.wikipedia.org/wiki/Wikipedia:Bots/Requests_for_approval, consulté le 3 juin 2014.

les discussions n'ont pas eu lieu directement sur la page de discussion du bot, mais sur celle de son propriétaire, l'ensemble des commentaires qui n'étaient pas directement reliés au bot a été exclu. Si les commentaires ont été réalisés dans une autre langue que l'anglais ou ont été dédiés par erreur à un autre bot, ils ont aussi été exclus. Les réponses des propriétaires des bots ainsi que des veilleurs de pages ont été codées séparément. Dans ce cadre, nous avons comptabilisé, pour chacune des pages de discussion associées aux agents logiciels, les discussions exclues, les commentaires d'utilisateurs, de propriétaires et de veilleurs de pages, par bots et par discussions (incluant la moyenne \pm ESM).

Par la suite, nous avons mesuré si certaines fonctions réalisées par les bots provoquaient des réactions différenciées chez les utilisateurs, ce qui est notre troisième objectif spécifique. Après y avoir quantifié la distribution générale des commentaires au sein de notre échantillon, une deuxième phase a permis de qualifier le ton des commentaires laissés par les contributeurs sur les pages de discussion des bots. Ces commentaires ont été classés entre *positif*, *neutre* et *néгатif*. Un test intra-juge réalisé sur 10,4% de l'échantillon permet d'observer une précision de 95,8% dans le codage de ces commentaires. Ces différents éléments nous permettent d'observer s'il y a des différences observables sur les pages des agents logiciels quant à l'interaction que les contributeurs ont avec eux. Pour chaque page de discussion de notre échantillon, nous avons comparé le ton des commentaires aux différentes caractéristiques dichotomiques des agents logiciels (incluant la moyenne \pm ESM).

Dans le cadre de notre quatrième objectif, une seconde catégorie de traitement de nos données a été réalisée en utilisant seulement le premier commentaire de chaque discussion, tel que divisé au sein des pages de discussion des bots. Spécifiquement, nous avons traité nos données en utilisant une méthode qualitative d'analyse de contenu, soit l'analyse structurale (Remy et Ruquoy, 1936). Cette méthode vise à « repérer les représentations mentales au travers desquelles l'acteur social définit son environnement et se donne une identité » (Remy et Ruquoy, 1936). Les discussions sont donc caractérisées en fonction de l'intention perçue par rapport aux commentaires initiaux. Un codage progressif a permis à l'auteur de résumer à neuf (9) les types de réaction des utilisateurs. Ceux-ci peuvent représenter des intentions positives avec, 1) des remerciements quant au travail du bot, 2) des demandes d'aide ou de conseils, et 3) des utilisateurs informant le bot d'actions qu'ils entreprennent par eux-mêmes. Elles peuvent aussi représenter des intentions neutres avec, 4) la détection d'erreurs ou de faux positifs dans le travail du bot, 5) des suggestions d'améliorations, et 6) le questionnement d'une édition ou d'une fonction du bot). Enfin, elles peuvent représenter des intentions négatives avec, 7) des critiques générales, 8) la contestation d'une édition spécifique, et 9) des utilisateurs venant justifier leurs actions. Des exemples pour chacun de ces neuf types de réaction sont fournis dans la table 4.3. Un test intra-juge réalisé sur 10,4% de l'échantillon permet d'observer une précision de 91,2% dans le codage de ces discussions. À cette fin, nous avons comptabilisé

FIGURE 4.3 – Exemples de réactions d'utilisateurs pour les neuf types de réaction observés.

Types de réaction	Exemple	Bot en question
1. Remerciements	Good catch you silly robot C6541 (T C) 21 :12, 27 November 2012 (UTC)	28Bot
2. Demandes d'aide ou conseils	Hi, I'd need the script you used to leave a notice to all the users which participated to the SOPA's discussion, Basically I have to do almost the same for the italian one. Thank you! – Vituzzu (talk) 12 :22, 17 January 2012 (UTC)	Thehelpfulbot
3. Utilisateur informant d'une action	Hi, I think that that bug in your bot is actually Py-wikimedia bug (you are using this bot, right?). So, I posted the bug here : https://sourceforge.net/[...] Safinaskar (talk) 10 :09, 7 June 2012 (UTC)	AvicBot
4. Détection d'erreurs ou de faux-positifs	It appears that RFC bot (talk · contribs) is maintaining redlinks at Wikipedia :Dashboard/Requests for comment again. (See diff) Please check to see what is causing it and if you can correct the problem. – Allen4names 15 :17, 25 May 2012 (UTC)	RFC Bot
5. Suggestions d'améliorations	Regarding this diff, it would be possible to reduce the number of similar false positives by searching matched content for 'Wikipedia'. — C M B J 05 :52, 14 December 2012 (UTC)	MadmanBot
6. Questions sur bot ou ses éditions	should this bot revert deletion of copyvio text at Lucy Morton including addition of a cospypaste template? — Preceding unsigned comment added by 82.132.232.222 (talk) 16 :15, 30 July 2012 (UTC)	ClueBot NG
7. Critiques générales	Why are you spamming this nonsense to hundreds/thousands of people? Stop doing that. I didn't sign up for this crap, and I'm sure most/all others didn't either. Dream Focus 19 :54, 4 September 2012 (UTC)	EdwardsBot
8. Contestations spécifiques	Hello, there was a reason for the space. With it there Ivanhoe sorts before Ivanhoe East, which would seem to be the normal sorting convention; looks ugly the other way. Regards Crusoe8181 (talk) 07 :57, 27 September 2012 (UTC)	BG19bot
9. Justifications d'un utilisateur	I have check on this page article and in fact it is real see on this web site Official website I will add more source on this. Thank you Errorzerol (talk) 07 :25, 1 September 2012 (UTC)	AnomieBot

le ton du premier commentaire des discussions ainsi que le type de réaction des utilisateurs (moyenne \pm ESM).

4.5 Constitution d'idéaux types de bots

Afin d'identifier le rôle perçu des agents logiciels, autant chez les éditeurs de Wikipédia que pour les opérateurs (cinquième objectif spécifique), nous avons utilisé deux types de traitement de données. Dans le premier, nous avons procédé à l'identification de deux idéaux types d'agents logiciels. Dans la section 4.3, nous avons identifié 5 différentes caractéristiques des bots, soit la propriété, le fonctionnement du bot, la localisation de ses éditions, son mode d'applications et le degré de complexité des décisions. Bien que cette classification apporte des informations fortement utiles, nous avons noté que les agents logiciels possédaient des appartenances spécifiques avec ces cinq caractéristiques. Nous avons retenu deux groupes de caractéristiques récurrents, représentant chacun plus de 15 % de notre échantillon. À partir de cette constatation, nous avons construit deux stéréotypes opposés, présentant des schémas comportementaux identifiables et spécifiques. D'un côté, certains bots réalisent des tâches laborieuses et répétitives à la place d'utilisateurs humains. Ces bots facilitent le travail des utilisateurs et agissent principalement en arrière-plan, automatiquement « *opt-out* » ainsi qu'avec une basse complexité de l'évaluation. Nous les avons nommés « bots servants ». D'un autre côté, certains bots agissent proactivement afin de faire respecter et appliquer les normes et règles de conduite de Wikipédia. Ils sont contrôlés par des administrateurs, contraignent le travail des utilisateurs, agissent en avant-plan, automatiquement ainsi qu'avec une haute complexité de l'évaluation. Nous les avons nommés « bots policiers ». Ces deux types de schémas comportementaux globaux représentent, dans nos résultats, des idéaux types distincts de fonctions, portant la communauté à avoir des interactions fortement distinctes entre les deux groupes. Ces idéaux types ont été comparés aux autres variables déjà exposés.

4.6 Traitement du « panneau d'affichage des propriétaires de bots »

Pour compléter de répondre à notre cinquième objectif spécifique, nous avons, dans un deuxième temps, réalisé un traitement des données présentes dans le « Panneau d'affichage des propriétaires de bots »⁷. Il s'agit d'une page d'arrière-plan dédiée aux discussions et à la coordination entre propriétaires de bots ainsi que d'un endroit de discussion pour la communauté, avec les propriétaires, permettant de traiter des questions importantes touchant les bots et leur travail. Spécifiquement, il s'agit de discussions ayant eu lieu entre le 10 janvier 2012 et le 27 décembre 2012. En vue du traitement de ces données, nous avons classé les interventions en quatre catégories distinctes : celles touchant la légitimité des bots et de leurs tâches, celles

7. http://en.wikipedia.org/wiki/Wikipedia:Bot_owners'_noticeboard, consulté le 24 août 2014.

touchant les processus (la façon concrète de réaliser une tâche), celles touchant des aspects strictement techniques (tels la programmation spécifique d'une tâche ou un questionnement sur la faisabilité d'une autre) ainsi que les discussions touchant directement le maintien des capacités techniques apportées par des bots (souvent dans le cas d'un remplacement d'un bot ou de son propriétaire). Un type supplémentaire d'interventions a été observé avec des discussions portant sur la communauté propre de Wikipédia. Celui-ci regroupe les discussions sur les activités de la communauté interne de Wikipédia ainsi que de la sous-communauté des propriétaires des bots. Bien que traitées, ces discussions ne sont donc pas directement reliées aux tâches et à la gestion des bots. Cette deuxième approche entreprise afin de répondre à notre cinquième objectif spécifique nous permet d'approfondir notre compréhension des phénomènes sociaux sous-jacents à l'utilisation des bots, notamment dans la coordination et l'esprit d'association présents chez les propriétaires de bots.

De ce « Panneau d'affichage des propriétaires de bots », nous avons manuellement codé chacun des contributeurs afin d'identifier sa position sociale dans Wikipédia. Nous avons mesuré si celui-ci est propriétaire d'un ou de plusieurs bots, s'il est administrateur, s'il est membre du BAG et/ou s'il est un contributeur de l'encyclopédie sans être membre d'aucune des catégories précédemment mentionnées. Lorsque le contributeur possède plusieurs rôles à la fois, nous avons codé séparément chacun de ces rôles. De plus, nous y avons comptabilisé le nombre de discussions et le nombre total de commentaires en fonction des types d'interventions.

4.7 Analyse statistique

Les données ont été codées et classifiées en utilisant le logiciel ouvert RQDA (la version 0.2-3 avec la version 2.15.2 du logiciel libre R), qui insère les données dans une base de données SQLite. Le codage a été réalisé par un seul observateur, aveugle aux hypothèses, sur une période de 40 jours. Un test intra-juge a été réalisé sur 10,4% de l'échantillon afin de valider le codage des différents éléments de l'étude. Les analyses statistiques ont été réalisées en utilisant les logiciels SPSS (version 20, IBM, Armonk), SigmaPlot (version 11, Systat Software, Chicago) et Origin (version 8.1, OriginLab, Northampton). Puisque la distribution de nos données (que cela soit le nombre de commentaires, le nombre de discussions, le nombre d'éditions des bots, etc.) respectait des lois de puissance, l'ensemble de nos données a été sujet à une transformation logarithmique avant les analyses statistiques. Cette manière de faire se retrouve régulièrement dans les études traitant de Wikipédia (Ciampaglia, 2011; Muchnik et collab., 2013; Nazir et Takeda, 2008; Voss, 2005). Toutes les analyses ont donc été réalisées à partir des données avec transformation logarithmique et non sur les données brutes. Tel qu'indiqué dans nos résultats, le test non paramétrique Mann-Whitney U (Armitage et Berry, 1994) a été utilisé pour répondre à la majorité de nos objectifs. Particulièrement, nous avons utilisé ce test afin de comparer la distribution des commentaires, des discussions, des tons

des commentaires et des types de réaction des utilisateurs. Des régressions linéaires simples ont été effectuées sur les neuf types de réaction des utilisateurs afin de les classer entre « positifs », « neutres » et « négatifs », selon la plus haute association retrouvée avec le premier commentaire de chaque discussion. Pour chacune des 5 caractéristiques de bots ainsi que pour les deux idéaux types, la répartition des neuf types de réaction des utilisateurs a été comparée en utilisant le test non paramétrique sans assumption de distribution Kolmogorov-Smirnov. Lorsque possible, les résultats sont présentés selon la moyenne \pm ESM.

Enfin, nous avons réalisé un traitement complémentaire de nos données en utilisant une ANOVA afin de valider nos résultats. En effet, la distribution de nos données nous amène à favoriser les analyses non paramétriques. Toutefois, la taille de notre échantillon (>30) et la nature indépendante de nos échantillons nous permettent d'assumer la normalité de nos données, et par conséquent, l'utilisation d'ANOVA. Ces traitements complémentaires de nos données se regroupent en deux ensembles. En premier lieu, nous avons analysé nos cinq caractéristiques dichotomiques selon les neuf types de réaction des utilisateurs à l'aide d'une ANOVA à un facteur. Ensuite, nous avons réalisé une MANOVA univariée avec deux covariables, soit la caractéristique touchant la façon dont le bot réalisera ses tâches (automatiquement ou par inscription) et le lieu dans lequel se réaliseront les discussions (soit sur la page de discussion du bot, du propriétaire ou des deux à la fois). Afin de mesurer les différences spécifiques touchant les neuf types de réaction par les utilisateurs, une analyse ANOVA à un facteur, accompagnée d'un test post-hoc de Tukey, ont été utilisés afin de comparer les différences significatives spécifiques de nos deux idéaux types et des bots ne faisant partie d'aucun des deux idéaux types.

Chapitre 5

Résultats

5.1 Caractéristiques de notre échantillon

Dans un premier temps, nous avons voulu quantifier notre échantillon. Un total de 872 bots ont été référencés au sein de la version anglaise de Wikipédia le 2 janvier 2013. Entre le 4 janvier 2012 et le 2 janvier 2013, 170 bots étaient actifs, représentant un total de 8 985 353 éditions (52 855 \pm 10 437 éditions par bot, médiane : 10 037, incluant un maximum de 975

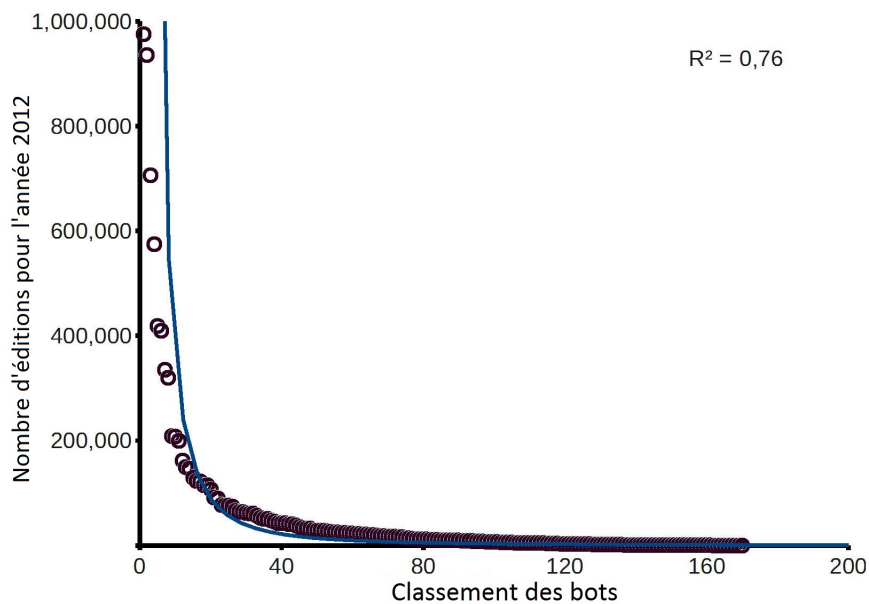
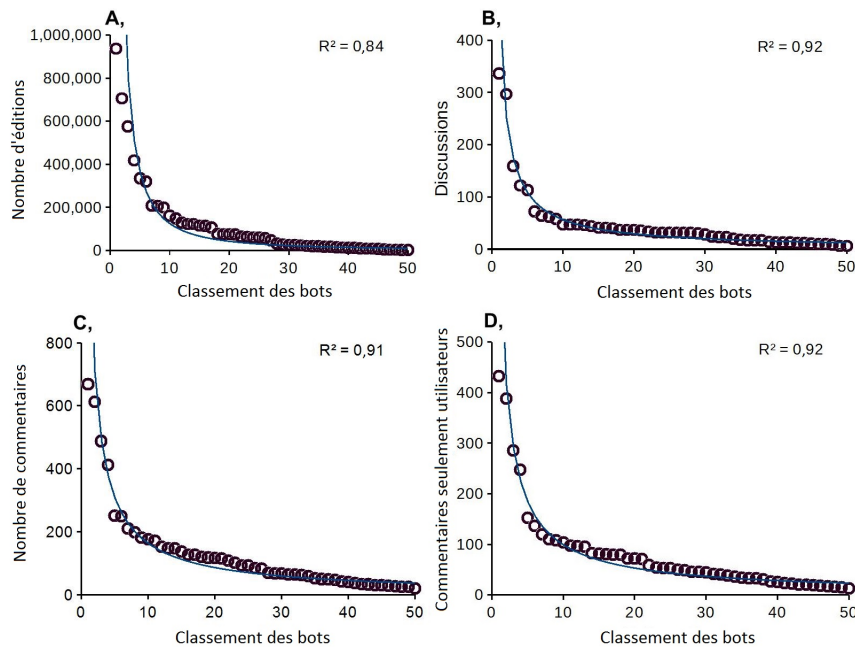


FIGURE 5.1 – **Activité de la population totale des bots de la version anglaise de Wikipédia.** Distribution du nombre d'éditions en fonction du rang du bot (ordonné selon le nombre décroissant d'éditions réalisées par celui-ci) pour les 170 bots actifs durant la période d'observation (du 4 janvier 2012 au 2 janvier 2013) sur le Wikipédia de langue anglaise. Cette distribution suit une loi de puissance ($R^2=0,76$, $p<0,001$).

FIGURE 5.2 – **Activité de l'échantillon de bots étudié.** Indiquées sont diverses mesures de l'activité pour les 50 bots de l'échantillon étudié durant notre période d'observation. Le classement des bots est réalisé selon le nombre décroissant d'éditations comptabilisées sur l'axe des ordonnées pour chacun des graphiques présentés. Notez que les différentes distributions observées suivent des lois de puissance. **A.** Pour la distribution du nombre d'éditions selon le classement du bot ($R^2=0,84$, $p<0,001$). **B.** Pour la distribution du nombre de discussions sur la page de discussion du bot en fonction du classement du bot. ($R^2=0,92$, $p<0,001$). **C.** Pour la distribution du nombre total d'interactions sur la page de discussion d'un bot en fonction du classement du bot ($R^2=0,91$, $p<0,001$). **D.** Pour la distribution du nombre d'interactions sur la page de discussion du bot, en excluant les commentaires des propriétaires et des veilleurs de pages, en fonction du classement du bot ($R^2=0,92$, $p<0,001$).



675 et un minimum de 20). La distribution de leurs contributions suivait de près une loi de puissance ($r^2 = 0,759$, $p<0,001$, Figure 5.1). Une distribution similaire a été mise en évidence dans notre échantillon de 50 bots pour leurs 5 918 024 éditions ($r^2 = 0,837$, $p<0,001$ avec une moyenne de $118\,360 \pm 26\,372$ éditions, médiane : 58 399,5, maximum : 935 971, minimum : 1828, Figure 5.2 A, le nombre total de commentaires par page de discussion ($r^2 = 0,915$, $p<0,001$ avec une moyenne de $130,56 \pm 19,80$ commentaires, médiane : 89,5, maximum : 668, minimum : 20, Figure 5.2 B), le nombre de commentaires d'utilisateurs, excluant les réponses des propriétaires et des « veilleurs de pages » ($r^2 = 0,921$, $p<0,001$ avec une moyenne de $79,52 \pm 12,31$ commentaires, médiane : 51,5, maximum : 433, minimum : 13, Figure 5.2 C), et pour le nombre de discussions sur les pages de discussion pendant la période d'observation ($r^2 = 0,916$, $p<0,001$ avec une moyenne de $47,06 \pm 8,91$ discussions, médiane : 32, maximum : 336, minimum : 6, Figure 5.2 D). Au 4 janvier 2012, les 50 bots de notre échantillon final avaient en moyenne $1110,8 \pm 92,4$ jours d'activités (médiane : 1040, maximum : 2415, minimum :

6). Au 2 janvier 2013, ils avaient une moyenne de $5,48 \pm 0,95$ tâches actives (médiane : 4, maximum : 41, minimum : 1).

En lien avec notre premier objectif spécifique, la répartition des bots en fonction des cinq caractéristiques dichotomiques possibles était la suivante : 26 bots étaient contrôlés par un administrateur (52 %) alors que 24 ne l'étaient pas (48 %), 17 bots contraignaient le travail des éditeurs (34 %) alors que 33 bots le facilitaient (66 %), 29 bots agissaient en avant-plan (58 %) et 21 agissaient en arrière-plan (42 %), 35 bots agissaient automatiquement (70 %) et 15 l'étaient par inscription seulement (30 %). Finalement, 16 bots avaient une haute complexité d'évaluation avant d'éditer (32 %) et 34 bots avaient une basse complexité d'évaluation (68 %).

5.2 Quantification des interactions au sein de l'échantillon

Notre deuxième objectif spécifique était de quantifier les interactions qu'ont les contributeurs de Wikipédia avec les agents logiciels et leurs propriétaires sur les pages de discussion associées aux bots de l'encyclopédie. Tel que mentionné auparavant, certaines pages de discussion étaient partagées entre plusieurs bots. De ces cas précis, seul le bot principal a été compté. Dû à ce phénomène, nous avons donc exclu 150 discussions de 10 pages de discussion différentes (20 % des pages de discussion de notre échantillon). Puisqu'environ la moitié des pages de discussion étaient partagées entre le bot et son propriétaire (54 % de l'échantillon), et que dans certains cas les deux pages étaient utilisées en même temps (14 %), 3066 discussions ne concernaient pas les bots directement et ont donc été exclues ($87,60 \pm 11,45$ discussions par bots pour les 34 bots concernés spécifiquement par ce phénomène). Nous avons donc obtenu un échantillon final de 2353 discussions ($47,06 \pm 8,91$ discussions par bot). Celles-ci représentaient 6528 commentaires ($130,56 \pm 19,80$ commentaires par bot et $3,14 \pm 0,13$ commentaires par discussion), incluant 3976 commentaires d'utilisateurs généraux de Wikipédia ($79,52 \pm 12,31$ commentaires par bot et $1,86 \pm 0,07$ commentaire par discussion), 1939 réponses de propriétaires de bots ($38,78 \pm 5,66$ commentaires de propriétaires par bot et $1,05 \pm 0,08$ commentaire de propriétaires par discussion) et 613 commentaires de « veilleurs de pages » ($12,26 \pm 4,78$ commentaires de « veilleurs de pages » par bot et $0,18 \pm 0,04$ commentaire de « veilleurs de pages » par discussion). Ces deux éléments indiquent une présence régulière de responsables sur les pages de discussion des bots afin de répondre aux questionnements des contributeurs. De plus, sauf rares exceptions, les commentaires laissés par les contributeurs démontrent que ceux-ci sont très conscients qu'ils s'adressent à un éditeur humain, responsable et redevable des actions du bot. Certains propriétaires de bots le mentionnent d'entrée de jeu sur leur page de discussion :

Je suis un bot, je ne peux pas lire! Pour des demandes de fonctions, ou si ce bot n'agit pas de la bonne façon, s'il-vous-plaît aller sur la page de discussion

du propriétaire à l'adresse `commons:User%_talk:Siebrand` ou communiquez à Siebrand par `commons:Special:Emailuser/Siebrand`.

(en-tête de la page de discussion du bot *CommonsDelinker* : https://meta.wikimedia.org/wiki/User_talk:CommonsDelinker, consulté le 15 juin 2014¹)

Quant aux éditeurs expérimentés, encore plus chez les propriétaires de bots, les interventions s'adressent toujours directement aux propriétaires en question.

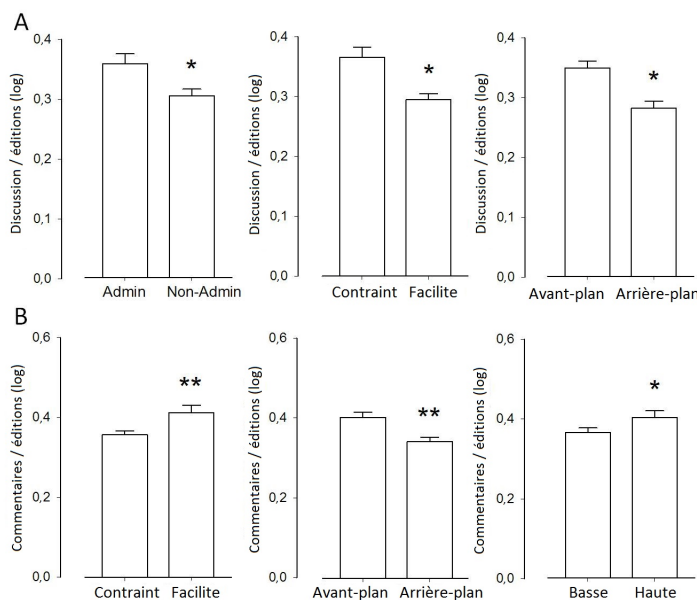
5.3 Variation des commentaires selon les variables dichotomiques caractéristiques des bots

Nous avons mesuré si certaines caractéristiques dichotomiques des agents logiciels provoquent des réactions différenciées chez les utilisateurs (ce qui était notre troisième objectif spécifique). Dans un premier temps, nous avons donc analysé le ton des réactions des contributeurs sur les pages de discussion des agents logiciels en fonction de leurs caractéristiques dichotomiques. Sur les 3976 commentaires observés, 1346 étaient positifs ($26,92 \pm 5,59$ commentaires positifs par bot, représentant $33,7 \% \pm 2,2 \%$ des commentaires), 1704 étaient neutres ($34,08 \pm 5,21$ commentaires neutres par bot, représentant $42,5 \% \pm 1,7 \%$ des commentaires) et 926 étaient négatifs ($18,52 \pm 2,77$ commentaires négatifs par bot, représentant $23,8 \% \pm 1,5 \%$ des commentaires).

Il nous est en effet possible de constater des différences significatives dans les réactions des utilisateurs en fonction des caractéristiques dichotomiques des agents logiciels. Le nombre moyen de discussions par bot était supérieur pour les bots contraignants contre facilitateurs ($75,50 \pm 22,89$ contre $32,80 \pm 5,42$, Mann-Whitney U, $p < 0,001$), avant-plan contre arrière-plan ($62,00 \pm 14,64$ contre $26,40 \pm 3,42$, $p < 0,05$) et haute contre basse complexité d'évaluation ($75,40 \pm 24,39$ contre $33,7 \pm 5,4$, $p < 0,05$) (Figure 5.3 A à la page 65). Le nombre total de commentaires par bot était relativement stable dans chacune des catégories, avec seulement significativement plus de commentaires pour les bots contraignants contre facilitateurs ($180,88 \pm 45,46$ contre $104,64 \pm 17,76$ commentaires par bot, $p < 0,05$) et pour les bots d'avant-plan contre d'arrière-plan ($165,24 \pm 31,82$ contre $82,67 \pm 11,47$ commentaires par bot, $p < 0,01$). Cette relation était aussi présente si l'on considérait seulement les commentaires des utilisateurs (excluant les propriétaires et les « veilleurs de pages ») et ce, pour les bots contraignants vis-à-vis des bots facilitateurs ($113,94 \pm 28,92$ contre $61,79 \pm 10,34$ commentaires par bot, $p < 0,01$), avant-plan contre arrière-plan ($100,90 \pm 19,89$ contre $50,0 \pm 6,6$ commentaires par

1. Traduction libre.

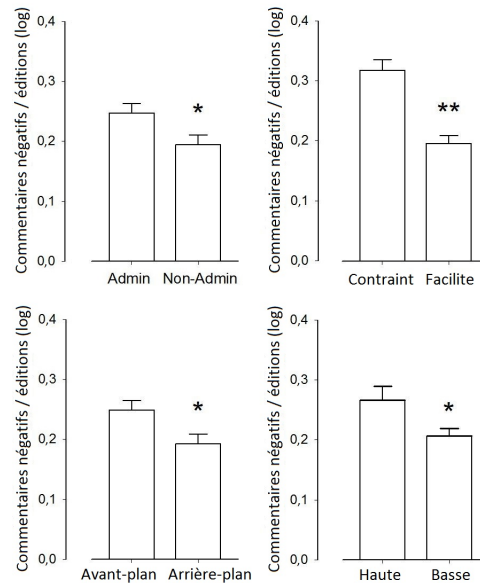
FIGURE 5.3 – **Distributions des interactions en fonction des caractéristiques des bots.** Quantification des interactions des éditeurs sur les pages de discussion du bot mesurées selon le ratio entre le nombre de discussions / le nombre d’édérations (**A**) et le ratio du nombre de commentaires des éditeurs (excluant les commentaires des propriétaires ainsi que des veilleurs de pages) / le nombre d’édérations (**B**) en fonction des cinq caractéristiques dichotomiques des bots. **A.** Des différences ont été observées dans le ratio du nombre de discussions / nombre d’édérations pour trois caractéristiques : le propriétaire du bot, le fonctionnement du bot et la location des éditions du bot. **B.** Des différences ont été observées dans le ratio du nombre de commentaires des utilisateurs (excluant les commentaires des propriétaires ainsi que des veilleurs de pages) / nombre d’édérations pour trois caractéristiques : le fonctionnement du bot, la localisation des éditions du bot, ainsi que le degré de complexité d’évaluation par le bot. Les différences ont été évaluées avec un test U de Mann-Whitney après une transformation logarithmique du ratio. * indique $p < 0,05$, et ** indique $p < 0,01$. Seulement les analyses significatives sont illustrées dans le présent graphique.



bots, $p < 0,01$) et basse contre haute complexité d’évaluation ($64,21 \pm 10,34$ contre $112,06 \pm 30,69$ commentaires par bot, $p < 0,05$) (Figure 5.3 B).

Les bots contrôlés par des administrateurs obtenaient plus de commentaires neutres ($46,35 \pm 8,97$ contre $20,79 \pm 3,32$ commentaires par bot, pour les bots contrôlés et non contrôlés par un administrateur, respectivement, Mann-Whitney U, $p < 0,05$) et de commentaires négatifs ($24,85 \pm 4,71$ contre $11,67 \pm 2,03$ commentaires par bot, $p < 0,05$). Plus de commentaires négatifs étaient observés pour les bots contraignants que pour les bots facilitants ($28,88 \pm 6,07$ contre $13,18 \pm 2,39$ commentaires par bot, $p < 0,01$). Les bots agissant en avant-plan provoquaient des commentaires davantage polarisés que ceux en arrière-plan, autant pour les commentaires positifs ($35,83 \pm 9,20$ contre $14,62 \pm 2,29$ commentaires par bot, $p < 0,01$), que pour les commentaires négatifs ($23,48 \pm 4,40$ contre $11,67 \pm 1,83$ commentaires par

FIGURE 5.4 – **Distribution des commentaires négatifs en fonction des caractéristiques des bots.** Des différences significatives ont été observées dans la distribution des commentaires négatifs laissés par les utilisateurs sur la page de discussion des bots pour quatre caractéristiques des bots : le propriétaire du bot (en haut à gauche), le fonctionnement du bot (en haut à droite), la localisation des éditions (en bas à gauche) et le degré de complexité d'évaluation du bot (en bas à droite). Les différences ont été évaluées avec un test U de Mann-Whitney après une transformation logarithmique du ratio. * indique $p < 0,05$, et ** indique $p < 0,01$.



bot, $p < 0,05$). Finalement, les bots avec une haute complexité d'évaluation généraient plus de commentaires négatifs comparés aux bots avec une basse complexité d'évaluation ($28,81 \pm 6,52$ contre $13,68 \pm 2,33$ commentaires par bot, $p < 0,05$) (Figure 5.4 à la page 66). Aucune différence significative n'était observée entre les bots « *opt-in* » et « *opt-out* ».

Les « *veilleurs de pages* » étaient significativement plus actifs chez les bots contrôlés par des administrateurs que par des non-administrateurs ($18,92 \pm 8,91$ contre $5,04 \pm 1,85$ commentaires par bots, Mann-Whitney U, $p < 0,05$) et chez les bots d'avant-plan par rapport à ceux d'arrière-plan ($18,48 \pm 8,03$ contre $3,67 \pm 1,48$ commentaires par bot, $p < 0,05$).

5.4 Analyse des types de réaction des utilisateurs

Cette section d'analyse nous permet de mesurer la distribution des discussions selon divers types de réaction des contributeurs, qui fait partie de notre quatrième objectif spécifique. Nous quantifions dans un premier temps ces discussions : un total de 2,353 discussions s'étaient déroulées au sein de notre échantillon, incluant 723 discussions positives ($15,26 \pm 4,52$ discussions par bot, $29,86 \% \pm 2,69 \%$), 932 discussions neutres ($18,64 \pm 3,06$ discussions par bot,

40,78 % \pm 1,93 %) et 657 discussions négatives (13,14 \pm 2,19 discussions par bot, 29,40 % \pm 2,15 %). La discussion moyenne contenait 1,86 \pm 0,47 commentaire, incluant 1,05 \pm 0,54 commentaire de propriétaire (généralement sous forme de réponse) et 0,18 \pm 0,27 intervention de « veilleurs de pages ». Ces résultats continuent d'éclairer notre troisième objectif spécifique.

Au sein de l'échantillon, 430 discussions concernaient des remerciements et des hommages (8,60 \pm 4,07 discussions par bot, 9,7 % \pm 1,6 %), 245 discussions étaient enclenchées par des demandes d'aide (4,90 \pm 0,79 discussions par bot, 11,8 % \pm 1,5 %), 170 discussions étaient initiées par un utilisateur informant d'une action qu'il a entrepris (3,40 \pm 0,69 discussions par bot, 7,30 % \pm 0,78 %), 385 discussions signalaient des erreurs et des faux-positifs (7,70 \pm 1,18 discussions par bot, 21,1 % \pm 2,2 %), 244 discussions offraient des suggestions (4,88 \pm 0,82 discussions par bot, 11,5 % \pm 1,0 %), 357 questionnaient le sens d'une édition du bot (7,14 \pm 1,01 discussions par bot, 17,3 % \pm 1,2 %), 229 étaient des critiques générales (4,58 \pm 9,30 discussions par bot, 9,0 % \pm 1,1 %), 198 consistaient en la contestation d'une édition particulière (3,96 \pm 0,70 discussions par bot, 9,5 % \pm 1,6 %) et 95 utilisateurs se justifiaient de leurs actions (1,90 \pm 0,65 discussion par bot, 2,7 % \pm 0,9 %).

La répartition des neuf types de réaction des utilisateurs en fonction des caractéristiques des bots présentait des effets complexes. L'analyse de cette répartition permet de répondre à nos questionnements concernant l'influence des différentes caractéristiques des bots sur l'interaction avec les contributeurs (en continuité avec notre quatrième objectif spécifique). Alors qu'il n'y avait pas de différence significative dans la répartition des neuf types de réaction des utilisateurs pour les bots contrôlés par des administrateurs et par des non-administrateurs (selon le test non-paramétrique Kolmogorov-Smirnov) (Figure 5.5 A à la page 68), les bots contrôlés par un administrateur obtenaient davantage de réactions pour certains types de réaction, tant positives (remerciement, Mann-Whitney U, $p < 0,01$, utilisateur informant d'une action $p < 0,05$) que négatives (contestation spécifique, $p < 0,05$; justification, $p < 0,05$). En contraste, une différence significative globale (observée grâce au test Kolmogorov-Smirnov) dans la répartition des neuf types de réaction des utilisateurs était observée dans le cas de la fonction du bot ($p < 0,05$, figure 5.5 B) et de la localisation de leurs éditions ($p < 0,05$, figure 5.5 C). Les bots contraignants provoquaient plus d'interventions que les bots facilitants pour un type de réaction positive (utilisateur informant d'une action, $p < 0,05$), un type de réaction neutre (questions, $p < 0,01$) et pour les trois types de réaction négatives (critique générale, $p < 0,01$; contestation spécifique, $p < 0,01$; justifications, $p < 0,001$). Notons que les bots agissant en avant-plan provoquaient plus de demandes d'aide ($p < 0,01$), mais aussi de questionnements ($p < 0,05$) et de critiques générales ($p < 0,05$), que les bots d'arrière-plan. La répartition des neuf types de réaction des utilisateurs n'était que peu affectée par le fait que les bots agissent automatiquement ou seulement avec inscription, les utilisateurs informant d'une action ($p < 0,05$) et venant se justifier ($p < 0,05$) plus fréquemment pour les bots agissant automatiquement (Figure 5.5 D).

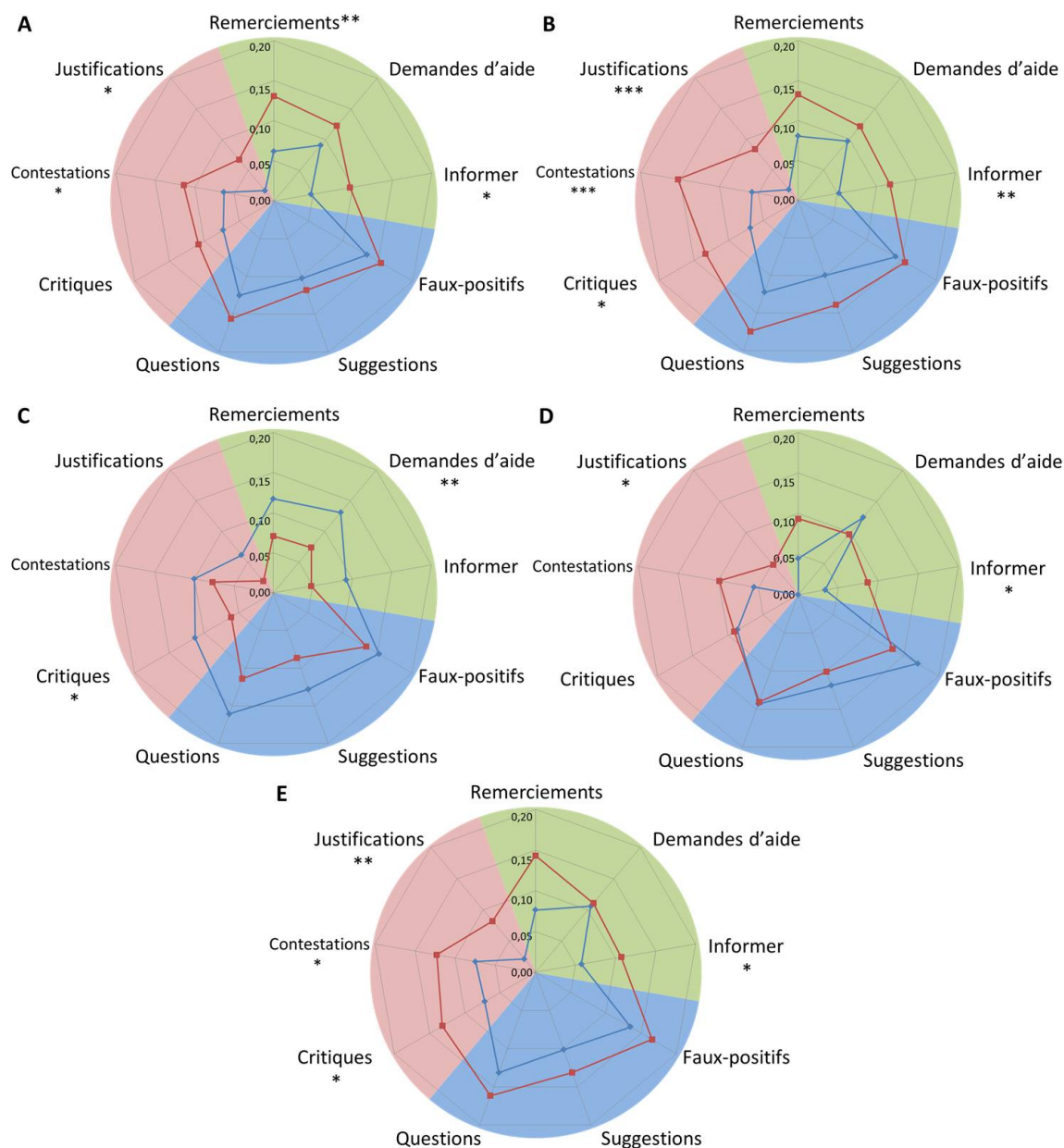


FIGURE 5.5 – Les types de réaction des utilisateurs selon les caractéristiques des bots.

Les graphiques polaires représentent la proportion des 9 types de réaction des utilisateurs aux bots (3 types de réaction positive : Remerciements, Demandes d'aide, Informer d'une action (vert pâle); 3 types de réaction neutre : Faux-positifs, Suggestions, Questions (bleu pâle)); et 3 types de réaction négative : Critiques générales, Contestations spécifiques, Justifications (rouge pâle)), selon les différentes caractéristiques des bots. Pour chaque type de réaction d'utilisateurs, l'axe radial expose la transformation logarithmique pour le ratio du nombre de commentaires de ce type de réaction / nombre total d'éditions pour le bot. Pour chaque caractéristique, les différences totales dans la répartition des types de réaction ont été mesurées grâce à un test Kolmogorov-Smirnov. * indique une différence de $p < 0,05$ et ** indique une différence de $p < 0,01$ pour le type de réaction spécifique à une caractéristique (test statistique Mann-Whitney). **A.** Réactions d'utilisateurs selon la nature du propriétaire du bot (administrateurs (gros points et ligne rouge), contre les non-administrateurs (petits points et ligne bleue)). **B.** Réactions d'utilisateurs selon le fonctionnement du bot (bots contraignant (gros points et ligne rouge), contre les bots facilitant (petits points et ligne bleue), Kolmogorov-Smirnov, $p < 0,05$). **C.** Réactions d'utilisateurs selon la localisation des éditions (arrière-plan (gros points et ligne rouge), contre d'avant-plan (petits points et ligne bleue), Kolmogorov-Smirnov, $p < 0,05$). **D.** Réactions d'utilisateurs selon le degré d'automatisation des actions du bot (bots « *opt-out* » (gros points et ligne rouge), contre les bots « *opt-in* » (petits points et ligne bleue)). **E.** Réactions d'utilisateurs selon le degré de complexité d'évaluation du bot (haut degré d'évaluation (gros points et ligne rouge), contre un bas degré d'évaluation (petits points et ligne bleue), Kolmogorov-Smirnov, $p < 0,05$).

La répartition des réactions des utilisateurs selon le degré de complexité d'évaluation des bots ressemblait sensiblement à ce qui fut observé avec le fonctionnement des bots (Figure 5.5 E). Les bots ayant une forte complexité dans l'évaluation des éditions provoquaient donc globalement plus de réactions que ceux avec basse évaluation (Mann-Whitney U, $p < 0,05$), avec autant de réactions positives (utilisateur informant d'une action, $p < 0,05$) que négatives (critique générale, $p < 0,05$; contestation spécifique, $p < 0,05$; justification $p < 0,01$).

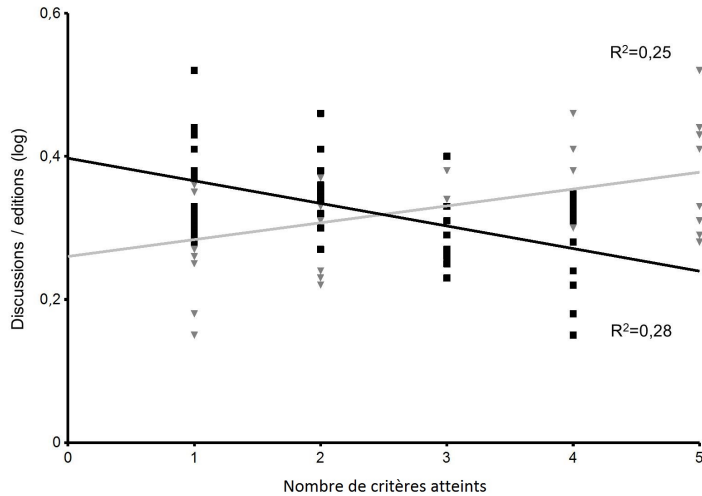
La contre-analyse, réalisée à l'aide d'une ANOVA à un facteur, modulait légèrement les résultats exposés précédemment. Nous notons que le type de réaction « remerciement » devenait significativement plus fréquent pour les bots contraignants par rapport aux bots facilitant le travail des utilisateurs, ainsi que pour les bots ayant une haute complexité d'évaluation avant de réaliser une action, face aux bots ayant une faible complexité d'évaluation ($p < 0,05$ pour chacune des deux relations). Inversement, le fait que les utilisateurs posaient plus de questions quand le bot agissait en avant-plan par rapport aux bots agissant en arrière-plan n'était plus une relation significative.

Comme résultats complémentaires, nous avons observé que les bots réalisant beaucoup d'éditions recevaient plus de commentaires négatifs (Mann-Whitney U, $p < 0,01$) ainsi que de critiques générales ($p < 0,001$).

5.5 Analyse des réactions face aux idéaux types de bots

Tel qu'exposé dans notre méthode, nous avons construit deux idéaux types distincts. Nous avons procédé ainsi à la suite du constat de la présence de certains schémas comportementaux identifiables et spécifiques. Cette classification est particulièrement utile pour exposer des éléments nous permettant de répondre à notre cinquième objectif spécifique, soit d'identifier le rôle perçu des agents logiciels, tant chez les éditeurs de Wikipédia que pour les opérateurs des agents logiciels. Cette catégorisation en idéaux types permet particulièrement de répondre à la première partie de cet objectif. Dans notre échantillon, 13 bots se retrouvaient dans l'idéal type des bots « servants » et 9 dans l'idéal type des bots « policiers ». D'un côté, les bots « servants » réalisaient des tâches laborieuses et répétitives à la place d'utilisateurs humains. Ils sont constitués des caractéristiques suivantes : ils facilitent le travail des utilisateurs, agissent principalement en arrière-plan, automatiquement « *opt-out* » ainsi qu'avec une basse complexité de l'évaluation. De l'autre côté, les bots « policiers » agissaient proactivement afin de faire respecter et appliquer les normes et règles de conduite de Wikipédia. Ils sont contrôlés par des administrateurs, contraignent le travail des utilisateurs, agissent en avant-plan, automatiquement ainsi qu'avec une haute complexité de l'évaluation. Nous les avons nommés « bots policiers ». Bien que l'ensemble des bots répondait à certains de ces critères, seulement ceux les regroupant totalement sont considérés dans les idéaux types.

FIGURE 5.6 – **Nombre de discussions en fonction du nombre de critères atteint pour les deux types d'idéaux types de bots identifiés.** Régression linéaire entre le nombre de discussions (exprimé selon la transformation logarithmique de la proportion du nombre de discussions / nombre total d'édérations pour le bot) et le nombre de critères rencontré pour les bots « servants » (ligne noire, $R^2=0.28$, $p<.05$), et pour les bots « policiers » (ligne grise, $R^2=0.25$, $p<.01$).

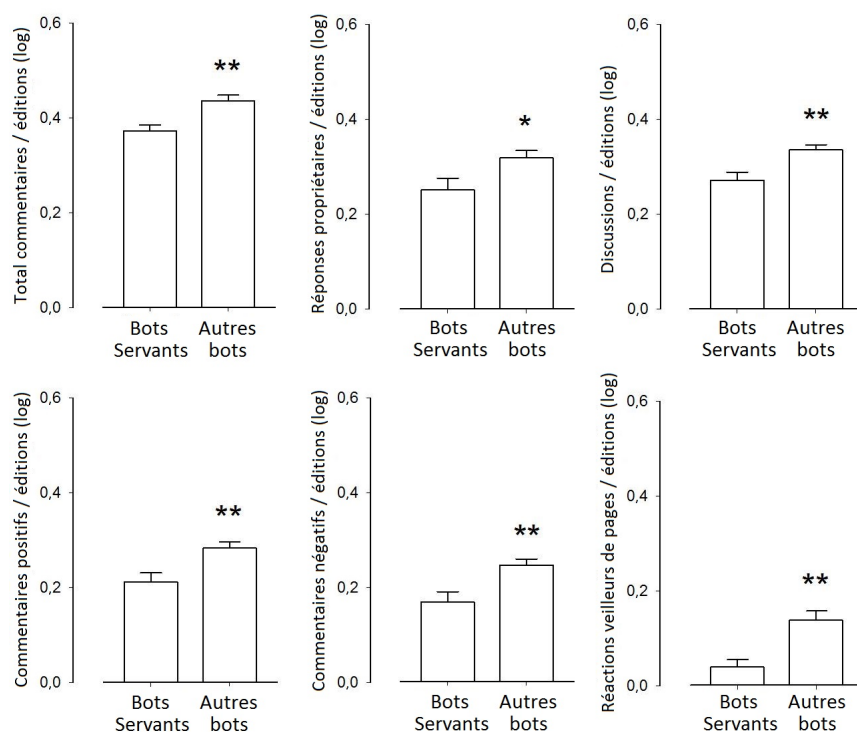


Nous avons d'abord mesuré si l'adhésion à ces idéaux types influençait quantitativement le nombre d'interactions réalisées avec les contributeurs. À cette fin, nous avons vérifié si l'adhésion à ces critères possède une influence directe sur l'interaction. Une régression linéaire nous révèle une relation inverse entre les bots « servants » et les bots « policiers » lorsque l'on compare le nombre de critères d'adhésion à l'idéal type et le nombre de discussions observées sur la page de discussion d'un bot ($r^2 = 0,28$, $p < 0,05$ et $r^2 = 0,25$, $p < 0,01$ respectivement pour les bots « servants » et les bots « policiers » (Figure 5.6 à la page 70).

De plus les bots « servants » provoquaient généralement moins d'interventions totales (Mann-Whitney U, $p < 0,01$), tant pour les commentaires d'utilisateurs ($p < 0,01$), le nombre total de discussions ($p < 0,05$), les réponses des propriétaires ($p < 0,05$), que pour les interventions des « veilleurs de pages » ($p < 0,01$). Il y eut moins de commentaires positifs ($p < 0,01$) et négatifs ($p < 0,01$) que pour les bots n'appartenant pas à cet idéal type (Figure 5.7 à la page 71). En contraste, aucune différence significative n'était observée sur le total d'interventions pour les bots « policiers ». Il s'agit d'une observation directe à l'effet que certaines fonctions réalisées par les bots provoquent des réactions différenciées chez les utilisateurs.

Répondant à notre cinquième objectif, des différences significatives étaient observées dans la répartition globale des neuf types de réaction des utilisateurs pour les bots appartenant à un idéal type et les autres bots, tant pour les bots « servants » (Kolmogorov-Smirnov,

FIGURE 5.7 – Différences au niveau de l’intensité des interactions entre les utilisateurs vis-à-vis des bots «servants» contre les autres bots. **A.** Différences pour la proportion du nombre total d’interventions / le nombre total d’éditations du bot. **B.** Différences pour la proportion du nombre de réponses du propriétaire / nombre d’éditations du bot. **C.** Différences dans la proportion du nombre de discussions / le nombre d’éditations du bot. **D.** Différences dans la proportion du nombre de commentaires positifs / le nombre d’éditations du bot. **E.** Différences dans la proportion du nombre de commentaires négatifs / le nombre d’éditations du bot. **F.** Différences dans la proportion du nombre de commentaires des veilleurs de pages / le nombre d’éditations du bot. Les différences ont été évaluées avec un test U de Mann-Whitney après une transformation logarithmique du ratio. * indique $p < 0,05$, ** indique $p < 0,01$ et *** indique $p < 0,001$.



$p < 0,05$) que pour les bots « policiers » ($p < 0,01$). Lorsque l’on considérait les réactions face aux actions des bots selon les neuf types de réaction des utilisateurs, les bots « servants » provoquaient moins de demandes d’utilisation par les utilisateurs (Mann-Whitney U, $p < 0,05$), moins de questionnements des éditions ($p < 0,05$) et moins d’informations d’erreurs ou de faux positifs ($p < 0,05$) que les bots n’appartenant pas à cet idéal type (Figure 8). Les bots « policiers » provoquaient plus de réactions polarisées significativement différentes, tant positives (remerciements, $p < 0,05$ et informant d’une action, $p < 0,05$) et négatives (critiques générales, $p < 0,05$ et justifications, $p < 0,01$) que les bots n’appartenant pas à cet idéal type (Figure 5.8 à la page 72).

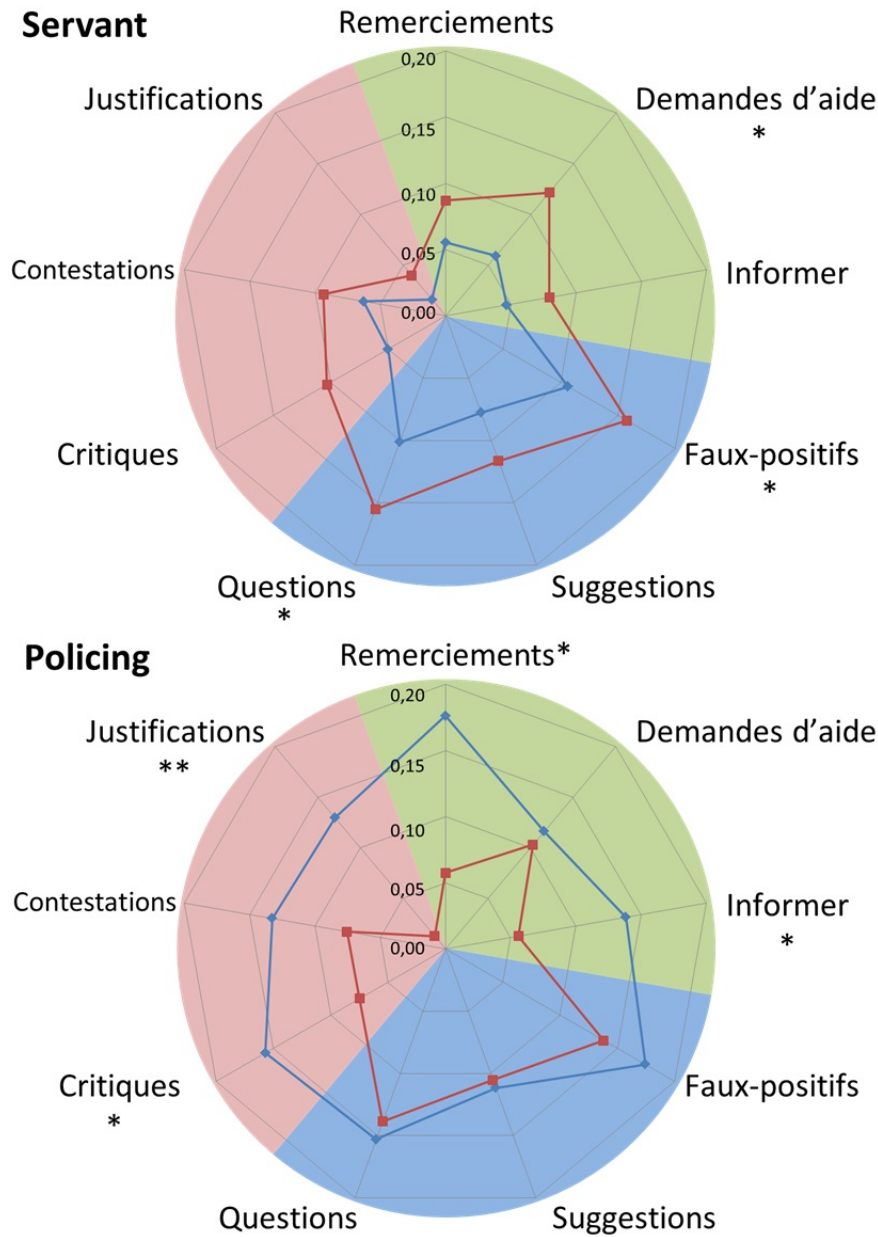


FIGURE 5.8 – **Distribution générale des types de réaction des utilisateurs selon l'idéal type des bots.** Les graphiques polaires représentent la proportion des 9 types de réaction des utilisateurs aux bots (3 types de réaction positives : Remerciements, Demandes d'aide, Utilisateur informant d'une action (vert pâle); 3 types de réaction neutres : Faux-positifs, Suggestions, Questions (bleu pâle); et 3 types de réaction négatives : Critiques générales, Contestations spécifiques, Justifications (rouge pâle)), selon les différentes caractéristiques des bots. Pour chaque type de réaction d'utilisateurs, l'axe radial expose la transformation logarithmique pour le ratio du nombre de commentaires de ce type de réaction / nombre total d'édits pour le bot. Pour chaque idéal type, les différences globales dans la répartition des neuf types de réaction ont été évaluées avec un test Kolmogorov-Smirnov. * indique une différence significative à $p < 0,05$ et ** une différence de $p < 0,01$ pour le type particulier de réaction pour la caractéristique du bot (Test U de Mann-Whitney).

Nous notons que la réalisation par les bots de tâches contraignantes touchant les contributeurs représente un indicateur clair au sein de notre échantillon quant à l'intensité des interactions. Toutefois, les réactions les plus négatives sont observées dans le cas de bots réalisant de nombreuses éditions mineures sur une même page. Par exemple, le cas de *Yobot* peut nous aider à comprendre les attentes des contributeurs. *Yobot* est en effet exclu des analyses détaillées que nous avons réalisées bien qu'il fût le bot ayant accompli le plus d'éditions dans l'encyclopédie en 2012. Nous avons observé sur sa page personnelle que celui-ci était très critiqué sur deux points. D'une part, lorsqu'il est question de changements cosmétiques et, d'autre part, quand les en-têtes des éditions sont mal identifiés (évoqué précédemment). Puisque les bots sont généralement représentés comme étant des agents facilitant le travail, nous constatons que les éditeurs deviennent rapidement agressifs s'ils se font immerger d'éditions superficielles et confondantes. Notre discussion traitera cet aspect de notre cinquième objectif spécifique.

Une analyse supplémentaire réalisée grâce à une MANOVA nous a donné des informations complémentaires sur la relation entre nos deux idéaux types et le reste de notre échantillon. Premièrement, notons que cette analyse par ANOVA était valide, tel qu'indiqué par le test de Wilks' Lambda ($p < 0,05$). L'analyse des co-variables nous permettait de constater que le lieu des discussions, à savoir si elles avaient lieu sur la page de discussion des bots, des propriétaires ou sur les deux à la fois, n'avait aucune influence sur la distribution des 9 types de réaction des utilisateurs quand il était question de nos idéaux types. Cette analyse montre toutefois que la méthode d'application des bots, soit qu'ils agissaient automatiquement ou par inscription, influençait significativement l'un des types de réaction, soit les utilisateurs informant le bot et son propriétaire d'une action qu'ils entreprennent ($p < 0,05$). Cette analyse a permis d'observer une différence significative dans le cas des remerciements ($p < 0,05$), informés d'une action ($p < 0,05$), de critique générale ($p < 0,05$) ainsi que pour un utilisateur se justifiant ($p < 0,001$). Une analyse complémentaire d'une ANOVA à un facteur usant d'un test post-hoc de Tukey a montré les différences spécifiques de ces relations. Dans le cas des remerciements, l'idéal type des bots « policiers » était significativement différent, autant qu'avec l'idéal type « servants » ($p < 0,05$) qu'avec le reste de l'échantillon ($p < 0,01$). Pour ce qui était des utilisateurs informant d'une action qu'ils entreprenaient (dont la co-variable touchant les bots automatiques et par inscription s'était révélée statistiquement significative), la même tendance a été observée pour les bots « policiers » face aux bots « servants » ($p < 0,05$) et au reste de l'échantillon ($p < 0,05$). De plus, nous avons constaté que c'était spécifiquement vis-à-vis de l'idéal type de bots « servants » que les « policiers » regroupaient davantage de critiques générales ($p < 0,01$). Enfin, dans le cas d'utilisateurs venant se justifier, les bots « policiers » étaient significativement plus fréquents que les bots « servants » ($p < 0,001$) et les bots du reste de l'échantillon ($p < 0,001$).

5.6 Analyse des discussions sur le panneau d’affichage des propriétaires de bots

Afin de terminer de répondre à la deuxième partie de notre cinquième objectif spécifique, soit le rôle perçu des agents logiciels chez les opérateurs d’agents logiciels, nous avons analysé l’ensemble des discussions ayant eu lieu sur le panneau d’affichage des propriétaires de bots durant l’année 2012. Spécifiquement, il s’agissait de discussions ayant eu lieu entre le 10 janvier 2012 et le 27 décembre 2012. Celles-ci représentaient un total de 744 commentaires répartis en 79 discussions différentes ($9,42 \pm 0,42$ commentaires par discussion). Ce panneau était aussi particulièrement utile pour approfondir notre compréhension des phénomènes sociaux sous-jacents, notamment dans la coordination et l’esprit d’association présents chez les propriétaires de bots

Logiquement, la forte majorité des interventions observées était constituée par des propriétaires de bots, ce lieu de discussions étant en effet dédié à leur coordination. Dans un premier temps, nous avons quantifié ces interactions. Les interventions réalisées par les propriétaires de bots représentaient un total de 486 interventions (soit 65,3 % du nombre total de commentaires, ainsi que $6,15 \pm 0,38$ commentaires par discussion). Parmi ceux-ci, certains se retrouvaient au même moment dans la posture d’administrateur, ce qui représentait 85 interventions (soit 11,4 % du nombre total de commentaires, représentant $1,08 \pm 0,27$ commentaires par discussion) ainsi que 29 propriétaires de bots membres du *Bot Approvals Group* (BAG) (soit 3,9 % du nombre total de commentaires, pour $0,37 \pm 0,20$ commentaires par discussion). De plus, 142 interventions étaient réalisées par des propriétaires membres du BAG et ayant à la fois les pouvoirs d’administrateurs (soit 19,1% du nombre total de commentaires, ainsi que $1,80 \pm 0,25$ commentaires par discussion). Chez les utilisateurs de Wikipédia n’étant pas propriétaires de bots, 252 ont réalisé une intervention sur ce panneau d’information (soit 33,9 % du nombre total de commentaires, représentant $3,27 \pm 0,38$ commentaires par discussion). De ceux-ci, 91 interventions étaient réalisées par des administrateurs de Wikipédia n’ayant pas un statut particulier dans la gestion des bots (soit 12,2 % du nombre total de commentaires, ainsi que $1,15 \pm 0,28$ commentaires par bot). Enfin, les utilisateurs de Wi-

Statut	Nombre d’interventions	Pourcentage
BAG-Admin-Proprio	142	19,1 %
Admin-Proprio	85	11,4 %
Admin	91	12,2 %
BAG-Proprio	29	3,9 %
Proprio	230	30,9 %
Utilisateur inscrit	167	22,4 %

Tableau 5.1 – Répartition des interventions selon le statut de l’utilisateur dans Wikipédia.

Tableau 5.2 – Répartition des discussions selon leur orientation.

Type de discussion	Sous-type	Nombre de discussions	Pourcentage
Communauté de Wikipédia		11	14,3 %
Maintient des capacités techniques		18	23,4 %
Type de débats et conflits	Légitimité	19	24,7 %
	Processus	14	18,2 %
	Technique	15	19,5 %

kipédia non-proprétaires de bots et n’ayant pas le statut d’administrateur ont participé à hauteur de 167 commentaires (représentant 22.4 % du nombre total de commentaires, ainsi que $2,11 \pm 0,35$ commentaires par discussion). Ces résultats sont résumés dans la table 5.1.

Un nombre élevé de discussions touchait le maintien de capacité technique des propriétaires de bots (18 discussions représentant 23,4 % du nombre total). Ces discussions étaient souvent reliées au départ, volontairement ou non, à un propriétaire de bots et aux inquiétudes liées à la pérennité des fonctions que ces bots apportent à la communauté. Justement, 11 discussions (14,3 % du nombre total) touchaient spécifiquement la communauté (qu’il s’agisse de celle de Wikipédia ou de la sous-communauté des propriétaires de bots). Cela pouvait consister en une invitation pour un événement extérieur ou une discussion sur un débat affectant une autre section de l’encyclopédie. Toutefois, la plupart des discussions observées pouvaient être classées selon trois types de débats/conflits : celles touchant la légitimité (19 cas se questionnant sur la légitimité et le consensus social entourant les actions des bots et de leur propriétaire, pour 24,7 % des discussions totales), celles touchant les processus (14 cas proposant et débattant de la façon de réaliser une tâche, pour 18,2 % des discussions totales) ainsi que celles touchant des aspects strictement techniques (15 cas touchant surtout la faisabilité technique de tâches particulières et des limitations présentes à même l’infrastructure de Wikipédia, pour 19,5 % des discussions totales). Ces résultats sont présentés dans la table 5.2.

Le rôle perçu des agents logiciels se traduit pour la communauté par l’adoption de normes venant encadrer le travail dit légitime. Considérant la formation des normes, plusieurs éléments surprenants nous sont apparus à la lecture du panneau de discussions des propriétaires de bots. En effet, nous avons observé l’évolution de certaines normes au cours de l’année 2012. D’un premier aspect, la norme touchant le nombre maximal d’édicions par minute autorisée aux bots n’est plus appliquée avec la même force qu’annoncée dans les politiques officielles. Bien que cette règle soit encore présente chez les contributeurs réguliers et les utilisateurs de logiciels d’assistance à l’édition, encore plus chez les nouveaux utilisateurs, la limite n’existe plus formellement pour l’utilisation des bots. En fait, ceux-ci sont autorisés à éditer aussi vite que ce que peuvent exécuter les serveurs de MediaWiki. Cette capacité est permise par l’option « *Maxlag* » du logiciel, permettant de calculer les ressources disponibles et de ralentir les bots lors des heures d’achalandage. Toutefois, les propriétaires eux-mêmes se questionnent

sur la légitimité d'un tel principe. En effet, l'inquiétude d'un bot perdant le contrôle reste bien présente chez ceux-ci, le volume de dommages potentiels étant proportionnel à la rapidité d'éditions du bot. Certains dénotent qu'une capacité d'édition absolue d'une édition par seconde devrait être imposée. Par ce questionnement, nous constatons à quel point l'acceptation des bots s'est étendue, puisque dans le premier cas connu de leur utilisation (avec l'importation de la *Easton's Bible Dictionary*), la communauté suggérait un maximum d'une édition par 20 minutes².

À l'inverse, cette analyse des discussions est aussi l'occasion d'observer l'instauration progressive d'une norme « *opt-in* » au cours de l'année. Le principal protagoniste est ici *Edward-Bot*, bot réalisant la distribution de bulletins d'information « *newsletters* » à partir de listes d'envoi, parfois réalisées par d'autres éditeurs que le propriétaire de bots. Dans un cas précis, un éditeur a ajouté 985 individus participants à un type particulier de pages de discussions dans l'encyclopédie, et ce, sans leur en demander explicitement l'autorisation. Ceci a donné lieu à des accusations d'émissions de messages indésirables :

Sérieusement. Participer à un « *WP board* » n'est pas une autorisation pour recevoir du spam. Ce n'est pas correct, et je suis surpris qu'il n'y ait rien dans les règles de conduite interdisant des utilisateurs d'en inscrire d'autres pour recevoir du spam sans leur permission.

Roscelese,
sur le panneau d'affichage des propriétaires de bots,
4 septembre 2012³.

Croyant suivre le consensus social en cours (la norme indique en effet qu'il est nécessaire d'offrir un mécanisme « *opt-out* », même pour les bots distribuant le courrier), cette tâche est autorisée en 24 heures sur l'avis de seulement quatre individus. Toutefois, cette liste d'envoi tant réalisée sans considération que même certains bots se retrouvent abonnés à ces courriels. Après coup, un courriel d'excuse est envoyé selon le même procédé, rien pour calmer les participants :

L'ironie est un bulletin d'information « résolution de dispute » causant conflits et perturbations, mais c'est là que nous en sommes. Avec d'autres, je suis d'accord que les bulletins d'information devraient toujours être seulement « *opt-in* ». Toutefois, cela n'est pas la faute à *EdwardsBot*. Il s'agit d'un processus de décisions pauvres de la part des organisateurs de ce bulletin d'information.

2. <http://lists.wikimedia.org/pipermail/wikipedia-l/2002-August/003597.html>, consulté le 4 septembre 2014.

3. Traduction libre.

Resolute,
sur le panneau d'affichage des propriétaires de bots,
4 septembre 2012⁴.

Tout en reconnaissant la bonne volonté du propriétaire d'*EdwardsBot*, il est décidé à ce moment que tous bots dédiés au courrier non essentiel, tels les bulletins d'information, seront strictement « *opt-in* ». À partir de ces éléments de réponse, nous allons dresser dans la discussion un portrait davantage étayé de la perception que la communauté possède des agents logiciels.

4. Traduction libre.

Chapitre 6

Discussion

L'objectif principalement poursuivi par ce mémoire était d'établir un portrait général de l'interaction entre les contributeurs de la version anglaise de Wikipédia, qu'ils soient propriétaires de bots, contributeurs récurrents de l'encyclopédie ou éditeurs occasionnels, et les agents logiciels, nommés bots au sein de l'encyclopédie. À cette fin, nous avons successivement étudié les caractéristiques descriptives des bots au sein du Wikipédia anglais, le ton des commentaires laissés sur la page de discussion de ces agents logiciels, le sujet de ces mêmes discussions ainsi que celles réalisées sur le panneau d'affichage des propriétaires de bots.

Nos résultats démontrent que les bots représentent des acteurs incontournables de l'encyclopédie, comme en témoigne l'impressionnant dispositif encadrant leur utilisation que nous avons pu observer. Alors qu'il serait aisé de croire que les agents logiciels agissent sans supervision une fois autorisés, leur utilisation démontre sur bien des points un rôle d'accompagnement des éditeurs. Beaucoup d'éléments sont en effet retrouvés afin de limiter la capacité réelle d'action de ces bots, telles la capacité des administrateurs de les bloquer, l'assiduité des propriétaires à répondre aux commentaires ainsi que les discussions entre les propriétaires et administrateurs portant sur la limite des droits des agents logiciels. Toutefois, ils ne sont pas tous perçus sur le même pied d'égalité et les interactions complexes qu'ils ont avec la communauté sont partiellement déterminées par le type de tâches qu'ils réalisent. À cet égard, nous identifions deux idéaux types qui regroupent, à divers degrés, une proportion significative des bots observés lors de cette étude. Cette catégorisation se base essentiellement selon le type de tâches réalisées par le bot (l'objectif de standardisation poursuivi). Cette catégorisation est validée par notre observation de fortes différences statistiques dans la distribution des interactions qu'ils ont avec les participants. Tel que présenté précédemment, nous avons nommé bots « servants » ceux qui réalisent principalement la standardisation d'éditions mineures, mais généralement répétitives et laborieuses pour les éditeurs. Les bots « policiers », pour leur part, regroupent ceux qui réalisent la standardisation de comportements et types de contribution légitime, ayant ainsi une portée beaucoup plus intrusive pour le travail des usagers.

Contrairement à nos attentes initiales, les bots apparaissent comme étant largement acceptés au sein de la communauté interne de l'encyclopédie, tel que démontré par le nombre relativement peu élevé de types de réaction négative observé dans notre échantillon. Ceci, même si la forte majorité des contributeurs interagissent avec les bots en réaction au travail de ces derniers. Les bots sont en effet largement perçus comme un élément important de la viabilité de l'encyclopédie, étant régulièrement promus et défendus par les membres de la communauté interne (notamment par la participation des veilleurs de pages). Cette acceptation peut paraître étonnante étant donné les risques que représentent les bots relativement à la balance des pouvoirs au sein de l'encyclopédie. Ceci nous permet d'observer la fracture existante entre ces contributeurs fréquents (qui sont partie prenante des discussions) et le reste des participants, dont la contribution reste faible au niveau organisationnel.

6.1 Considération sur la validité générale de notre modèle

Wikipédia constitue un modèle de recherche ayant une popularité grandissante (Jullien, 2012 ; Lawler, 2006 ; Martin, 2010 ; Yasseri et János Kertész, 2013 ; Okoli et collab., 2012, 2014 ; Mesgari et collab., 2014). Ceci est dû autant à son statut de réussite par excellence d'une communauté collaborative en ligne, que par l'accessibilité totale qu'elle autorise de ses archives. Cependant, les recherches touchant spécifiquement l'utilisation des bots restent à ce jour limitées. Ce n'est qu'avec leur récente montée en importance que des chercheurs y ont porté leur intérêt (Geiger, 2011 ; Halfaker et Riedl, 2012 ; M. Livingstone, 2012 ; Müller-Birn et collab., 2013 ; Niederer et van Dijck, 2010). Jusqu'à maintenant, l'accent de ces recherches a surtout porté sur la place structurelle qu'occupent les bots au sein de l'encyclopédie, abordant souvent la structure sociotechnique dans laquelle les bots s'insèrent (Halfaker et Riedl, 2012 ; M. Livingstone, 2012 ; Niederer et van Dijck, 2010). De plus, certains auteurs ont analysé les réactions de la communauté lors de certains événements charnières touchant l'établissement des bots (Geiger, 2011 ; M. Livingstone, 2012). Toutefois, peu se sont concrètement intéressés à analyser la perception et l'interaction que les contributeurs entretiennent avec les bots au jour le jour. Wikipédia étant constitué d'une communauté en constant développement, comprendre ces dynamiques nous aide à bien identifier l'importance structurelle que représentent les bots pour l'encyclopédie.

Wikipédia n'est pas le premier projet à faire une utilisation large de bots, ceux-ci étant régulièrement utilisés dans une majorité de canaux IRC (Golbeck et Mutton, 2004 ; Canavan, 2005 ; Geiger, 2014). Ceci étant dit, Wikipédia constitue un terrain privilégié afin de les étudier, tant par leur nombre brut que par l'amplitude et l'utilisation massive qui est faite de ceux-ci. En ce sens, ce terrain représente sans doute le plus imposant à notre disposition, puisqu'une majorité de contributeurs sont exposés aux actions des bots. Rappelons que des 151 bots identifiés en 2006 (Niederer et van Dijck, 2010), puis des 700 en septembre 2011

(M. Livingstone, 2012), nous avons identifié, lors de notre recherche, un ensemble de 872 bots désormais autorisés sur la version anglaise de l'encyclopédie en date du 2 janvier 2013. Ceci représente 1789 différentes tâches¹. Plus que leur nombre brut, nous avons estimé, en utilisant les statistiques mensuelles compilées par Wikipédia même², que les éditions réalisées par les bots représentent un impressionnant 22 % du nombre total d'éditions de Wikipédia anglais durant l'année 2012, en croissance face aux années précédentes (Geiger, 2009). Ce qui vient confirmer notre première hypothèse à l'effet que : *Les agents logiciels continuent d'augmenter leur importance au sein de l'encyclopédie*. Notons qu'avec l'émergence de Wikidata, la proportion d'éditions réalisée par les bots semble avoir significativement diminué dans la version anglaise de Wikipédia (Steiner, 2014). Ceci est dû à la récupération de l'ensemble des tâches touchant les liens inter-wikis par Wikidata.

C'est que nous retrouvons de particulier est l'utilisation de ces agents logiciels par la communauté même qui constitue Wikipédia. En effet, les agents logiciels sont habituellement utilisés sous deux aspects bien différents dans les médias sociaux : soit par l'organisme même, dans l'objectif d'améliorer le service offert ainsi que de réaliser des cueillettes de données et de statistiques ; ou par des contributeurs malveillants, dont l'objectif est souvent lié à des activités de perturbation ou de pourriels. En ce sens, Wikipédia est bel et bien un terrain sans équivalent nous offrant la possibilité d'analyser l'utilisation des bots ainsi que des rôles qu'une communauté collaborative pourra progressivement leur déléguer.

6.1.1 Composition de l'échantillon

Malgré ce fort nombre de bots accepté dans l'encyclopédie, seul un faible pourcentage est effectivement actif durant la fenêtre de temps de notre étude. En effet, entre le 4 janvier 2012 et le 2 janvier 2013, seuls 170 bots réalisent des éditions dans l'encyclopédie. Ce chiffre est légèrement en-dessous du nombre estimé actuellement (188 bots en date du 14 avril 2014³). Ce qui signifie un imposant nombre de bots ayant l'autorisation d'agir, mais actuellement inactifs. Quelle est l'importance qu'occupe cette imposante masse passive de bots ? Tout d'abord, il semble qu'un grand nombre réalisent des tâches uniques ou d'usages occasionnels. Cela peut être le déplacement massif d'une catégorie d'articles au sein d'une autre, le remplacement d'un gabarit qui n'est plus à jour, etc. Sur certains aspects, les actions de ces bots font écho aux importations massives réalisées dans les premières années de l'encyclopédie. De plus, nous savons qu'un certain nombre de bots autorisé ne l'est que dans la perspective de remplacer, temporairement ou non, un bot tombé hors service (Geiger, 2014). Dans notre échantillon, nous notons entre autres que *VWBot* (2496 éditions en 2012) remplace *DumbBOT* (21 137 éditions) lorsque ce dernier tombe hors ligne.

1. <https://en.wikipedia.org/wiki/Wikipedia:Bots/Status>, consulté le 6 novembre 2014.

2. <http://stats.wikimedia.org/EN/TablesDatabaseEdits.htm>, consulté le 6 novembre 2014.

3. Notons que la section Wikipedia : Bots/status est reconnue pour avoir quelques erreurs dans ses décomptes.

La construction de notre échantillon montre une forte disparité entre la présentation et le fonctionnement des bots : malgré l'importance structurelle des bots, ceux-ci sont loin d'afficher une présentation standard pour la communauté. Entre autres, nous notons *ZéroBot* (6^e bot en importance en 2012 avec 409 491 éditions) et *Yfdyh-bot* (14^e avec 146 311 éditions), qui malgré leur importance, ont un propriétaire dont la page principale n'est pas située dans la version anglaise de Wikipédia. Leur page de discussion réalise en effet une redirection vers une autre version de l'encyclopédie. Ce phénomène est loin d'être anodin, puisque ceci réduit énormément la capacité de la communauté à encadrer l'utilisation de ces bots. Cette variabilité dans la présentation et le maintien des bots semble correspondre au concept de code fait sur mesure « *bespoke code* » (Geiger, 2014). Cette conception nous permet de comprendre qu'une grande partie de la multitude de nouvelles fonctions émergentes, composant le cadre sociotechnique de l'encyclopédie, est produite et maintenue par des individus n'ayant pas de lien direct avec MediaWiki, dépositaire de l'infrastructure de l'encyclopédie. Ceux-ci ne peuvent donc pas offrir le même genre de garantie et de fiabilité que l'encyclopédie. D'ailleurs, il semble en effet bien difficile de maintenir ces fonctions, puisque 23,4 % des discussions du tableau d'affichage des propriétaires de bots de 2012 touchent directement la capacité de la communauté à s'assurer de leur maintien. La majorité de ces discussions concerne le remplacement de bots ou de leurs propriétaires, désormais absents de l'encyclopédie pour diverses raisons. Cette mobilisation constante de la communauté montre bien que les fonctions remplies par les bots sont considérées comme importantes, bien que nécessitant un entretien constant pour en assurer leur maintien.

Une question émergente est donc de savoir si les interactions éditeurs-bots dans Wikipédia sont représentatives des phénomènes cognitifs habituellement observés dans les communautés virtuelles, tant en termes de magnitude que de qualité. Valider la composition de notre échantillon peut nous aider à répondre à cette interrogation. Dans un premier temps, l'analyse de notre échantillon nous permet de constater que la distribution de nos données correspond aux mêmes phénomènes que ceux généralement observés dans Wikipédia. Que des distributions concernant la participation et la croissance de l'encyclopédie suivent une loi de puissance est en effet constaté à de multiples occasions (Muchnik et collab., 2013 ; Nazir et Takeda, 2008 ; Voss, 2005 ; Yasseri et collab., 2012). Ce type de distribution est connu pour être fréquemment retrouvé sur Internet (Albert et collab., 1999 ; Adamic et Huberman, Mars 2000) et dans les réseaux sociaux. Le fait de retrouver le même type de distribution au sein de notre échantillon suggère fortement que celui-ci reflète correctement la population des bots, tout en nous confortant sur notre capacité à analyser les interactions que ceux-ci ont avec les contributeurs. Toutefois, et tel que mentionné dans notre méthode, il s'agit seulement ici des contributeurs ayant participé aux pages de discussion des bots. Il serait donc difficile de généraliser nos résultats à l'ensemble des contributeurs aux articles de Wikipédia.

6.2 La présentation des bots face à la communauté

Les études réalisées sur l'interaction utilisateurs-bots, voir homme-machine, le sont souvent auprès de bots ayant un objectif déclaré de maintenir des conversations convenables avec leurs interlocuteurs (Nass et collab., 1996 ; Aharoni et Fridlund, 2007 ; Lortie et Guitton, 2011 ; Epley et collab., 2007). Bien que la perception des contributeurs quant aux bots semble aller dans le même sens que la littérature, la seule interaction systématique qu'il est possible de considérer comme étant « communicative » de la part des bots de Wikipédia, est généralement le message laissé en en-tête des éditions sur l'encyclopédie. Volontaire dans le cas humain, celle-ci est obligatoire dans le cas des bots. Comme l'expose le cas de *Yobot*, bloqué deux fois durant l'année 2012, les éditeurs montrent une très forte intolérance envers les en-têtes n'exposant pas clairement l'objectif et le contenu de l'édition réalisée par le bot.

Toutefois, malgré cette absence d'interaction communicative avec les bots, le premier contact social des nouveaux contributeurs est réalisé, deux fois sur trois, soit par un humain utilisant un logiciel d'assistance à l'édition, usant d'un message standardisé, soit par un bot (Geiger et collab., 2012). Puisque les messages de bienvenue ne sont pas autorisés dans le cas des bots, considérés trop impersonnels, ce premier contact est souvent de nature autoritaire et normative. Ces nouveaux contributeurs se voient souvent signifier que leur contribution n'a pas atteint les standards de l'encyclopédie (quand ils ne sont pas directement accusés de vandalisme). Pour beaucoup d'éditeurs, les interactions qu'ils ont avec les bots sont donc déclenchées en constatant le contrôle des bots sur leurs propres actions. Par contre, cette interaction avec la communauté est très variable selon le type de bots. Dans notre échantillon notamment, 40 % des bots ayant plus de 1000 éditions n'ont même pas reçu 12 messages sur un an (34 sur 85 bots ayant rempli toutes les autres conditions de notre échantillon). Toutefois, si l'on considère ces résultats à la lumière de la distribution en loi de puissance ainsi que selon leur idéal type, cette variabilité n'est pas surprenante. Ainsi, il est possible d'affirmer que de manière générale, l'ampleur des éditions réalisées par les bots impose certaines attentes spécifiques des éditeurs, ce qui est analysé à la prochaine section.

De plus, bien que les agents logiciels se doivent de laisser des messages respectant des standards humains de qualité et de compréhensibilité à chacune de leurs éditions, aucun des bots n'est conçu pour passer un test de Turing. Ce test permet de mesurer la capacité d'une machine à réaliser des comportements intelligents, indistinguables de ceux d'un humain (Saygin et collab., 2000). Nous pouvons alors supposer que l'implication fréquente et constante des propriétaires, administrateurs et veilleurs de page permet de réduire la perception des autres éditeurs de bots trop indépendants, en leur rappelant constamment leur statut d'aide éditoriale.

6.3 L'interaction entre les éditeurs et les bots

Alors que le contrôle humain est apparent pour la plupart des bots considérés dans cette étude, la possibilité d'actions massives des bots sur le contenu de Wikipédia défie cette simple vision. Dans ce contexte, est-ce que les éditeurs ne les considèrent seulement qu'en tant que servants automatisés, ou des perceptions plus complexes peuvent-elles s'instaurer ?

Bien qu'à priori intuitive, notre catégorisation de bots portant sur le degré de visibilité que leurs éditions a prédit, à tout le moins partiellement, les interactions qu'ils ont avec les membres de la communauté. En effet, les bots agissant à l'avant-plan de l'encyclopédie regroupent la majorité des réactions observées et comparées aux bots d'arrière-plan. Ceci n'est pas étonnant, car il est observé que la connaissance du fonctionnement de l'encyclopédie décroît avec la teneur administrative et d'arrière-plan de celles-ci (Antin et Cheshire, 2010). Si le bot occupe des fonctions structurelles ou organisationnelles, au contraire d'agir à même les articles, il est peu probable que beaucoup de participants – souvent plus attentifs au « contenu » immédiatement visible – s'aperçoivent de leur existence. Puisque les bots sont généralement représentés comme étant des agents facilitant le travail, nous constatons que les éditeurs deviennent rapidement agressifs s'ils se font immerger d'éditations superficielles et confondantes, tel qu'observé sur la page de discussion de *Yobot*. Cette affirmation est aussi valide dans le cas des logiciels d'assistance à l'édition, d'autant plus qu'AWB (*Auto-Wiki Browser*, logiciel d'assistance à l'édition) est censé inclure ces modifications mineures au sein d'éditations regroupant des changements notables. Il n'est alors pas étonnant de constater que le contributeur nommé Magioladitis, propriétaire de *Yobot*, est aussi un des développeurs d'AWB. Les éditeurs humains, tout comme les bots, sont encouragés à réaliser ce type de modifications mineures au sein d'une édition touchant plusieurs éléments à la fois. Par contre, dû au caractère anecdotique que le travail manuel sur un article peut représenter, l'éditeur ne sera pas réprimandé si sa contribution se résume en la suppression d'un caractère invisible. Les bots au contraire, avec leur pouvoir d'édition, sont aptes à transformer rapidement cette situation et à entraver la capacité de suivre le fil des événements aux éditeurs, en les saturant d'informations vides de signification et finalement, superflues.

Lorsque les bots sont considérés au sein d'un continuum défini par le type d'actions qu'ils réalisent, la perception des contributeurs change selon la fonction du bot. Ceci permet d'éclairer sous un autre angle notre cinquième objectif spécifique : *Identifier le rôle perçu des agents logiciels, tant chez les éditeurs de Wikipédia que chez les propriétaires des agents logiciels*. En termes de perception, les bots répondant fortement à l'idéal type « servant » reflètent effectivement le stéréotype généralement retrouvé de servant automatisé. En réalisant des tâches répétitives et laborieuses, ces bots réduisent le poids des tâches nécessaires aux contributeurs (Geiger, 2011 ; Halfaker et Riedl, 2012) et sont perçus clairement comme des aides automatisées, que l'on peut percevoir comme une simple amélioration du logiciel. Cette catégorie

d'agents logiciels semble en effet largement acceptée par les contributeurs. Comme résultat, ces bots offrent de très peu d'interactions avec la communauté. L'observation à l'effet qu'il y a moins de faux positifs identifiés, de questionnements et même de demandes d'utilisation, indique que ces bots réalisent bien le travail pour lequel ils sont assignés. En effet, ce type de bot semble bien là pour être oublié (Geiger, 2011). Toutefois, bien qu'à priori mineur, le nombre de corrections petites, répétitives et principalement invisibles faites par ces bots (mais finalement aussi par l'ensemble des bots), représente un grand impact structurel sur l'encyclopédie. La standardisation des gabarits, catégories et autres éléments n'est pas à prendre à la légère, puisque l'encyclopédie repose sur ces éléments afin de maintenir une classification cohérente et être apte à répartir les efforts et tâches entre les contributeurs. L'allègement qu'ils représentent quant au fardeau de tâches des utilisateurs semble expliquer cette grande acceptation, ce qui semble répondre positivement à notre deuxième hypothèse, soit *les agents logiciels représentent un outil important pour leur propriétaire leur permettant d'alléger le poids des tâches nécessaires qu'ils ont à réaliser pour l'encyclopédie.*

La situation est plus complexe quand il est question des bots invasifs « policiers ». En effet, les bots correspondant à cet idéal type agissent principalement en intervenant directement dans le travail des individus, au contraire des bots de type « servant », qui eux réalisent des tâches complémentaires et souvent parallèles à l'éditeur. À première vue, et comme nous en émettions l'hypothèse au début de ce mémoire, ceux-ci risquent d'être perçus comme étant agressifs ou à tout le moins, contraignants. Très souvent, les bots de cette catégorie exécutent systématiquement (ou « forcent » les éditeurs à appliquer) des sélections de normes. Ceci standardise *de facto* certains types de contributions et de comportements considérés comme légitimes. Normes pourtant initialement adoptées dans la perspective d'être appliquées volontairement (en lien avec l'« *Ignore all rules* », cinquième pilier de Wikipédia). Nous constatons que ces bots concentrent la majorité des interactions avec les contributeurs ainsi que des rétroactions de la communauté. Ces interactions peuvent être décrites selon deux hypothèses opposées, bien que complémentaires. Une première possibilité consiste en la considération de ces bots par les contributeurs comme étant des intrus externes, déclenchant massivement des critiques et rejets. Cette perspective semble particulièrement valide quand l'on considère les contributeurs sporadiques, puisque ceux-ci semblent peu au courant de l'existence des normes et encore moins de leur application par des agents logiciels. Alternativement, ces actions sur le contenu de Wikipédia peuvent être perçues comme le reflet de l'application voulue des normes de l'encyclopédie. Dans ce cadre, les débats recensés sur les pages de discussion des bots s'orientent souvent autour de la légitimité des normes, au contraire de l'action même de ces bots. Cette seconde possibilité apparaît valide, puisque les bots « policiers » provoquent à la fois beaucoup de commentaires positifs et négatifs, consistant autant en des remerciements qu'en des critiques générales.

On comprend alors que, pendant que les bots « servants » permettent concrètement d’alléger les tâches et d’accumuler quantité d’informations autorisant des décisions et des actions autrement impensables chez les éditeurs, les bots « policiers », quant à eux, sont d’une autre nature. Les tâches qu’ils réalisent sont, par leur formulation même, principalement orientées pour appliquer des normes chez des éditeurs peu au fait de celles-ci. La vigueur constatée chez les opérateurs de bots, les administrateurs et les veilleurs de pages pour défendre les bots soutient cette interprétation. Il s’agit donc d’une confirmation supplémentaire de notre deuxième hypothèse : *Les agents logiciels représentent un outil important pour leur propriétaire leur permettant d’alléger le poids des tâches nécessaires qu’ils ont à réaliser pour l’encyclopédie.* Alors que ceci est évident pour les bots «servants», cette deuxième forme de bots rend effectives des décisions prises socialement par une communauté « précise ». Ces contributeurs réguliers n’ont souvent plus la patience d’enseigner et de ré-enseigner les normes de l’encyclopédie à la multitude d’utilisateurs lambda de Wikipédia, souvent perçus comme simples passagers. Ceci est appuyé par la régulière mention, par ces propriétaires, des normes déjà en vigueur lorsqu’ils défendent leurs bots. Cette hypothèse n’est pas entièrement validée pour autant, puisque cette manière de faire peut cacher une contestation réelle des contributeurs. En effet, il est commun pour les propriétaires d’affirmer leur ouverture à la possibilité d’établir un nouveau consensus lorsqu’ils constatent beaucoup de critiques (Geiger, 2011). En effet, un propriétaire peut toujours affirmer qu’il changera volontiers le « comportement », le fonctionnement ou les tâches de son bot, si un nouveau consensus quant à une norme vient remplacer celui déjà en place et appliqué par un bot.

Ce lien maintenu entre les éditeurs et les bots s’expose aussi au travers de débats légaux touchant la propriété intellectuelle reliée au travail des bots. Bien qu’abstrait à première vue dans le cadre de notre sujet, il doit être envisagé que l’auteur du bot n’est pas toujours le pur auteur du contenu, puisque bien souvent, le bot agit en recueillant et compilant des informations provenant de diverses sources (Kennedy, 2009). Le bot doit être considéré second à son propriétaire, ce dernier étant lui-même second à Wikipédia quand il est question de la « propriété », tel le cas des employés d’une entreprise. Afin d’être apte à se voir reconnaître des droits d’auteurs, un pouvoir décisionnel et éditorial se révèle nécessaire (Kennedy, 2009). Autant dans le cas des bots que des articles, le droit d’auteur, s’il peut être question d’un tel principe dans Wikipédia (puisque’il y a absence de travaux originaux), repose sur l’employeur. L’apport du bot doit alors être perçu dans son processus collaboratif, c’est-à-dire dans sa performance. Celui-ci est donc entièrement considéré sous la responsabilité de son propriétaire et de Wikipédia en général.

6.4 Les bots et la formation des normes

Nos résultats font écho sur plus d'un point à la perspective dialectique développée par Freitag (2011). On y observe entre autres les trois pôles de ce processus lors de l'écriture des articles (figure 1.2 à la page 19) : un sujet (le contributeur), un objet (la norme), ainsi qu'une médiation (sa reproduction). En effet, une norme prend son sens et son autorité à chaque fois qu'un contributeur la respecte à travers ses agissements. Ces normes restent donc en place jusqu'à ce qu'un nombre suffisamment élevé d'individus les contestent, que ce soit parce que les pratiques, les conditions ou la vision du projet ont changé. Toutefois, cette donne semble passablement différente avec l'utilisation massive des bots. En effet, ceux-ci appliquent systématiquement des normes, ce qui semble avoir pour effet de déplacer les débats quant à la légitimité de ces normes. De pages spécifiquement dédiées à ces débats, les pages de discussion de bots deviennent le lieu même de la discussion, réalisant dans ce passage un lien entre la capacité du bot à bien réaliser sa tâche et l'appréciation générale que les contributeurs possèdent de cette norme. Bien entendu, le niveau de consensualité d'une norme rend son application par un bot d'autant plus facile à implanter et à maintenir. Ainsi, l'édition des articles, dans la perspective du contributeur, n'est plus tant régulé par son application individuelle des normes que par son application systématique et *a posteriori* par des bots, réactualisant donc la relation précédemment exposée.

Ce processus commence par la standardisation d'éléments organisationnels d'arrière-plan pour progressivement découler vers des normes dont l'impact sur le travail individuel est proportionnellement plus important. C'est ainsi que l'on arrive au conflit d'*Hagermanbot* : pour les promoteurs de bots ainsi que pour la communauté interne, la signature des commentaires n'est qu'une formalité, alors que la non signature complique les interactions et l'organisation commune. Cependant, il s'agit d'un choc pour tous les contributeurs qui ne font pas partie des discussions internes. Nous devons donc répondre à notre quatrième hypothèse par oui et non. Formulé ainsi : *Les agents logiciels occupant des tâches touchant l'imposition et le respect des normes de l'encyclopédie seront perçus négativement par la communauté, puisqu'ils participent activement à la stabilisation dans le contenu de Wikipédia et donc, à une perte d'autonomie des utilisateurs en général*, il est désormais évident que cette perception est fortement modulée selon la position sociale d'un individu au sein de l'encyclopédie. Effectivement, il s'avère que beaucoup de gens ne sont pas familiers avec l'émergence des normes (Antin et Cheshire, 2010), il peut donc être déstabilisant de constater qu'un bot peut agir sur sa propre action, allant même jusqu'à l'annuler (Geiger, 2011).

Supposant que le consentement, joint à une application systématique d'une norme, est central à sa légitimité (Suzor et Woodford, 2013), il est permis de nous questionner sur l'acceptation des normes dans Wikipédia. Particulièrement par leur application par des bots, quand une majorité des éditeurs de Wikipédia restent anonymes et/ou contributeurs spo-

radiques. La particularité du travail des bots dans ce cadre est justement de systématiser la reproduction des normes. Nous supposons donc que les critiques les plus sévères reçues par les bots, particulièrement pour les « policiers », sont principalement constituées de ces contributeurs sporadiques. Toutes ces raisons nous portent à considérer les bots comme des agents sociaux, qui appliquent systématiquement certaines normes, reproduisant une vision particulière du monde qu'est Wikipédia. Ceci implique donc d'établir une distance rendant la norme imperméable, jusqu'à un certain niveau, aux changements effectivement vécus par la communauté.

L'ensemble de ce processus nous porte à nous questionner sur les orientations générales de Wikipédia. Animée d'un idéal humaniste et peuplée d'érudits au cours de sa première phase, il nous semble évident que désormais, la structure de pouvoir se concentre davantage autour d'un paradigme technocentriste. Nous constatons en effet que la plupart des positions de pouvoir sont occupées par des individus avec des fortes habiletés techniques, ce qui influence fortement les orientations de l'encyclopédie. Bien que les normes se développent avec une participation large de la communauté, la puissance d'action des bots impose beaucoup d'efforts afin d'encadrer leur utilisation. Beaucoup d'efforts qui risquent d'apparaître tel un mur de contraintes infranchissable pour le contributeur qui n'est pas au fait de l'ensemble de la structure sociale de l'encyclopédie.

La critique de cette conception passe par une analyse de la balance des pouvoirs au sein de l'encyclopédie. Déjà, notons de nouveau que les bots de l'encyclopédie sont autorisés par le BAG. Ceux-ci autorisent spécifiquement certaines tâches, tout en obligeant les propriétaires d'être redevables auprès des membres de la communauté interne de Wikipédia, pour la durée totale de leur période d'activités. Avec le travail des bots, la bonne entente entre les divers acteurs et ceux-ci est donc d'une importance primordiale, le cas de *Yobot* nous fournissant un exemple non négligeable.

6.5 Les bots et la communauté; l'équilibre des pouvoirs

Avant de conclure ce mémoire, certaines questions restent donc encore en suspens. Bien qu'il ait été identifié dans la littérature que les individus sont plus susceptibles d'accepter le travail de bots si ces derniers sont en position de pouvoir (Hinds et collab., 2004), ce pouvoir doit leur être reconnu en premier lieu. Autant les préjugés que les attentes des contributeurs envers les bots ont un impact sur ces perceptions (Srivastava et collab., 2010; Mateas et Sengers, 1998; Paechter et collab., 2010; Paepcke et Takayama, 2010; Zdenek, 2003).

Par convention, les propriétaires se doivent d'être contributeurs actifs et reconnus de Wikipédia, avoir réalisé un minimum de 500 éditions et démontré qu'ils sont au fait des normes de

l'encyclopédie. Cela est confirmé par la prépondérance des administrateurs dans notre échantillon (environ 50 %). Le fait d'avoir des bots clairement contrôlés par des éditeurs favorise donc la reconnaissance de la position de ces premiers. Là-dessus, les nombreux propriétaires de bots qui partagent directement leurs pages de discussion avec leur bot renforcent cette perception de proximité.

Toutefois, l'utilisation de forces de travail automatisé, représenté par les bots, conduit à un impact massif sur le contenu de Wikipédia. Ceci conduit aussi à des questionnements touchant la balance des pouvoirs au sein de l'encyclopédie. En effet, bien qu'ils soient clairement perçus comme n'étant pas équivalents en termes de statut social, les bots de Wikipédia possèdent toutefois davantage de liberté d'action sur l'encyclopédie (en termes de pouvoir d'altération et de contrôle sur le contenu), qu'un contributeur enregistré non-administrateur (Niederer et van Dijck, 2010). Pour les utilisateurs, les bots semblent donc posséder *de facto* une prépondérance sur une grande partie des contributeurs (Niederer et van Dijck, 2010). La réalité est toutefois passablement différente, puisqu'il s'agit concrètement d'une forte augmentation de la capacité d'action des propriétaires. Toutefois, un bot n'est autorisé à avoir des droits d'administrateurs que si son propriétaire est lui-même un administrateur. En ce sens, notons qu'il semble exister un sentiment d'appropriation chez les propriétaires des tâches réalisées par leurs bots. On le voit par exemple par les inquiétudes manifestées par le propriétaire de *DumbBOT* face à *VWBot*, ce dernier semblant profiter de l'absence du premier pour récupérer une tâche particulière du bot. En réponse, le propriétaire de *VWBot* se veut rassurant : il ne s'agit que d'un bot de remplacement permettant de maintenir les tâches habituellement en place.

Dans ce contexte, il est intéressant d'observer comment ce conflit apparent de distribution des pouvoirs est abordé par la communauté. Une seconde inquiétude est de savoir si ce questionnement est effectivement résolu dans l'encyclopédie. Cette question n'est pas secondaire, puisque les deux tiers des nouveaux éditeurs reçoivent leur premier message d'un bot ou d'un logiciel d'assistance à l'édition (Niederer et van Dijck, 2010). Considérant la tendance à la baisse du taux de rétention des nouveaux contributeurs, ceci constitue une inquiétude majeure de la part de la communauté (Halfaker et collab., 2012). Cet aspect de la perception des bots ne devrait donc pas être négligé.

Avec les bots possédant des pouvoirs éditoriaux importants, la perception des usagers risque d'être sensible à la présence de contrôles sur l'action des bots. L'un des débats initiaux touchant l'acceptation des bots a résulté en l'adoption du gabarit « *no-bot* » en 2006. Consistant en un mécanisme de « *opt-out* », ce gabarit autorise un éditeur à empêcher un bot d'agir sur ses propres éditions ou sur des pages précises (Geiger, 2011 ; Halfaker et Riedl, 2012). Ces mécanismes apparaissent avoir rempli leurs objectifs. En effet, la quasi-absence de résultats significatifs entre nos catégories « *opt-in* » et « *opt-out* » suggère que ce critère ne représente plus un angle nous permettant d'interpréter les interactions entre les éditeurs et les bots.

Ceci infirme donc notre troisième hypothèse qui se formulait ainsi : *Parmi les différentes classifications catégorielles, le mode d'application des agents logiciels, allant de l'application automatique jusqu'aux agents logiciels fonctionnant uniquement par abonnement, va influencer l'appréciation que les éditeurs auront de ceux-ci.* La seule relation significative observée consiste en des utilisateurs se justifiant plus souvent de leurs actions quand le bot agit automatiquement. Cela peut indiquer que les éditeurs possèdent le désir de ne pas perturber le travail des bots. Il s'agit d'une reconnaissance aux bots d'une certaine autorité sur les tâches qu'ils réalisent. Comme il est observé dans l'établissement de la norme « *opt-in* », les propriétaires peuvent avoir la tendance à se déresponsabiliser du travail de leur bot en invoquant la possibilité pour un éditeur d'utiliser le gabarit « *no-bot* ». Confirmant ce phénomène et servant une mise en garde au même instant, l'administrateur Headbomb, évoque le problème en ces termes :

[...] Je vous rappelle aussi que la possibilité d'ignorer votre bot n'est pas une excuse pour rejeter les inquiétudes des utilisateurs concernant *WP :COSMETICBOT*, ni de créer une mentalité « Vous ne l'aimez pas ? Ignorez-le ! »

Headbomb,

sur le panneau d'affichage des propriétaires de bots,

14 avril 2012⁴.

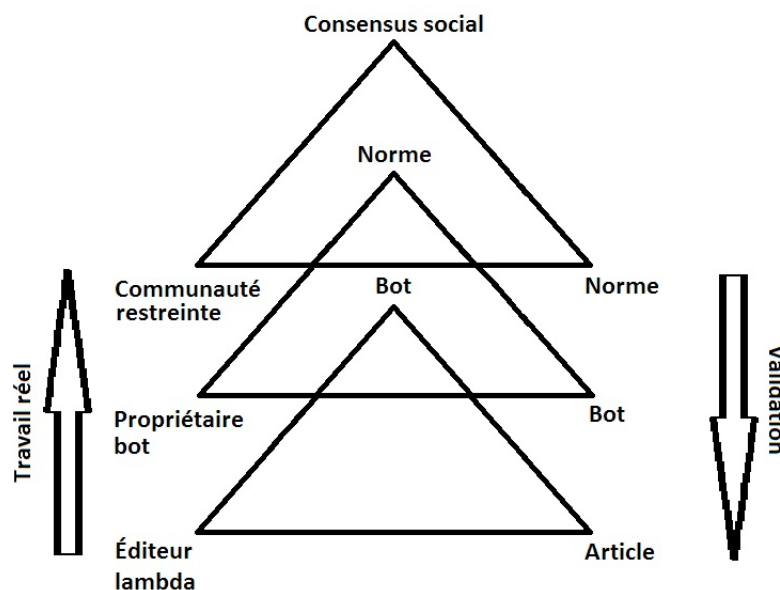


FIGURE 6.1 – **Processus dialectique de l'écriture d'un article.** Description des trois étapes principales régulant le travail des contributeurs sur les articles.

4. Traduction libre.

La défense des bots par les propriétaires et les administrateurs ne doit pas être perçue comme désintéressée. En effet, bien que la littérature identifie que les administrateurs perdent de l'importance en proportion du nombre total d'éditions (Kittur et collab., 2007b), la structure sociale s'est profondément renforcée durant les mêmes années. En ce sens, la montée des éditions réalisées par les bots devrait être rajoutée à celles réalisées par les administrateurs, puisque leur nature est très semblable. Les bots correspondent en effet à l'acceptation sociale de certaines normes par la communauté et ce sont les administrateurs qui sont responsables de leur application. Ces normes sont acceptées par la communauté, car elles reproduisent la représentation qu'elle se fait d'elle-même. Bien que nous n'ayons observé que peu de discussions de nature communautaire (ne touchant pas directement les bots et leur gestion) au sein du panneau d'affichage des propriétaires de bots, l'esprit de corps est fortement présent lors de la défense de leur intérêt collectif. Cette conception est représentée dans un triple processus de médiation, imagé dans la figure 6.1. Selon cette représentation, nous comprenons que la médiation, aussi considérée en tant que régulation, du travail de l'éditeur sur l'article est souvent réalisée par l'intermédiaire de bots. Dans un processus de deuxième niveau, ces bots constituent l'objet du travail de leur propriétaire. Dans les faits, ceux-ci se retrouvent donc dépositaires d'une traduction des normes de l'encyclopédie vers le travail des éditeurs. Enfin, ces mêmes normes sont l'objet d'intérêt de la communauté restreinte de l'encyclopédie. Ceux-ci prennent leur décision au travers d'une idée particulière qu'ils se font du consensus social en cours dans l'encyclopédie, vision interprétée et validée par les administrateurs. Nous exposons donc ici un processus fortement social qui percole ensuite vers un mouvement inverse en une série de mécanismes techniques, avec les bots comme fer de lance, venant encadrer le travail légitime des éditeurs de l'encyclopédie.

Pour beaucoup de nouveaux éditeurs, le premier moment qui les confrontera avec l'existence de normes sera lors de la suppression d'une de leurs premières éditions (Schneider et collab., 2012). « *Ceci risque d'aliéner les potentiels et nouveaux contributeurs tout en épuisant ceux existants (qui doivent trier, non seulement le vandalisme délibéré, mais aussi les contributions bien intentionnées, mais erronées), et gaspiller le temps des novices dans l'écriture et le temps des évaluateurs dans la suppression de contenu inapproprié.* » (Schneider et collab., 2012). En ce sens, les bots semblent en effet consister en une réponse de la communauté face à la montée massive de la participation de contributeurs qui ne sont pas suffisamment au courant des standards de l'encyclopédie (Butler et collab., 2008). Avec la régularité à laquelle les contributeurs intensifs viennent défendre les bots lors de conflits, la question mérite d'être posée. En ce sens, ce problème de rétention peut bel et bien sembler un fait accepté par la communauté. Difficile de voir autrement quand *ClueBot NG* vise spécialement les nouveaux contributeurs et les anonymes.

De plus, nous observons que certains propriétaires de bots utilisent les IRC pour leurs communications. Bien que nous estimons que cela renforce l'esprit de communauté, nous

observons qu'il y a toutefois un risque de prendre des décisions qui ne sont pas exposées et explicitées pour tous, ce qui est contraire au mode de fonctionnement caractéristique de Wikipédia. En effet, sauf dispositions spéciales, IRC est un mode de communication éphémère, au contraire des discussions systématiquement archivées propres à Wikipédia. Par la nature de ce médium, nous n'avons pas eu la possibilité d'analyser *a posteriori* le contenu des discussions qui y ont eu lieu. Toutefois, IRC est tout de même utilisé dans un cas précis afin d'autoriser à un propriétaire, Kumioko, de tester un bot sans qu'il ait reçu formellement l'autorisation du BAG. Ceci entraîna une forte réponse négative des autres propriétaires de bots. Suite à un débat houleux, il est décidé que Kumioko se voit enlever son droit d'utiliser des bots. Dans ce cas particulier, comme dans celui de Rich Farmbrough (dont les droits d'utilisation de bots ont été retirés au cours de l'année), ces propriétaires ont décidé d'agir sans consulter la communauté, croyant respecter le consensus en cours. La révocation de leurs droits n'est pas mineure puisque ces deux contributeurs en particulier possédaient l'autorisation d'utiliser des bots pour un total cumulé d'une trentaine de tâches différentes. À la suite de ces exemples, il est observé que la récupération des tâches désormais vacantes nécessite de passer au travers d'un nouveau processus d'autorisation individuelle des tâches. Ceci est fait dans l'objectif de spécifiquement renforcer le lien entre une tâche et un propriétaire de bot, explicitant d'autant plus la responsabilité de celui-ci à respecter les consensus en cours dans l'encyclopédie. Dans le cas simple d'un transfert de propriétaire pour cause de « retraite », le processus qui y est observé est beaucoup moins rigide.

Dans tous les cas, le rôle structurel des bots semble largement reconnu par la communauté. Cela se note entre autres par le fort nombre d'administrateurs et de veilleurs de pages venant défendre les bots dans les discussions. Alors que l'on peut supposer que des propriétaires fuient leurs responsabilités en se cachant derrière le travail des bots, ceux-ci mobilisent très souvent un argumentaire invoquant la légitimité sociale des normes appliquées afin de justifier leurs actions. De même, certains veilleurs de pages vont rappeler aux critiques de ne pas s'adresser aux bots lors de mécontentement :

Bonjour. Premièrement, *ClueBot* est un bot, non pas un humain, et ne peut donc pas répondre aux messages. En second, il n'y a pas de place pour le vandalisme dans Wikipédia. L'édition que vous avez fait à l'article touchant Yoko Ono est du pur vandalisme, rien à voir avec la correction de faits. Pour citer Jimmy Wales, pourquoi ne changez-vous pas afin de devenir un éditeur de bonne foi. *ClueBot* ne peut vous bloquer, mais si vous continuez à éditer de la façon que vous avez édité l'article de Yoko Ono, vous allez finir bloqué indéfiniment d'éditions par l'un de nous, administrateurs.

5 albert square,
sur la page de discussion de *ClueBot NG*,
21 :25, 2 novembre 2012.

Dans ce cas précis, le rappel à « l'ordre » est réalisé par un autre administrateur, renforçant d'autant plus la structure sociale de droits et pouvoirs en place dans l'encyclopédie. De plus, notons que si l'une des éditions du bot est contestée par un contributeur, le bot n'est pas autorisé à la contester à son tour. Ceci indique donc clairement que le jugement final reste aux mains de l'éditeur, réduisant non seulement le pouvoir actuel possédé par le bot, mais aussi nombre de critiques potentielles qu'ils peuvent recevoir.

En fait, nous observons davantage un désir de la communauté de voir des bots efficaces que de les contester par principe, comme le montre la proportion d'interventions touchant l'identification de faux positifs (21.1 %) et les questionnements sur le fonctionnement de leurs tâches (17.3 %). Nous observons ainsi que la communauté de Wikipédia n'est pas absente de la reproduction des représentations symboliques que sont les normes, comme pourrait le laisser entendre une lecture simpliste de la structure sociotechnique. Au contraire, il nous est permis de constamment observer la communauté se questionner sur la pertinence des normes et encore plus de leur application. La communauté est pleinement consciente des possibilités offertes par un système sociotechnique. Elle en profite constamment afin d'ajouter des fonctions *ad hoc* à celles existantes. Absente du cadre initial, ces normes leur permettent d'actualiser et de matérialiser la vision commune de la communauté, créée au fil des expériences, face à un environnement extérieur généralement jugé hostile à la perpétuité du projet. En ce sens, nous pouvons répondre par l'affirmative à notre dernière hypothèse : *Dans Wikipédia, les agents logiciels représentent la formalisation de sa communauté originellement informelle, en traduisant les normes sociales qu'elle adopte en politiques et règles de conduite, qui sont ensuite appliquées à même l'infrastructure de Wikipédia.*

Nous concluons donc cette discussion en actualisant le schéma exposé en figure 6.1. En effet, ce schéma explique bien le processus que traverse l'éditeur lambda. Ainsi donc, nous retrouvons un processus relativement complexe laissant bel et bien le contrôle de Wikipédia à ses utilisateurs, mais seulement chez les plus actifs. Ce processus semble mettre en danger la pérennité de l'encyclopédie, puisque cette communauté interne tend à diminuer face à une utilisation toujours plus grande de Wikipédia par le public. C'est donc dire qu'effectivement, le poids et l'inertie de Wikipédia pèsent de plus en plus sur le travail individuel. Inertie qui ressemble fortement à un déterminisme technique pour les contributeurs concernés, rendant ainsi bien difficile pour eux d'accepter la maxime d'accueil de Wikipédia : « *Bienvenue à Wikipédia : le projet d'encyclopédie libre que « vous » pouvez améliorer.* ».

Conclusion

L'Encyclopédie peut aisément s'améliorer ; elle peut aussi aisément se détériorer. Mais le danger auquel il faudra principalement obvier, & que nous aurons prévû, c'est que le soin des éditions subséquentes ne soit pas abandonné au despotisme d'une société, d'une compagnie, quelle qu'elle puisse être. Nous avons annoncé, & nous en attestons nos contemporains & la postérité, que le moindre inconvénient qui put en arriver, ce seroit qu'on supprimât des choses essentielles ; qu'on multipliât à l'infini le nombre & le volume de celles qu'il faudroit supprimer ; que l'esprit de corps, qui est ordinairement petit, jaloux, concentré, infectât la masse de l'ouvrage ; que les Arts fussent négligés ; qu'une matiere d'un intérêt passager étouffât les autres ; & que l'Encyclopédie subît le sort de tant d'ouvrages de controverse.

« Encyclopédie », *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers*, Diderot, 1755, p.635

Ce mémoire a constitué une occasion unique d'observer l'interaction entre une communauté collaborative en ligne et les agents logiciels. Ceci en ayant un regard particulier sur les stratégies qu'elle a utilisées pour tirer avantage du cadre technique à sa disposition. À cette fin, nous avons utilisé un terrain fort intéressant pour notre sujet, soit l'incontournable Wikipédia. Ce terrain nous était particulièrement pertinent puisque Wikipédia consiste en l'un des projets collaboratifs en ligne le plus dynamique à ce jour. De plus, cette encyclopédie possède une approche unique dans sa gestion, par la communauté même qui la compose, des dispositifs techniques sous-tendant le projet. Un peu plus de 10 ans après sa création, ce terrain de jeu aux règles minimalistes a en effet évolué vers une communauté dotée d'une identité complexe possédant une vision désormais explicite de ce à quoi doit ressembler l'encyclopédie. Au cœur de ce processus, les agents logiciels, ici appelés bots, représentent des acteurs incontournables pour l'encyclopédie. Comment ces derniers sont-ils perçus par la communauté ? Est-ce en simples automates, leur permettant de réaliser facilement des tâches auparavant considérées comme fastidieuses ? Ou entraînent-ils des considérations plus complexes ?

Pour les contributeurs réguliers, les bots sont souvent perçus comme des outils facilitant ou permettant la réalisation de certaines tâches qui épuiserait autrement la communauté, par l'augmentation soutenue du temps nécessaire à administrer l'encyclopédie. Accumulant

les succès, l'utilisation des bots glissera de la réalisation de tâches simples, mais répétitives, jusqu'à l'application systématique de normes consensuelles. Toutefois, l'impact de ces deux types de tâches est bien différent pour l'encyclopédie. Quand, dans le premier cas, il est plutôt question de standardisation du contenu, la deuxième forme suppose une détermination au second degré des comportements et contributions légitimes des éditeurs de l'encyclopédie. Second degré, car il ne s'agit pas ici d'une action sur l'encyclopédie, mais bien d'une action performative sur la possibilité d'actions d'autrui. Pour les contributeurs membres actifs de la communauté ou d'une des multiples sous-communautés de Wikipédia, l'utilisation de ces outils semble fortement apprécié. Quant à lui, le contributeur sporadique, éditeur lambda et souvent anonyme, confronté à un cadre technique et proactif régissant les contributions acceptées, semble souvent abandonner tout simplement sa participation. En effet, l'identité chez un groupe est à double tranchant : une forte identité de soi est corollaire à une exclusion renforcée de ceux ne partageant pas ces mêmes caractéristiques identitaires.

C'est à cette fin que nous avons posé la question suivante : *comment les contributeurs perçoivent-ils l'action des bots sur leur propre travail et, dans un second temps, comment le travail des bots s'intègre-t-il dans le fonctionnement et la structuration de la communauté ainsi que dans le contenu de Wikipédia ?* À cette fin, nous avons cherché à atteindre quatre objectifs. Nous avons voulu quantifier et qualifier les interactions que les contributeurs participants à l'écriture des articles ont avec les bots de Wikipédia. Il nous était aussi d'intérêt d'identifier les principales fonctions des bots afin d'en réaliser différentes classifications catégorielles. Puis, il a s'agit d'identifier si certaines fonctions réalisées par les bots provoquent des réactions différenciées chez les utilisateurs. Enfin, il a été question d'identifier la signification que prend le travail des bots pour les éditeurs de Wikipédia ainsi que pour leurs opérateurs (propriétaires de bots). À cette fin, nous avons réalisé un aller-retour entre les données disponibles et la littérature pertinente, dans une approche non sans rappeler la théorisation ancrée.

Nos résultats confirment que les bots sont principalement perçus comme des collaborateurs asservis et d'une potentielle grande valeur pour l'encyclopédie. Ils constituent le principal rempart afin d'assurer l'intégrité de la base de données. Bien que les bots occupent des tâches éditoriales et managériales importantes et répétitives, ils ne sont pas perçus comme des entités indépendantes. Le contrôle humain transpire en bien des points. Entre autres, les bots et leurs propriétaires se voient fortement réprimandés dès que ce contrôle est perçu déficient par la communauté. Ce phénomène est interrelié dans les trois catégories de membres que nous avons identifiées : la sous-communauté des propriétaires de bots, la communauté interne large de Wikipédia, ainsi que les contributeurs lambda, occasionnels et spontanés, que l'on pourrait qualifier d'environnement social externe de l'encyclopédie. Ce phénomène passe par l'incompréhension première de l'utilisateur lambda, au désir de la communauté interne d'avoir des outils d'assistance à la fois efficaces et efficients, jusqu'au questionnement récurrent des propriétaires de bots quant à savoir s'ils agissent dans le cadre normatif de l'encyclopédie. La

cohérence des réactions semble indiquer la présence d'une communauté mature qui partage désormais de nombreuses règles de vie commune.

Le fait le plus probant de cette recherche semble donc être le déplacement observé quant au débat sur la norme ainsi que le renforcement du pouvoir des *power users* dans la communauté. Loin de l'initiale « interprétation créative des règles » ainsi que de la perception d'un terrain de jeu libre de règles souvent retrouvée à propos de Wikipédia, nous constatons au contraire une limitation évidente des possibilités de participation chez les nouveaux utilisateurs. Entre autres, la récupération de diverses tâches par les bots, ainsi que par la nature systématique de leur travail, transforme la nature informelle de la communauté de Wikipédia. En ce sens, le déterminisme technique qui s'impose aux nouveaux éditeurs peut sembler infranchissable. Cependant, et contrairement aux lectures technocentristes simples, nous avons repéré, à tous les niveaux, un contrôle des contributeurs dans la détermination des tâches déléguées aux bots ainsi qu'une évaluation rétroactive de leur réalisation. Même avec la récupération de toujours plus nombreuses tâches, il ne s'agit pas d'un abandon de celles-ci par la communauté. Toutefois, cette détermination et supervision se retrouvent essentiellement réduites à la communauté interne, limitant *de facto* la capacité d'action de la multitude de contributeurs lambda. De plus, les propriétaires de bots ne sont pas en dehors des jeux de pouvoir au sein de l'encyclopédie. De par leurs capacités fortement techniques et organisationnelles, ils en sont même avantagés sur bien des points. Toutefois, notre étude nous a permis d'identifier, chez ceux-ci, un fort sentiment d'être les dépositaires de l'application de consensus sociaux précis.

Suite aux découvertes de notre étude, il n'est pas possible de parler de quelconque calcification de l'encyclopédie, évoquée dans notre troisième hypothèse : *les bots occupant des tâches touchant l'imposition et le respect des normes de l'encyclopédie seront perçus négativement par la communauté, puisqu'ils participent activement à la stabilisation dans le contenu de Wikipédia et donc, à une perte d'autonomie des utilisateurs en général*. Au contraire, le maintien de cette gamme d'outils et fonctions autonomisés (sic) est loin d'être aisé, la communauté interne ayant développé beaucoup d'attentes envers les bots. Cela témoigne encore une fois de la force de la communauté des propriétaires de bots, qui doit porter sur leurs épaules (à défaut de la maison-mère qu'est MediaWiki), la reproduction d'artefacts sociaux et techniques toujours plus nombreux. En contrepartie, leur influence sur l'encyclopédie s'en retrouve significativement augmentée. Ces affirmations font que nous tenons à limiter le constat d'autres auteurs quant à la rigidité du système sociotechnique en place dans Wikipédia, ayant observé des bots et administrateurs soucieux de répondre aux utilisateurs. Autant dans l'objectif d'un laminage intemporel de connaissances que dans le maintien d'une structure de normes, Wikipédia est loin d'être un processus automatisé hors de l'atteinte de ses participants.

Avec la montée en puissance des communautés de production collaborative, l'utilisation des agents logiciels s'ouvre sur de nouveaux horizons dont les conséquences ne sont pas encore

clairement exposées. De toute évidence, les conclusions de ce mémoire ne sauraient s'appliquer à des communautés dont la gestion est à caractère privé, tels Youtube ou la majorité des grands sites d'informations. Ces organisations observent souvent chez leur communauté d'utilisateurs un fort caractère chaotique et irrespectueux des interactions encourues. Ce phénomène n'est sans doute pas étranger à la conscience des participants de n'avoir aucun contrôle sur les règles de fonctionnement de celles-ci. Dans ces plateformes monétisant les contributions des individus qui y participent, peu d'espace est laissé à la communauté pour agir sur le cadre technique qui les limite. Toutefois, ce sentiment de rejet semble aussi être vécu de plus en plus par les nouveaux contributeurs de Wikipédia.

Il est nécessaire de se questionner sur la capacité à long terme de l'encyclopédie à maintenir une communauté active et dynamique. En effet, le temps consacré à des tâches administratives, de même qu'à simplement conserver les outils et fonctions qui ont émergé des pratiques sociales, reste en constante hausse. Alors, malgré cette acceptation des bots, nous notons que la communauté vit désormais beaucoup de débats normatifs en réaction à l'immense capacité d'action de ces bots. Étant donné que l'effort focal vise d'autant plus la saine gestion de ces agents logiciels, la difficulté de retenir de nouveaux usagers peut être causée par la perte d'attention touchant la cohésion de la communauté. Si un équilibre n'est pas atteint, et aucun signe ne va en ce sens pour l'instant, beaucoup d'éditeurs risquent d'oublier les raisons les ayant menés à être actifs sur Wikipédia.

Annexe

.1

Tableau résumé de la classification des bots

Bot	Date d'entrée en fonction	Nombre tâches actives à la fin de l'année 2012	Nombre d'éditions	Controle	Localisation des éditions	Autonomie	Fonction	Lieu de la page de discussion	Application de la tâche	Idéal-type
ClueBot NG	25 octobre 2010	1	935971	Administrateur	Avant-plan	Haute	Contraint	Bot	Op-out	Policier
EmausBot	4 mars 2010	1	706080	Membre enregistré	Arrière-plan	Basse	Facilite	owner	Op-in	Autre
Cydebot	7 avril 2006	2	575172	Administrateur	Arrière-plan	Basse	Facilite	two	Op-in	Autre
WP 1.0 bot	4 février 2007	1	418365	Membre enregistré	Arrière-plan	Basse	Facilite	Bot	Op-in	Autre
AnomieBOT	7 aout 2008	41	334684	Administrateur	Avant-plan	Basse	Facilite	Bot	Op-out	Autre
Legobot	13 juin 2008	19	319640	Administrateur	Arrière-plan	Basse	Facilite	Bot	Op-out	Servant
TedderBot	9 juillet 2009	5	208418	Administrateur	Arrière-plan	Basse	Facilite	owner	Op-in	Autre
SineBot	24 juillet 2007	1	207644	Administrateur	Avant-plan	Haute	Contraint	Bot	Op-out	Policier
AvicBot	21 mars 2010	9	199395	Membre enregistré	Arrière-plan	Basse	Facilite	owner	Op-out	Servant
RjwilmsiBot	11 novembre 2009	7	161430	Membre enregistré	Avant-plan	Haute	Contraint	owner	Op-out	Autre
AAlertBot	30 novembre 2010	1	149325	Membre enregistré	Avant-plan	Basse	Facilite	bot	Op-in	Autre
BG19bot	3 mai 2011	6	129163	Administrateur	Avant-plan	Basse	Facilite	owner	Op-out	Autre
BattyBot	30 décembre 2011	13	122913	Membre enregistré	Arrière-plan	Basse	Facilite	two	Op-out	Servant
Xqbot	16 novembre 2008	4	122391	Membre enregistré	Arrière-plan	Basse	Facilite	owner	Op-out	Servant
MiszaBot III	4 mars 2007	1	114872	Administrateur	Avant-plan	Basse	Facilite	bot	Op-in	Autre
FrescoBot	20 février 2009	10	114516	Membre enregistré	Avant-plan	Basse	Facilite	owner	Op-out	Autre
EdwardsBot	16 juin 2009	1	106651	Membre enregistré	Avant-plan	Basse	Facilite	bot	Op-in	Autre
DPL bot	2 octobre 2011	4	76926	Administrateur	Avant-plan	Haute	Contraint	two	Op-out	Policier
CommonsDelinker	23 septembre 2006	1	76037	Membre enregistré	Avant-plan	Haute	Contraint	Bot	Op-in	Autre
RussBot	26 mai 2005	5	75702	Administrateur	Arrière-plan	Basse	Facilite	owner	Op-in	Autre
SporkBot	20 mai 2010	3	74286	Administrateur	Arrière-plan	Basse	Facilite	owner	Op-out	Servant
XLinkBot	10 janvier 2008	1	64435	Administrateur	Avant-plan	Haute	Contraint	Bot	Op-out	Policier
DASHBot	17 decembre 2009	9	63759	Membre enregistré	Arrière-plan	Haute	Contraint	owner	Op-out	Autre
28bot	31 mars 2011	4	60735	Administrateur	Avant-plan	Haute	Contraint	bot	Op-out	Policier
The Anomebot2	11 aout 2006	1	60389	Administrateur	Avant-plan	Haute	Contraint	owner	Op-out	Policier
RFC bot	12 janvier 2007	1	56410	Administrateur	Avant-plan	Basse	Facilite	owner	Op-in	Autre
Citation bot	12 mars 2008	9	47323	Administrateur	Avant-plan	Haute	Contraint	two	Op-out	Policier
Thehelpfulbot	5 février 2009	11	32789	Administrateur	Arrière-plan	Basse	Facilite	Owner	Op-out	Servant
Amalthea (bot)	8 avril 2011	6	28006	Administrateur	Arrière-plan	Basse	Facilite	owner	Op-out	Servant
DYKUpdateBot	21 fevrier 2010	1	26381	Administrateur	Arrière-plan	Haute	Facilite	owner	Op-out	Autre
VeblenBot	13 mars 2007	3	25576	Administrateur	Arrière-plan	Basse	Facilite	owner	Op-out	Servant
Snotbot	8 janvier 2011	7	22974	Administrateur	Avant-plan	Basse	Contraint	owner	Op-out	Autre
ImageTaggingBot	24 novembre 2007	1	21174	Membre enregistré	Arrière-plan	Basse	Facilite	bot	Op-out	Servant
JL-Bot	3 juillet 2007	4	20238	Membre enregistré	Arrière-plan	Basse	Facilite	bot	Op-out	Servant

Tableau résumé de la classification des bots

Addbot	1 mars 2008	13	19185	Administrateur	Arrière-plan	Basse	Facilite	owner	Op-out	Servant
MadmanBot	9 juin 2007	2	17626	Administrateur	Avant-plan	Haute	Contraint	owner	Op-out	Policier
EarwigBot	28 avril 2009	18	17234	Administrateur	Avant-plan	Haute	Contraint	Owner	Op-out	Policier
Scsbot	9 juin 2006	1	15512	Membre enregistré	Arrière-plan	Basse	Facilite	two	Op-in	Autre
HostBot	4 avril 2012	2	13969	Membre enregistré	Avant-plan	Basse	Facilite	owner	Op-in	Autre
SDPatrolBot	22 mai 2009	3	12865	Membre enregistré	Avant-plan	Haute	Contraint	two	Op-out	Autre
SuggestBot	1 décembre 2005	6	11903	Membre enregistré	Avant-plan	Haute	Facilite	bot	Op-in	Autre
GimmeBot	7 janvier 2007	3	9262	Membre enregistré	Avant-plan	Basse	Facilite	Owner	Op-out	Autre
Mr.Z-bot	29 avril 2007	4	9116	Membre enregistré	Avant-plan	Basse	Facilite	Owner	Op-in	Autre
BernsteinBot	10 mars 2009	1	8358	Membre enregistré	Arrière-plan	Basse	Facilite	Bot	Op-out	Servant
OgreBot	25 avril 2009	1	7246	Administrateur	Avant-plan	Basse	Contraint	owner	Op-out	Autre
KingpinBot	22 aout 2009	5	5017	Membre enregistré	Avant-plan	Basse	Facilite	two	Op-in	Autre
RscprinterBot	7 novembre 2011	4	3410	Membre enregistré	Avant-plan	Haute	Contraint	bot	Op-out	Autre
Rick Bot	13 avril 2007	6	3227	Membre enregistré	Avant-plan	Basse	Facilite	owner	Op-out	Autre
VWBot	3 mai 2010	6	2496	Membre enregistré	Arrière-plan	Basse	Facilite	owner	Op-out	Servant
UcuchaBot	29 mars 2011	5	1828	Administrateur	Avant-plan	Basse	Contraint	owner	Op-out	Autre

.2

Log ratio des discussions	Propriétaire		Fonctionnement		Localisation des éditions		Application		Complexité de la tâche	
	Non-Admin (N=24)	Admin (N=26)	Facilite (N=32)	Contraint (N=17)	Avant-plan (N=29)	Arrière-plan (N=21)	Inscription (N=15)	Automatique (N=35)	Basse (N=34)	Haute (N=16)
(1) Remerciements	0,0401± 0,0671	0,1149± 0,0234	0,058± 0,0133	0,1231± 0,0325	0,1027± 0,023	0,0463± 0,0116	0,0446± 0,0179	0,0937± 0,0191	0,0580± 0,0133	0,1237± 0,0342
(2) Demandes d'aide ou de conseils	0,0869± 0,016	0,1223± 0,0181	0,0937± 0,0133	0,1209± 0,0212	0,1308± 0,0158	0,0701± 0,0172	0,1243± 0,0226	0,0972± 0,0147	0,1028± 0,0159	0,1107± 0,0192
(3) Utilisateurs informant d'une action entreprise	0,0432± 0,0562	0,0966± 0,0176	0,0488± 0,0108	0,1168± 0,0219	0,0891± 0,0164	0,0459± 0,0124	0,0338± 0,0122	0,0868±0,0144	0,0539± 0,0113	0,1071± 0,0235
(4) Erreurs ou faux-positifs	0,1341± 0,0138	0,1543± 0,0195	0,1413± 0,0135	0,1548± 0,0244	0,1522± 0,0178	0,134± 0,0151	0,1698± 0,0213	0,1338± 0,0145	0,1348± 0,0139	0,1653± 0,0234
(5) Suggestions d'améliorations	0,0983± 0,0161	0,1137± 0,0162	0,0966± 0,0152	0,1309± 0,0205	0,1236± 0,0158	0,0825± 0,015	0,119± 0,0196	0,1009± 0,014	0,0984± 0,0126	0,1231± 0,0234
(6) Questions sur le bots ou les éditions	0,1251± 0,0128	0,1561± 0,0198	0,1224± 0,0129	0,1745± 0,0237	0,1609± 0,0166	0,114± 0,016	0,1431± 0,0181	0,1404± 0,0155	0,1316± 0,0139	0,1616± 0,0232
(7) Critiques générales	0,0693± 0,0153	0,1083± 0,0181	0,0667± 0,0127	0,1335± 0,023	0,1127± 0,0168	0,0577± 0,0151	0,0865± 0,0186	0,0909± 0,0156	0,0697± 0,0125	0,1319± 0,0247
(8) Contestations spécifiques	0,0552± 0,0173	0,1139± 0,0178	0,0529± 0,0124	0,1525± 0,0224	0,0999± 0,0186	0,0661± 0,0167	0,0553± 0,0154	0,0988± 0,017	0,0682± 0,0143	0,123± 0,0234
(9) Justification	0,0113± 0,0535	0,0545± 0,0176	0,0109± 0,0075	0,0787± 0,024	0,0488± 0,016	0,0129± 0,0102	0± 0	0,0482± 0,0143	0,0135± 0,0079	0,0768± 0,254

.3



Interacting with bots online: Users' reactions to actions of automated programs in Wikipedia



Maxime Clément, Matthieu J. Guitton*

Faculty of Medicine, Laval University, Quebec City, QC, Canada
 Institut Universitaire en Santé Mentale de Québec, Quebec City, QC, Canada

ARTICLE INFO

Article history:

Keywords:

Community of practice
 Human–machine interactions
 Online collaboration
 Sociotechnical systems
 Virtual community
 Wikipedia

ABSTRACT

With the drastic rise of social media, large-scale collaborative online projects such as Wikipedia are now dealing with incredible large amount of data. This growth forces the community to provide tremendous efforts in order to maintain the accuracy and structure of the database. To deal with such amounts of data, Wikipedia users have developed automated programs – bots – to help them to do some of the maintenance tasks. However, it is unclear how human users react to the actions of these bots. Based on a corpus of 6528 interventions (2353 different discussions) of users on talk pages of 50 bots active on English-language Wikipedia pages between January 4, 2012 and January 2, 2013, we analyzed the reactions of users depending of the characteristics of the bots' actions. Bots activity was strongly associated with the functioning Wikipedia internal community. Bots whose activity was mostly related to the work of other users (e.g. high degree of constraint or visibility) elicited more responses. By combining the different characteristics of the bots, we were able to define two opposite "ideotypes" of bots with distinct behavior: "servant bots" which mainly do repetitive and laborious work instead of human users; and "policing bots" proactively enforcing Wikipedia's guidelines and norms, which elicited more polarized responses from users (either negative or positive rather than neutral). Our results demonstrated a surprisingly high level of acceptance of bots, modulated by differential reactions in function of the actual behavior of the bots.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The drastic increase of the size of social networks and virtual communities over the last decades has deeply affected the way humans interact online (Dunbar, 2012). Indeed, the global propensity of users for collaboration – emerging as a need to maintain efficacy – increases as the size of the virtual community increases (Yang et al., 2013). While optimizing strategies of communication within the community can help maintain social density to an extent (Guitton, 2015), ultimately, the growth of a virtual community forces its members to reduce their individual capacity of action to favor the emergence of managerial operative frameworks (Yang et al., 2013).

One of the most widely used solution that online communities have found to solve this issue was the development of specific tools

which could provide assistance of human users – whether standard users or system administrators. These tools often take the form of automated programs able to operate directly in the online media, referred as "bots". Internet users have long been exposed to bots, whether they were used for entertainment in IRC, supporting the community by doing administrative work, or gathering and providing particular sets of data to users. While conversational bots are easily identified as such and are mainly used for mundane purposes, bots having explicit duties have an important role in structuring both the information available for the community, and the community itself. Indeed, their programmed scripts make them act on a systematic basis, whether the action of the bot involves material added by the bot's owner or by other users, and, more importantly, whether the human users are aware of not of the action of the bot. Therefore, due to their function and despite the fact that bots obviously do not have any form of consciousness, they strongly contribute to automate and enforce rules. In a limited community, the actions of bots would be easily identified, and the bot's owner would be considered accountable for the bot's actions. Any malfunction of the bot would immediately trigger a human reaction; the bots being then perceived as mere tools.

* Corresponding author at: Institut Universitaire en Santé Mentale de Québec, 2601 Chemin de la Canardière (F-6500), Quebec City, QC G1J 2G3, Canada. Tel.: +1 418 663 5747; fax: +1 418 663 5971.

E-mail address: matthieu.guitton@fmed.ulaval.ca (M.J. Guitton).

The situation however gets considerably more complex for larger communities. Indeed, with the considerable increase of social media and online collaborative communities and projects, the interactions between human users and bots are becoming more and more common. More than a mere increase of the magnitude of interactions, the actions of the bots over the online material increases exponentially, and in any case considerably faster than the actions of human users, due to the automated characteristics of bots' actions. Therefore, modern bots are pushing the interactions to a brand new level. In this emerging reality, the question of how human users perceive bots, and perceive the interactions with them, becomes central. More specifically, while users tend to always perceive humans as superior than bots in computer-mediated communication (Aharoni & Fridlund, 2007; Edwards, Edwards, Spence, & Shelton, 2014; Lortie & Guitton, 2011; Mowbray, 2014), the questions remains of what kind of reactions and perceptions could be expected from human users confronted with massive actions of bots, and with the need to cooperate with them more tightly.

Among the different emerging communities making a heavy use of bots, Wikipedia represents a model of choice. Wikipedia is a community-based collaborative encyclopedia, in which users can implement knowledge in a massively shared process (Fallis, 2008; Giles, 2005; Loveland & Reagle, 2013; Zlatic, Bozicevic, Stefancic, & Domazet, 2006). With the constant implementation of its content, Wikipedia is dealing with incredible large amount of data, which forces the Wikipedia community to provide tremendous efforts in order to maintain the accuracy and structure of the database (Niederer & van Dijck, 2010). To deal with such amounts of data, Wikipedia users have developed bots to help them to do some of the maintenance tasks (Geiger, 2011; Halfaker & Riedl, 2012; Müller-Birn, Herbsleb, & Dobusch, 2013; Niederer & van Dijck, 2010). With approximately 400 bots in 2010 (Niederer & van Dijck, 2010), 700 bots by September 2011, and a total of 872 bots were referenced on English Wikipedia pages on January 2, 2013, the number of bots on Wikipedia has been constantly growing. While bots were responsible for only 3–5% of the edits of the English version of Wikipedia in 2005–06, they accounted for 16.33% in 2009 (Geiger, 2009), and above 20% of the total number of editions for English Wikipedia pages in the year 2012, as compiled using Wikipedia monthly statistics. Furthermore, two thirds of new Wikipedia users receive their first message from a bot or a semi-automated algorithm (Geiger, 2014), and therefore, bots impact the retention rates of new users one way or another.

In the context of the study of human–bots interactions in virtual spaces, Wikipedia bots represent an excellent model for several reasons. First, the combination of the size of the Wikipedia community, of the sheer number of Wikipedia bots, of the size of the corpus of texts which can be edited either by humans or by bots, and of the number of actual editions performed daily on Wikipedia results in a considerable amount of potential interactions. Second, and maybe more important, is the extremely prominent social factor underlying the process of knowledge accumulation, and therefore, the implications of human–bot interactions. Indeed, the social dimension of Wikipedia is still extremely prominent, since implementation of knowledge heavily relies on community-based decision processes and “consensus” (Jahnke, 2010; Viégas, Wattenberg, & McKeon, 2007). In other words, the decision processes in Wikipedia are in fact based on social-related factors (Black, Welser, Cosley, & Degroot, 2011; Chang & Chuang, 2011). However, while clearly perceived as not equivalent in terms of social status, the Wikipedia bots nonetheless have considerably more power in terms of their action over the content of Wikipedia (power of alteration and control over the content) than a standard registered, non-administrator user (Niederer & van Dijck, 2010). Thus, the actions of bots could bypass the conventional decision

process. In this context, the massively increasing actions of bots over Wikipedia content cannot be unnoticed by the community members and Wikipedia users, making Wikipedia a model of choice to study the interactions between bots and humans in an ecological large-scale setting, as well as the perception of the bots by the users.

Based on the analysis of comments on the talk pages of bots active in English Wikipedia, this study aims to characterize the interactions between users and bots. By aggregating different individual characteristics of bots, we were able to define opposite, yet complementary global profiles of bot behavior, and to demonstrate differential reactions of humans in function of these behavioral bot phenotypes. Our results lead to a better understanding of the way bots are perceived by the community, and demonstrate the importance of the reminiscence of human control over bots in their acceptance by others.

2. Material and methods

2.1. General protocol

The perception of bots by human users was evaluated on English Wikipedia active bots, via the analysis of the posts left on the bots' talk pages by users. Given that each individual language version of Wikipedia is supported by a slightly different community, the exact set of rules and conduct may slightly vary across languages. Hence, we focused on the mainstream and largest Wikipedia community, the English-language one. Bots were defined as automated software agents that perform algorithmically defined tasks involved with editing, maintenance, and administration, which can be interpreted as software tools or as managerial protocols in regards to the infrastructure (Geiger & Ribes, 2010; Müller-Birn et al., 2013; Niederer & van Dijck, 2010). The owner of a given bot is the registered Wikipedia user accountable for the actions of this bot. Each bot has a talk page, which is a background page dedicated specifically for discussions with or about the bot. In some cases, the bot's talk page is merged with the talk page of its owner. An edit consists in an edition or modification made by a user or a bot on the mainspace of Wikipedia, i.e. on Wikipedia articles. Instead, a post refers to an edit made on a talk page. The direct succession of posts concerning the same topic of interest in a talk page forms a discussion. The internal community of Wikipedia gathers the Wikipedia users who are editing with some regularity and who are at least partially aware and often part of the internal debates taking place in the collaborative project. Finally, talk page stalkers (sometimes referred as “WikiJaguar” or talk page watchers) are internal community members who survey particular talk pages (in the context of the present study, bots' talk pages) and act by answering or adding inputs to discussions on these pages, in which there were not initially directly involved.

2.2. Bots selection

Bots active on English Wikipedia pages between January 4, 2012 and January 2, 2013 were surveyed. Data related to bot activities were gathered from the “bots by number of edit” page of Wikipedia, which is automatically maintained by BernsteinBot, from the edit counts extracted by Toolserver (a specific software suite used by Wikipedia editors). Bot activity was assessed by calculating the difference between the number of edits done by a given bot between January 2, 2013 and January 4, 2012. A total of 170 active bots were thus identified. Pages that were directly linked to the bot talk page in order to let users post comments were retrieved, including, when relevant, the bot owner's personal talk page.

Only bots with at least 1000 edits and 12 user posts on their talk page (or the related pages) during the period of interest were

considered. Bots whose main contributions were in another language than English, or whose edits were mainly made in another language version of Wikipedia were excluded. When several bots were sharing the same talk page, only the bot with the higher edit count was considered. This resulted in a total of 51 bots. However, one of the bots (Yobot, ranking first in terms of number of edits between January 4, 2012 and January 2, 2013) was at the center of repeated conflicts, resulting in its temporary block twice over the period of sampling. Due to a problem of interpretation of bot tasks, its activity resulted in an overwhelming number of edits strongly, repeatedly, and consistently contested by the community in terms of relevance and legitimacy. Hence, the discussion pages of this bot could not be considered as a standard model of users' reactions to bots, and thus this particular bot was thus excluded from this study. Therefore, our final corpus consisted of 50 bots.

2.3. Bots characteristics

Bots were defined by five biphasic characteristics. The first characteristic of a bot was the nature of its owner (either an administrator of Wikipedia ("admin"), either a registered user not having administrator rights). The second characteristic was the function of the bot, defined as being mostly constraining the activity of other users (protecting from vandalism, patrolling for copyright material, warning users), or as facilitating their work (substituting templates, correcting redirect or false links). The third characteristic was the location of the modifications made by the bot, which could be either frontstage (either in article and user-space, including talk pages), or backstage (like categories, portals, help, template spaces and so on). The fourth characteristic was whether the bot acted automatically (in an opt-out way) or was only subscription-based (opt-in).

Finally, a fifth characteristic was a global evaluation of the degree of autonomy of the bot, qualified as either high (bot presenting some degree of evaluation of user contributions before acting) or low (purely factual or managerial editions, such as correcting false links or distributing email around Wikipedia users).

When bots realized multiple types of edits, characteristics were defined by the type of edits most commented by users.

2.4. Posts and discussions analysis

A first analysis was performed on the total number of posts on the bots associated talk pages. When discussions were not directly on the bot's talk page, but on the talk page of the bot's owner, all of the posts that were not related to the bot were excluded. If posts were not in English or were mistakenly meant for another bot, they were also excluded. Answers of bot owners and talk page stalkers were coded separately.

A second analysis was performed using only the first post of each discussion, as divided in the talk pages. Discussions were characterized as a function of the perceived intention of the initial post. Nine distinct types of users reactions were identified, reflecting either positive intention (1, acknowledgments of the bot's work; 2, help and advice requests; and 3, user informing of an action), neutral intention (4, mistake or false-positive detection; 5, enhancement suggestions; and 6, question about the bot or the edits), and negative intention (7, general criticism; 8, specific edit contestation; and 9, user justification). Examples for each of these nine types are provided in [Table 1](#).

2.5. Constitution of bot ideotypes

While the analysis the five different characteristics of the bots brought key information, we noted that similar patterns of characteristics emerged on larger population of bots. Therefore, we built

two opposite stereotypes, which presented specific and identifiable behavioral patterns. On the one hand, some bots mainly did repetitive and laborious work instead of human users. These bots facilitated the work of users, acted backstage, automatically (opt-out), and with a low degree of autonomy. They could be qualified as "servant bots". On the other hand, some bots proactively enforced Wikipedia's guidelines and norms. They were controlled by administrators, constraining the work of users, acting front-stage, automatically and with a high degree of autonomy. They could be referred to as "policing bots". These two types of global behavioral patterns thus represented global ideotypes of bot function and functioning.

2.6. Coding validity

All the initial coding was performed by a single observer, blind to the hypotheses, over a period of 40 days. Coding validity was assessed by an intra-judge evaluation: a second coding was performed on 10% of the corpus, by the same judge, with an interval period between the two coding sessions of more than 6 months.

2.7. Statistical analysis

Data were coded using the RQDA open source software (version 0.2-3 with R version 2.15.2), which classify the data in a SQLite database, allowing further statistical compilation from our coding. Statistical analyses were made using SPSS (version 20, IBM, Armonk), SigmaPlot (version 11, Systat Software, Chicago) and Origin (version 8.1, OriginLab, Northampton). Due to the fact that the distribution of all the recorded variables followed power laws, all data were subjected to logarithmic transformations before statistical analysis. Hence, all statistical analyses were performed on the logarithmic transformed data and not on the raw data. Analyses were performed using the non-parametric Mann-Whitney *U* test ([Armitage & Berry, 1994](#)). Simple linear regressions were performed for the nine types of users reactions in order to validate them as "positive", "neutral", or "negative" according to the highest association with the valence of the first posts of each thread. For each of the five characteristics of bots and the two ideotypes, the repartitions of the nine types of users reactions were compared using non-parametric distribution free Kolmogorov-Smirnov analyses. When applicable, results are presented as mean \pm SEM.

3. Results

3.1. Characteristics of the bots population

A total of 872 bots were referenced on English Wikipedia pages on January 2, 2013. Between January 4, 2012 and January 2, 2013, 170 bots have been active, for a total of 8,985,353 edits ($52,855 \pm 10,437$ edits per bot, median: 10,037, maximum: 975,675, minimum: 20). The distribution of their contributions follows a power law ($R^2 = 0.759$, $p < .001$, [Fig. 1](#)). A similar distribution was evidenced in our final sample of 50 bots for the total number of edits ($R^2 = 0.837$, $p < .001$ with an average of $118,360 \pm 26,372$ edits, median: 58,399.5, maximum: 935,971, minimum: 1828, [Fig. 2A](#)), the total number of posts on talk pages ($R^2 = 0.915$, $p < .001$ with an average of 130.56 ± 19.8 posts, median: 89.5, maximum: 668, minimum: 20, [Fig. 2B](#)), the number of posts by users on talk pages excluding the bot's owner and bot's talk page stalkers ($R^2 = 0.921$, $p < .001$ with an average of 79.52 ± 12.31 posts, median: 51.5, maximum: 433, minimum: 13, [Fig. 2C](#)), and the number of discussions on talk pages during the sampling period ($R^2 = 0.916$, $p < .001$ with an average of 47.06 ± 8.91 discussions, median: 32,

Table 1

Examples of users' reaction for the nine types of observed reactions.

1. <i>Acknowledgments of the bot's work</i>	Good catch you silly robot C6541 (T ↔ C) 21:12, 27 November 2012 (UTC) – 28 Bot
2. <i>Help and advice requests</i>	Hi, I'd need the script you used to leave a notice to all the users which participated to the SOPA's discussion, Basically I have to do almost the same for the italian one. Thank you! – Vituzzu (talk) 12:22, 17 January 2012 (UTC) – Thehelpfulbot
3. <i>User informing of an action</i>	Hi, I think that that bug in your bot is actually Pywikipedia bug (you are using this bot, right?). So, I posted the bug here: https://sourceforge.net/tracker/?func=detail&aid=3532712&group_id=93107&atid=603138 Safinaskar (talk) 10:09, 7 June 2012 (UTC) – AvicBot
4. <i>Mistake or false-positive detection</i>	It appears that RFC bot (talk · contribs) is maintaining redlinks at Wikipedia:Dashboard/Requests for comment again. (See diff) Please check to see what is causing it and if you can correct the problem. – Allen4names 15:17, 25 May 2012 (UTC) – RFC Bot
5. <i>Enhancement suggestions</i>	Regarding this diff, it would be possible to reduce the number of similar false positives by searching matched content for 'Wikipedia'.—C M B J 05:52, 14 December 2012 (UTC) – MadmanBot
6. <i>Question about the bot or the edits</i>	Should this bot revert deletion of copyvio text at Lucy Morton including addition of a cospypaste template?—Preceding unsigned comment added by 82.132.232.222 (talk) 16:15, 30 July 2012 (UTC) – ClueBot NG
7. <i>General criticism</i>	Why are you spamming this nonsense to hundreds/thousands of people? Stop doing that. I didn't sign up for this crap, and I'm sure most/all others didn't either. Dream Focus 19:54, 4 September 2012 (UTC) – EdwardsBot
8. <i>Specific edit contestation</i>	Hello, there was a reason for the space. With it there Ivanhoe sorts before Ivanhoe East, which would seem to be the normal sorting convention; looks ugly the other way. Regards Crusoe8181 (talk) 07:57, 27 September 2012 (UTC) – BG19bot
9. <i>User justification</i>	I have check on this page article and in fact it is real see on this web site Official website I will add more source on this. Thank you Errorzerol (talk) 07:25, 1 September 2012 (UTC) – AnomieBot

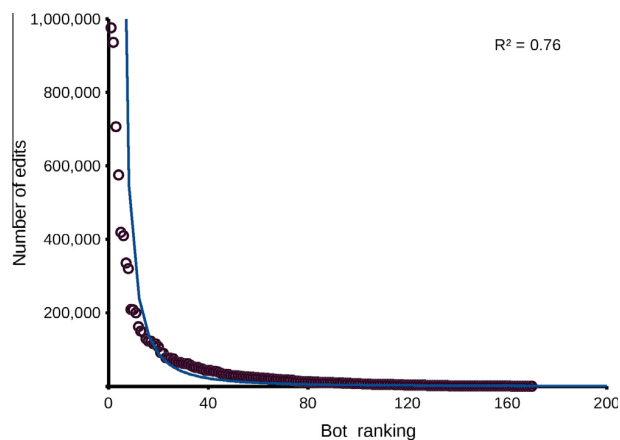


Fig. 1. Activity of the global bot population of the English Wikipedia. Distribution of the number of edits as a function of the bot ranking for the 170 bots active during the observation period (January 4, 2012 and January 2, 2013) on the English-language Wikipedia. This distribution follows a power law ($R^2 = 0.76$, $p < .001$).

maximum: 336, minimum: 6, Fig. 2D). At January 4, 2012, the 50 bots of our final sample were in average 1110.8 ± 92.4 days old (median: 1040, maximum: 2415, minimum: 6). At January 2, 2013, they had an average of 5.48 ± 0.95 active tasks (median: 4, maximum: 41, minimum: 1).

As mentioned earlier, some talk pages were shared between several bots. In such case, only the principal bot was counted. Therefore, 150 discussions were excluded from the 10 relevant talk pages (this phenomenon concerned 20% of the bots of our sample). Since half of the talk pages were shared between the bot and the owner of this bot (54% of our sample), and in some cases the talk page of the owner and the talk page of the bot were used concomitantly (14%), 3066 conversations were not related to the bot itself, and were thus excluded (87.6 ± 11.45 conversations per bot for the 34 bots specifically concerned by this phenomenon). Therefore, our final sampling consisted of 2353 conversations (47.06 ± 8.91 conversations per bot), for 6528 posts (130.56 ± 19.8 posts per bot and 3.14 ± 0.13 posts per conversation), including 3976 posts

from general Wikipedia users (79.52 ± 12.31 posts per bot, and 1.86 ± 0.07 posts by conversation), 1939 answers from the bot's owner (38.78 ± 5.66 owner posts per bot, and 1.05 ± 0.08 owner posts per conversation), and 613 answers from talk page stalkers (12.26 ± 4.78 talk page stalker's posts per bot, and 0.18 ± 0.04 talk page stalkers' posts per conversation).

The repartitions of the bots according to the five biphasic characteristics were as follow: 26 bots were controlled by an administrator (52%) and 24 were not (48%), 17 bots constrained the users' work (34%) and 33 bots facilitated it (66%), 29 bots acted frontstage (58%) and 21 acted backstage (42%), 35 bots acted automatically (70%) and 15 were subscription-based only (30%). Finally, 16 bots had high autonomy (32%) and 34 bots had low autonomy (68%).

3.2. Analysis of the total number of posts

On the 3976 users' posts observed, 1346 were positive (26.92 ± 5.59 positive posts per bots, accounting for $33.7\% \pm 2.2\%$ of the posts), 1704 were neutral (34.08 ± 5.21 neutral posts per bot, accounting for $42.5\% \pm 1.7\%$ of the posts), and 926 were negative (18.52 ± 2.77 negative posts per bot, accounting for $23.8\% \pm 1.5\%$ of the posts). The number of discussions is higher for constraining vs. facilitating (75.5 ± 22.89 vs. 32.8 ± 5.42 , $p < .001$), frontstage vs. backstage (62 ± 14.64 vs. 26.4 ± 3.42 , $p < .05$) and high autonomy vs. low autonomy (75.4 ± 24.39 vs. 33.7 ± 5.4 , $p < .05$) (Fig. 3A). The number of total posts from users was relatively stable across characteristics, with only significantly more posts observed from constraining vs. facilitating bots (180.88 ± 45.46 vs. 104.64 ± 17.76 posts per bot, $p < .05$), and frontstage vs. backstage bots (165.24 ± 31.82 vs. 82.67 ± 11.47 posts per bot, $p < .01$). This relation is still present when considering only posts from users, and so for constraining vs. facilitating bots (113.94 ± 28.92 vs. 61.79 ± 10.34 posts per bot, $p < .01$), frontstage vs. backstage bots (100.9 ± 19.89 vs. 50 ± 6.6 posts per bot, $p < .01$) and low vs. high autonomy (64.21 ± 10.34 vs. 112.06 ± 30.69 posts per bot, $p < .05$) (Fig. 3B).

Bots controlled by administrator elicited more neutral (46.35 ± 8.97 vs. 20.79 ± 3.32 posts per bot, for bots controlled and not controlled by an administrator, respectively, $p < .05$) and

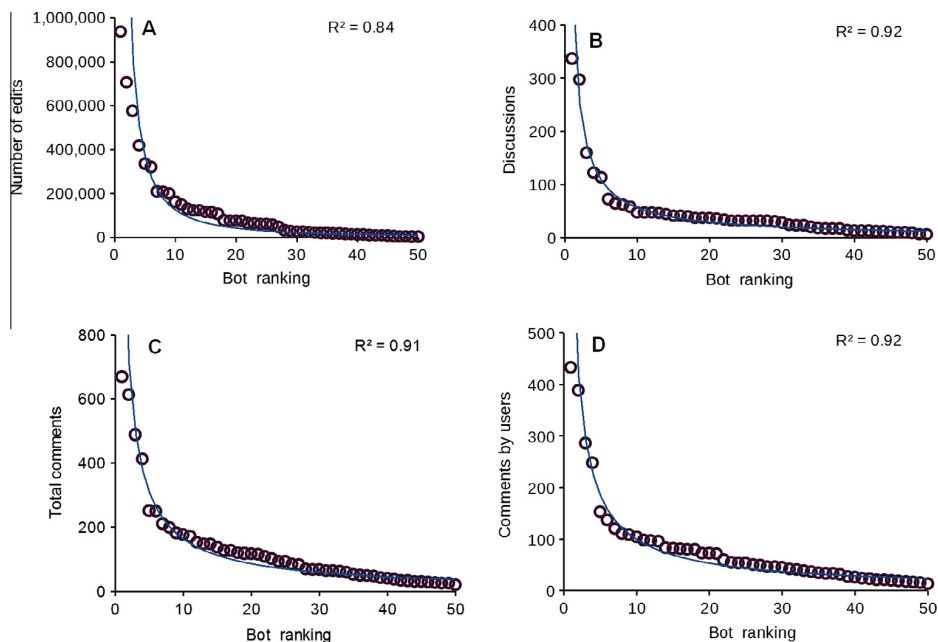


Fig. 2. Activity of the bot population studied. Shown are various measures of activity for the 50 bots of the studied sample. Note that all the observed distributions followed power laws. A. Distribution of the number of edits as a function of bot ranking ($R^2 = 0.84$, $p < .001$). B. Distribution of the number of discussion on bot-related talk pages as a function of bot ranking ($R^2 = 0.92$, $p < .001$). C. Distribution of the total number of interactions on bot-related talk pages as a function of bot ranking ($R^2 = 0.91$, $p < .001$). D. Distribution of the number of interactions on bot-related talk pages excluding the comments by owners and talk page stalkers, as a function of bot ranking ($R^2 = 0.92$, $p < .001$).

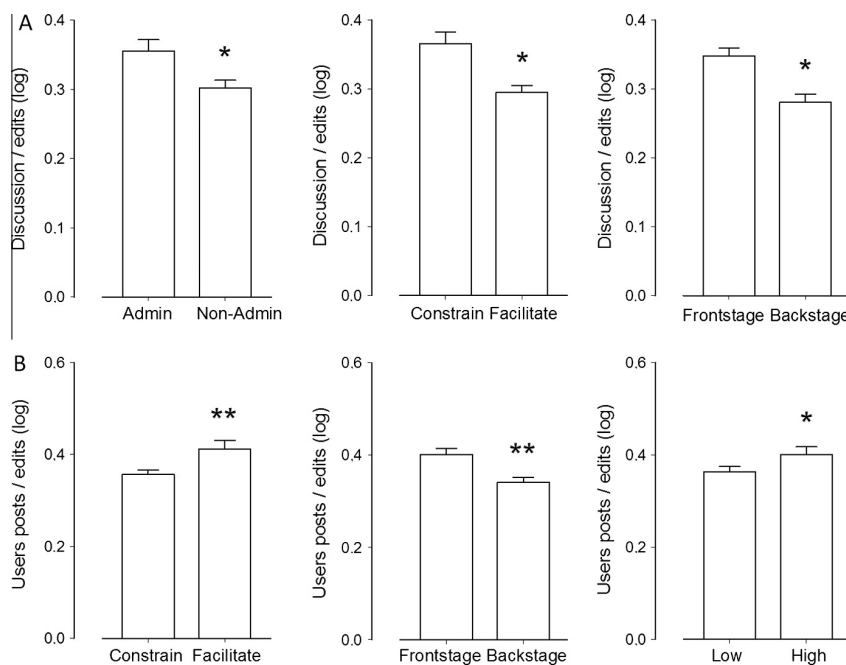


Fig. 3. Distributions of the interactions depending on bot characteristics. Quantification of the interactions of users on bot-related talk pages assessed by the ratio of the number of discussions/number of edits (A.) and the ratio of the number of user posts/number of edits (B.) across the five dichotomic characteristics of bots. A. Differences were observed in the ratio of the number of discussions/number of edits for three characteristics: bot's owner, bot's function, and the location of the edits of the bot. B. Differences were observed in the ratio of the number of user posts (excluding comments by owners and talk page stalkers)/number of edits for bot's function, the location of the edits of the bots, and the degree of autonomy of the bot. Statistical differences were assessed using a Mann–Whitney U test after a logarithmic transformation of the ratio. Differences were assessed using a Mann–Whitney U test. * indicates $p < .05$, and ** indicates $p < .01$.

negative posts (24.85 ± 4.71 vs. 11.67 ± 2.03 posts per bot, $p < .05$). More negative posts were observed for constraining bots than facilitating bots (28.88 ± 6.07 vs. 13.18 ± 2.39 posts per bot,

$p < .01$). Bots acting frontstage elicited more polarized responses than bots acting backstage, with both more positive (35.83 ± 9.2 vs. 14.62 ± 2.29 posts per bot, $p < .01$), and negative posts

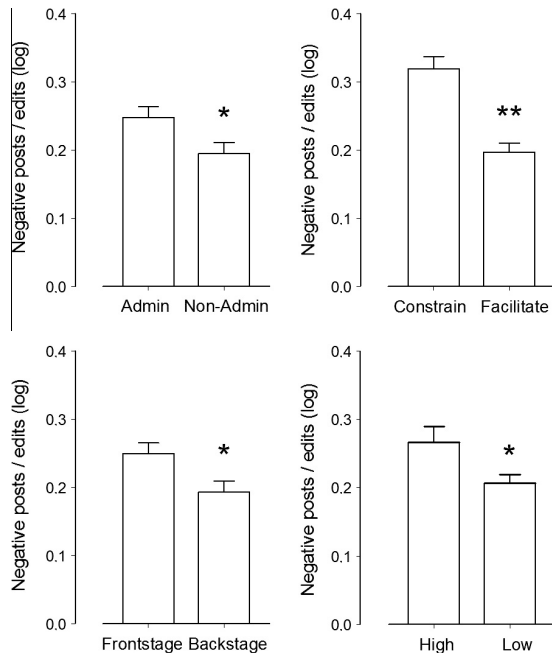


Fig. 4. Distribution of negative comments as a function of bots' characteristics. Significant differences were observed in the distribution of negative comments left by users on bot-related talk page for four characteristics of the bots: bot's owner (upper left), bot's function (upper right), the location of the edits (lower left), and the degree of autonomy of the bot (lower right). Statistical differences were assessed using a Mann–Whitney U test after a logarithmic transformation of the ratio. Differences were assessed using a Mann–Whitney U test. * indicates $p < .05$, and ** indicates $p < .01$.

(23.48 ± 4.4 vs. 11.67 ± 1.83 posts per bot, $p < .05$). Finally, bots with high autonomy elicited more negative posts compared to bots with low autonomy (28.81 ± 6.52 vs. 13.68 ± 2.33 posts per bot, $p < .05$) (Fig. 4). No significant difference was assessed between opt-in and opt-out bots.

Talk page stalkers were significantly more active for administrator than non-administrator (18.92 ± 8.91 vs. 5.04 ± 1.85 posts per bot, $p < .05$), frontstage bots rather than backstage bots (18.48 ± 8.03 vs. 3.67 ± 1.48 posts per bot, $p < .05$).

3.3. Analysis of users' types of reactions in discussions

A total of 2353 discussions was observed in our sample, including 723 positive discussions (15.26 ± 4.52 discussions per bots, $29.86\% \pm 2.69\%$), 932 neutral discussions (18.64 ± 3.06 discussions per bots, $40.78\% \pm 1.93\%$), and 657 negative discussions (13.14 ± 2.19 discussions per bots, $29.4\% \pm 2.15\%$). The average discussion contained 1.86 ± 0.47 posts, including 1.05 ± 0.54 owner posts (typically answers) and 0.18 ± 0.27 talk page stalkers interventions. The validity of the coding for the tone of the discussions was extremely high (intra-judge reliability: 95.8%).

Among the sample, 430 discussions reflected acknowledgments or praising (8.6 ± 4.07 discussions per bots, $9.7\% \pm 1.6\%$), 245 discussions were help requests (4.9 ± 0.79 discussions per bots, $11.8\% \pm 1.5\%$), 170 discussions were initiated by users informing of an action they undertook (3.4 ± 0.69 discussions per bot, $7.3\% \pm 0.78\%$), 385 discussions signalling an error (7.7 ± 1.18 discussions per bot, $21.1\% \pm 2.2\%$), 244 discussions making suggestions (4.88 ± 0.82 discussions per bot, $11.5\% \pm 1\%$), 357 questioning the logics of edits (7.14 ± 1.01 discussions per bot, $17.3\% \pm 1.2\%$), 229 were general criticism (4.58 ± 9.3 discussions per bot, $9\% \pm 1.1\%$), 198 were contestations of a specific edit (3.96 ± 0.7 discussions

per bot, $9.5\% \pm 1.6\%$), and 95 user justifications (1.9 ± 0.65 discussions per bot, $2.7\% \pm 0.9\%$). Once again, the validity of the coding was extremely high (intra-judge reliability: 91.2%).

The repartitions of the types of users reactions in discussions as a function of the characteristic of the bots unveiled interesting, yet complex effects. While no overall difference in the repartition of the nine types of users reactions was observed between bots controlled by an administrator and by bots controlled by another user (Fig. 5), bots controlled by an administrator elicited significantly more responses for some types of users reactions, both positive (acknowledgment $p < .01$, user informing $p < .05$) and negative (specific contestation $p < .05$, user justification $p < .05$). In contrast, overall significant differences in the repartition of the nine types of users reactions were assessed both regarding the function of the bots ($p < .05$), and the localization of their actions ($p < .05$). Constraining bots elicited significantly more feedback than facilitating bots for one positive type of users reaction (user informing, $p < .01$), one neutral type of users reaction (questions $p < .05$), and all of the three negative types of users reactions (general criticism $p < .01$, specific contestation $p < .001$, justifications $p < .001$). Note that the frontstage modifications elicited more help requests ($p < .01$), but also questionings ($p < .05$) and general criticism ($p < .05$) than backstage modifications. Repartitions of the types of users reactions were mostly not affected by whether the bots acted automatically or were subscription-based, with users informing more ($p < .05$) and providing more justification ($p < .05$) for automatically applying bots (Fig. 5).

The pattern of reactions according to the degree of autonomy of the bot was rather similar to what was observed regarding the function the bots (Fig. 5). Overall, bots with higher degree of autonomy elicited more reactions than bots with lower autonomy ($p < .05$), with both more positive (users informing, $p < .05$) and negative reactions (general criticism $p < .05$, specific contestation $p < .05$, justifications $p < .01$).

3.4. Bot ideotypes

In our sample, 13 bots belonged to the “servant bot” ideotype, and 9 to the “policing bot” ideotype. Linear regressions revealed an inverse relationship for “servant bots” and “policing bots” between the number of criterions met toward the ideotype and the number of discussions observed for the bot ($R^2 = 0.28$, $p < .05$ and $R^2 = 0.25$, $p < .01$ for “servant bots” and “policing bots” respectively, Fig. 6).

“Servant bots” elicited less total interventions ($p < .01$), either for user comments ($p < .01$), for the total number of discussions ($p < .05$), for the owner responses ($p < .05$), and for talk page stalkers interventions ($p < .01$). There were less positive ($p < .01$) and negative comments ($p < .01$) than for bots not belonging to this ideotype (Fig. 7). In contrast, no significant difference was assessed on global answers for “policing bots”.

Overall significant differences were found in the repartition of the nine types of users reactions between bots belonging to the ideotype and other bots, both for “servant bots” ($p < .05$), and for “policing bots” ($p < .01$). When considering the responses to the bots' actions accordingly to the nine identified types of users' reaction, “servant bots” elicited less requests for the use of the bot ($p < .05$), less edit questioning ($p < .05$) and less information on false positives or errors ($p < .05$) than bots not belonging to this ideotype (Fig. 8). “Policing bots” elicited more polarized responses, with either positive reactions (acknowledgments $p < .05$, and informing of an action $p < .05$) or negative reactions (general criticisms $p < .05$, and justification $p < .01$) significantly different than for bots not belonging to the ideotype (Fig. 8).

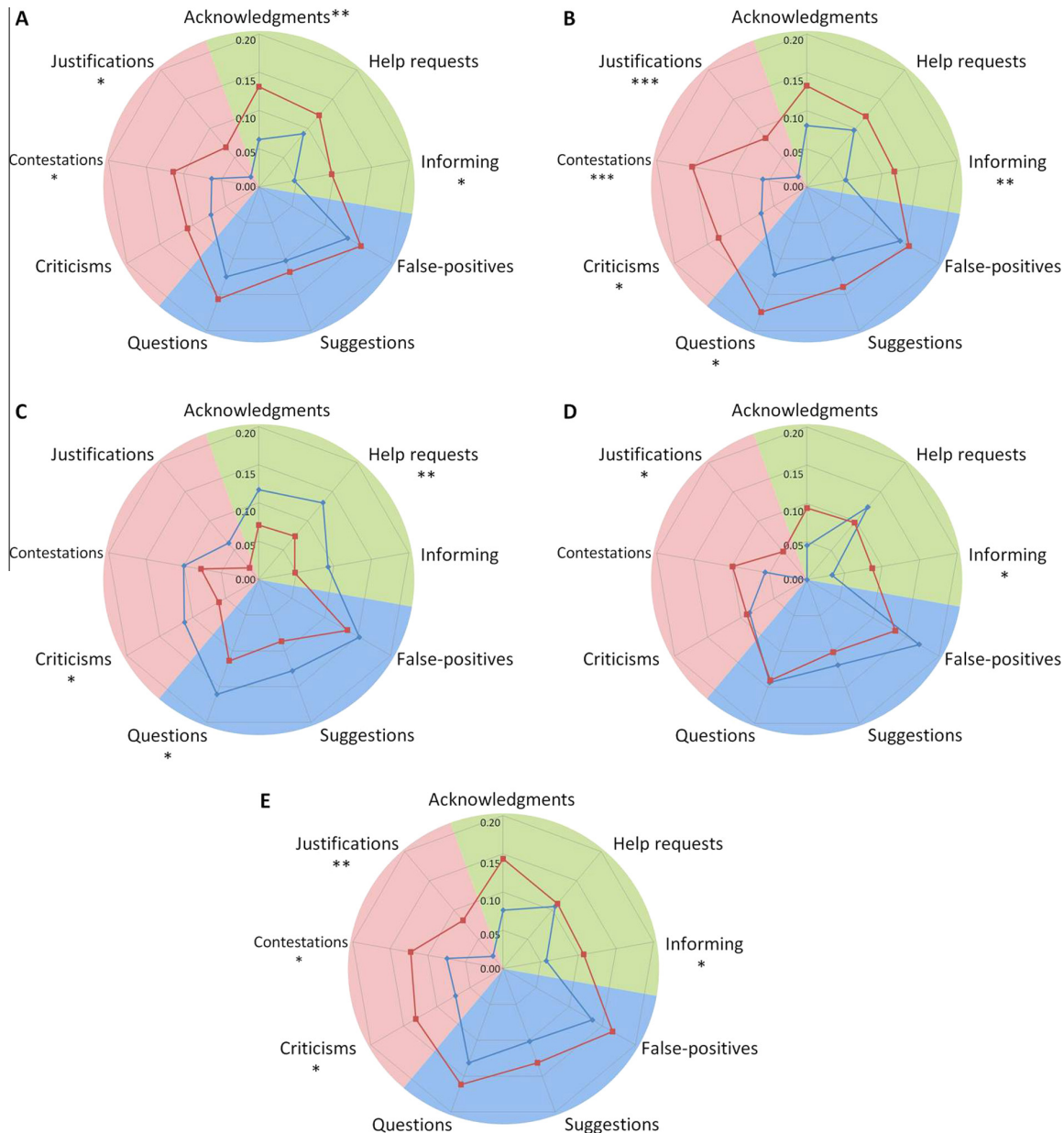


Fig. 5. Users reactions depending on the characteristics of the bots. Polar plots representing the proportion of the 9 identified types of user reactions to bots (3 types of positive reactions: Acknowledgments, Help requests, informing, light green; 3 types of neutral reactions: False-positives, Suggestions, Questions, light blue; and 3 types of negative reactions: Criticisms, Contestations, Justifications, light red), depending on the different characteristics of the bots. For each type of user reaction, the radial axis displays the logarithmic transformation of the ratio of the number of comments of this type of reaction/total number of edits of the bot. For each characteristic, overall differences in the repartition of types of reactions were assessed using a Kolmogorov–Smirnov test. * indicates a $p < .05$ difference and ** a $p < .01$ difference for the particular type of reaction for the characteristic (Mann–Whitney U test). A. User reactions depending on the nature of the bot's owner (administrators, big dots and red line, vs. non-administrators, small dots and blue line, Kolmogorov–Smirnov, $p < .05$). B. User reactions depending on the bot's function (constraining bots, big dots and red line, vs. facilitating bots, small dots and blue line, Kolmogorov–Smirnov, $p < .05$). C. User reactions depending on the location of the edits (backstage, big dots and red line, vs. frontstage, small dots and blue line, Kolmogorov–Smirnov, $p < .05$). D. User reactions depending on the degree of automatism of the bot's actions (opt-out bots, big dots and red line, vs. opt-in bots, small dots and blue line, Kolmogorov–Smirnov, $p < .05$). E. User reactions depending on the degree of autonomy of the bot (high degree of autonomy, big dots and red line, vs. Low degree of autonomy, small dots and blue line, Kolmogorov–Smirnov, $p < .05$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

4.1. Wikipedia bots as a model of perception of human/bot interactions

While Wikipedia bots have already been the target of research interest (Geiger, 2011; Halfaker & Riedl, 2012; Müller-Birn et al., 2013; Niederer & van Dijck, 2010), the focus so far has almost

always been on characterizing the putative role of the bots on structural levels – what is referred to by some authors as sociotechnical norms or system (Halfaker & Riedl, 2012; Niederer & van Dijck, 2010). The objective of the present study was fully different: we focused here on investigating how Wikipedia users reacted to the action that bots had over their work, in order to decipher the perception of the nature of the interactions between users and bots.

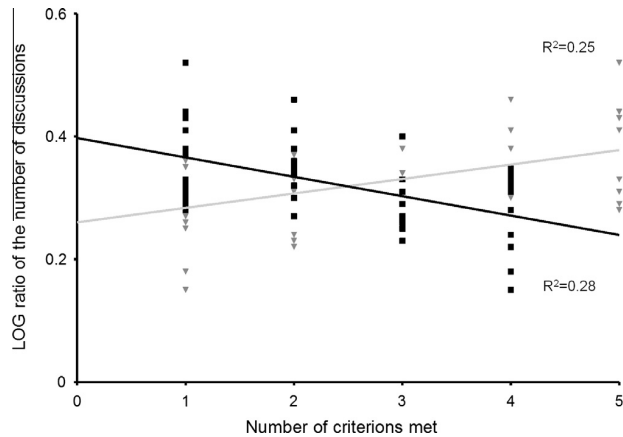


Fig. 6. Number of discussions as a function of the number of criterions met for each of the two identified ideotypes of bots. Linear regressions between the number of discussions (expressed as the logarithmic transformation of the ratio of the number of discussions/number of edits) and the number of criterions met for “servant” bots (black line, $R^2 = 0.28$, $p < .05$), and for “policing” bots (gray line, $R^2 = 0.25$, $p < .01$).

Our results show that not only all bots are not equal, but they also are not equally perceived. We also uncovered that bots broadly tend toward one of two stereotypes depending on the actual behavior they display, ranging from a “servant” type doing mostly repetitive minor corrections, to a more intrusive “policing” type. Bots appear to be surprisingly well accepted – even when it comes to the most intrusive ones. Furthermore, bots are largely perceived as valuable for the viability of Wikipedia, as they were consistently endorsed and defended by members of the internal community. This acceptance, which arises despite potentially

ambiguous issues regarding the balance of power, might originate in the fact that the bots are still perceived as an extension of human users rather than truly independent agents.

A legitimate question would be to know whether these human–bots interactions are representative of the kind of interactions which can occur in virtual spaces, both in terms of magnitude and of quality. Investigations of Wikipedia-based large datasets have consistently demonstrated that the distribution of user-generated events followed power laws (Muchnik et al., 2013; Nazir & Takeda, 2008; Voss, 2005; Yasseri, Sumi, Rung, Kornai, & Kertész, 2012). For instance, that is the case for the number of communications between users (Muchnik et al., 2013), the number of user edits for an article (Muchnik et al., 2013), the number of edited articles by a registered user (Nazir & Takeda, 2008), the number of authors per article (Voss, 2005), the number of ingoing and outgoing links per article (Voss, 2005), the number of broken links per article (Yasseri et al., 2012). Thus, the fact that this phenomenon was also observed for our data strongly suggests that our sample reflects the actual occurrence of spontaneously generated behavior in social networks.

Despite attributing human-like qualities to bots and displaying a willingness to collaborate with them, the general perception of users about bots remains as automated servants rather than partners equal in terms of social status (Aharoni & Fridlund, 2007). Indeed, both in experimental and ecological settings, human subjects confronted to discussion bots consciously display a higher quality of inter-personal behavior when the interlocutor is identified as human vs. as a bot (Aharoni & Fridlund, 2007; Lortie & Guitton, 2011). In the present case, several elements clearly reminded users that bots were controlled by humans rather than independent agents. Indeed, more than 50% of the bots of our sample were directly controlled by Wikipedia administrators. The fact

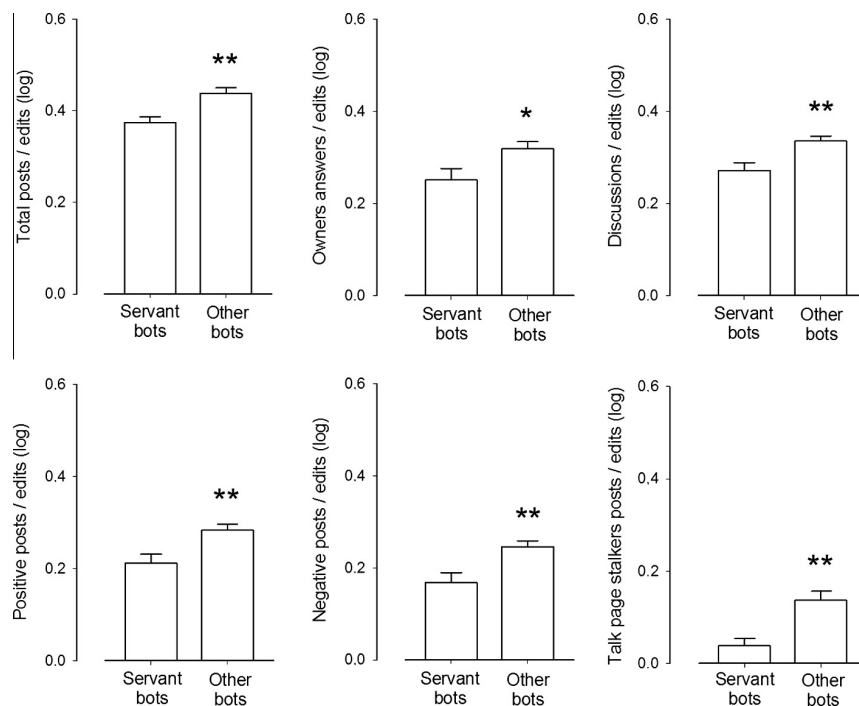


Fig. 7. Differences in the intensity of interactions between users and “servant” vs. other bots. A. Differences in the ratio of the total number of posts/number of edits of the bot. B. Differences in the ratio of the number of owner responses/number of edits of the bot. C. Differences in the ratio of the number of discussions/number of edits of the bot. D. Differences in the ratio of the number of positive posts/number of edits of the bot. E. Differences in the ratio of the number of negative posts/number of edits of the bot. F. Differences in the ratio of the number of talk page stalkers posts/number of edits of the bot. Statistical differences were assessed using a Mann–Whitney U test after a logarithmic transformation of the ratio. * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

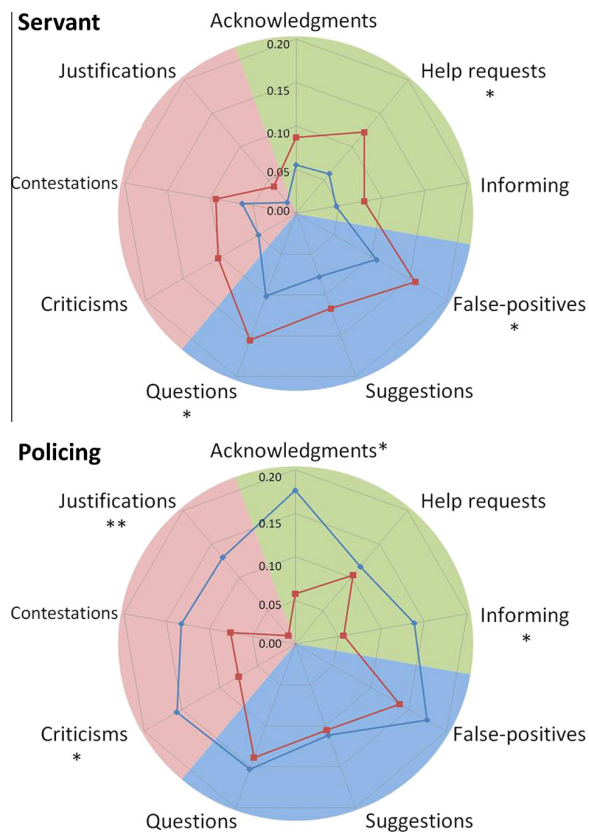


Fig. 8. Overall distribution of the users reactions depending on the ideotype of the bot. Polar plots representing the proportion of the 9 identified types of users reactions to bots (3 types of positive reactions: Acknowledgments, Help requests, Informing, light green; 3 types of neutral reactions: False-positives, Suggestions, Questions, light blue; and 3 types of negative reactions: Criticisms, Contestations, Justifications, light red), depending on the ideotype of the bot (servant bots, upper panel; and policing bots, lower panel). For each type of user reaction, the radial axis displays the logarithmic transformation of the ratio of the number of comments for this type of reaction/total number of edits of the bot. For each ideotype, overall differences in the repartition of types of reactions were assessed using a Kolmogorov–Smirnov test. * indicates a $p < .05$ difference and ** a $p < .01$ difference for the particular type of reaction for the characteristic (Mann–Whitney U test). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that most bots discussions were at least partially redirected under the bots owner's talk page probably reinforced this perception.

4.2. Human users' reaction to bots actions

While human control was apparent in most of the considered bots, the possibility of massive actions of bots over the content of Wikipedia challenges this simplistic view by pushing the interactions to another level of magnitude. In such a context, would users' views of bots be limited to seeing them as purely automated servants, or would more complex perceptions occur?

Generally speaking, a very important variability was observed when considering the reactions of users to Wikipedia bots, with some bots not gathering more than 12 comments in a year. However, when considered in the light of the distributions of Wikipedia-based behavior, such variability is not surprising. The characterization of bots according to the degree of visibility of the edits they perform predicted at least partially the future interactions with the members of the community. Indeed, the bots acting frontstage concentrated most of the reactions compared to the bots acting backstage. The most negative comments from users

were observed when the actions of a bot resulted in numerous very minor edits on a potentially large number of pages. Indeed, while a regular user doing such purely cosmetic edits would not be an hassle, bots doing so, with their power of treatment, can easily become overwhelming and saturate other users following a particular page with unnecessary and irrelevant information. While excluded from the present study, it is exactly what happened in the case of Yobot.

While considering bots on a continuum depending on their actualized behavior – a controlling bot vs. a helper – the users' perception evolves as a function of the role of the bots. In terms of perception, the bots closely matching the servant ideotype reflected the general stereotype of bots as automated agents. By performing repetitive and labor-intensive tasks, these bots decrease the burden born by users (Geiger, 2011; Halfaker & Riedl, 2012) and are clearly perceived as automated helpers, akin to some forms of software enhancement. Hence, these bots are the target of extremely few interactions with the community of users.

However, the situation was much more complex regarding the more invasive types of bots – the policing bots. Very often, the bots in this category systematically applied (or “force” users to apply) sets of norms which were initially supposed to be followed on a purely voluntary basis (the “ignore all rules” of the five pillars of Wikipedia). Therefore, the actions of these bots are more invasive than those of purely servant bots, and could be perceived by users as “aggressive” or, at the very least, constraining. One could have predicted that being “controlled” by a bot in terms of editorial policy would elicit massively negative reactions from human users. Indeed, as intuitively foreseen, these bots concentrated most of the interactions with users, and most of the feedbacks from the community. However, in contrast to a simplistic view, these feedbacks were not simply negative, but rather equally negative and positive.

These interactions should be read according to two opposite, yet complementary hypotheses. A first possibility would be for these bots to be perceived as external intruders by Wikipedia's contributors, thus eliciting massive criticism and rejection. Alternatively, the actions of the bots on Wikipedia pages could be perceived by users as reflecting the application of current norms, and thus, the debates over the bots' personal pages would revolve mainly around the legitimacy of the norms rather than on the actions of the bots themselves. Our data suggest that this second possibility is the correct one, since while policing bots elicited both positive and negative reactions, these reactions strongly revolve around acknowledgments or general criticisms. This might however bear a bias, since bots' owners will often use the application of the norm as an argument when facing criticisms (e.g., owners would claim that they would gladly modify the behavior of the bots once a different consensus emerges). Thus, the constant and massive implication of bot owners, administrators, and talk page stalkers that we evidenced may refrain other users from perceiving bots as too independent by constantly remaining them of their status of editorial helpers.

4.3. Bots and community: perception of the balance of power

The automated workforce of bots has a massive impact on the content of Wikipedia, and leads to putative questionings regarding the balance of power between bots and human users. In this context, a legitimate question is to inquire how this apparent conflict between the users' intuitive perception of bots and the reality of their power gets actualized in Wikipedia.

In Wikipedia, not all users have the same administrative powers or editorial rights. Indeed, a hierarchy can be established between the different types of users according to the actual powers they have access too, the system developers being at the top of the

pyramid while banned users are at the very bottom (Niederer & van Dijck, 2010). In this structure, the bots are just below the administrators in terms of power, but are above regular registered users, i.e., the vast majority of the active community of Wikipedia (Niederer & van Dijck, 2010). Interestingly, when a bot is controlled by a regular user instead of an administrator, the bot itself has more editorial power than its owner. It was thus not a surprise to see in the present study that administrators were highly represented among bot owners (with more than 50% of the bots of our sample directly controlled by an administrator).

With bots having major editorial power, user perceptions may be sensitive to the presence of limits to bot actions. One of the first debates on the acceptance of bots resulted in the implementation of a “no-bot” template as early as 2006, akin to a form of opt-out mechanism, allowing users to individually forbid bots to act on specific pages (Geiger, 2011; Halfaker & Riedl, 2012). Our results show that these mechanisms had the expected results. Indeed, the absence of differences between opt-in and opt-out bots in our study suggests that this parameter is not anymore a relevant criterion when interpreting the interactions between bots and users. Nonetheless, our results show that this concern about bot power is present both in the positive and the negative reactions of users, with users both informing about their actions (positive) and justifying their actions (negative) when interacting with opt-out bots. Obviously, these types of feedback decrease when interacting with opt-in bots, since they acts specifically over users demands. Finally, the fact that bots cannot revert the edition of a human having reverted the bots correction maintains the final judgement into human’s hands, reducing not only the actual powers of bots, but also the potential for criticisms and negative feedbacks, as observed in our sample.

While being *a priori* minor, the numerous small, repetitive, and mostly unnoticed corrections made by the bots (particularly of the servant type) have a very high structural impact, as they concern standardizing templates or categories, on which the whole system of classification of the encyclopedia depends. Our results demonstrate that this structural role of bots is largely acknowledged by the community, as evidenced by the number of administrators and talk page stalkers coming to the rescue of bots in discussions. In this view, bots could be perceived by internal members as answering a worry of the Wikipedia community facing a massive participation of users not fully aware of the encyclopedia’s standards (Butler, Joyce, & Pike, 2008).

In conclusion, our results confirm the notion that bots are currently perceived mostly as indentured collaborators and potentially valuable helpers by members of the Wikipedia community. While bots take on their own important part of the labor-intensive editorial work, they are not perceived by users as running on their own, and human control can always been felt so far. Similarly, retroactive evaluation of bot actions is the target of intense scrutiny by the community. The reactions of the community members translate a wish of having more efficient bots, more than general contestation or feelings of unease. Finally, the coherence of the reactions to bot actions could be the expression of a maturing process of the community including a reinforcement of shared sets of agreed rules.

Acknowledgments

MJG holds a Career Grant from the “Fonds de Recherche en Santé – Québec – Santé” (FRQS). This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC – Grant Number 371644).

References

- Aharoni, E., & Fridlund, A. J. (2007). Social reactions toward people vs. computers: How mere labels shape interactions. *Computers in Human Behavior*, 23, 2175–2189. <http://dx.doi.org/10.1016/j.chb.2006.02.019>.
- Armitage, P., & Berry, G. (1994). *Statistical methods in medical research*. Oxford: Blackwell.
- Black, L. W., Welsler, H. T., Cosley, D., & Degroot, J. M. (2011). Self-governance through group discussion in Wikipedia: Measuring deliberation in online groups. *Small Group Research*, 42, 595–634. <http://dx.doi.org/10.1177/1046496411406137>.
- Butler, B., Joyce, E., & Pike, J. (2008). *Don't look now, but we've created a bureaucracy: The nature and roles of policies and rules in Wikipedia*. CHI 2008 proceedings. Florence, Italy: ACM Press (pp. 1101–1110). Florence, Italy: ACM Press.
- Chang, H. H., & Chuang, S. S. (2011). Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator. *Information & Management*, 48, 9–18. <http://dx.doi.org/10.1016/j.im.2010.11.001>.
- Dunbar, R. I. (2012). Social cognition on the Internet: Testing constraints on social network size. *Philosophical Transactions of the Royal Society of London: B Biological Sciences*, 367, 2192–2201. <http://dx.doi.org/10.1098/rstb.2012.0121>.
- Edwards, C., Edwards, A., Spence, P. R., & Shelton, A. K. (2014). Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. *Computers in Human Behavior*, 33, 372–376. <http://dx.doi.org/10.1016/j.chb.2013.08.013>.
- Fallis, D. (2008). Toward an epistemology of Wikipedia. *Journal of the American Society for Information Science and Technology*, 59, 1662–1674. <http://dx.doi.org/10.1002/asi.20870>.
- Geiger, R. S. (2009). The social roles of bots and assisted editing programs. In *Proceedings of the 5th international symposium on wikis and open collaboration – WikiSym '09* (pp. 1–2). Orlando, Florida, USA: ACM Press. <http://dx.doi.org/10.1145/1641309.1641351>.
- Geiger, R. S. (2014). Bots, bespoke code, and the materiality of software platforms. *Information, Communication & Society*, 17, 342–356. <http://dx.doi.org/10.1080/1369118X.2013.873069>.
- Geiger, R. S. (2011). The lives of bots. In G. Lovink & N. Tkacz (Eds.), *Critical point of view: A Wikipedia Reader* (pp. 78–93). Amsterdam, The Netherlands: Institute of Network Cultures.
- Geiger, R. S., & Ribes, D. (2010). *The work of sustaining order in Wikipedia: The banning of a vandal*. Savannah, Georgia, USA: ACM Press (pp. 117–126). Savannah, Georgia, USA: ACM Press.
- Giles, J. (2005). Internet encyclopaedias go head to head. *Nature*, 438, 900–901. <http://dx.doi.org/10.1038/438900a>.
- Guittou, M. J. (2015). Swimming with mermaids: Communication and social density in the second life merfolk community. *Computers in Human Behavior*, 48, 226–235. <http://dx.doi.org/10.1016/j.chb.2015.02.004>.
- Halfaker, A., & Riedl, J. (2012). Bots and cyborgs: Wikipedia’s immune system. *Social Computing*, 12, 79–82.
- Jahnke, I. (2010). Dynamics of social roles in a knowledge management community. *Computers in Human Behavior*, 26, 533–546. <http://dx.doi.org/10.1016/j.chb.2009.08.010>.
- Lortie, C. L., & Guittou, M. J. (2011). Judgment of the humanness of an interlocutor is in the eye of the beholder. *PLoS One*, 6, 1–7. <http://dx.doi.org/10.1371/journal.pone.0025085>.
- Loveland, J., & Reagle, J. (2013). Wikipedia and encyclopedic production. *New Media & Society*, 15, 1–18. <http://dx.doi.org/10.1177/1461444812470428>.
- Mowbray, M. (2014). Automated Twitter accounts. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Eds.), *Twitter and society* (pp. 183–194). New York: Peter Lang.
- Muchnik, L., Pei, S., Parra, L. C., Reis, S. D. S., Andrade, J. S., et al. (2013). Origins of power-law degree distribution in the heterogeneity of human activity in social networks. *Scientific Reports*, 3, 1–7. <http://dx.doi.org/10.1038/srep01783>.
- Müller-Birn, C., Herbsleb, J. D., & Dobusch, L. (2013). *Work-to-rule: The emergence of algorithmic governance in Wikipedia*. C&T'13. Munich, Germany: ACM Press (pp. 1–10). Munich, Germany: ACM Press.
- Nazir, F., & Takeda, H. (2008). Extraction and analysis of tripartite relationships from Wikipedia. *Technology and Society*, 1–13.
- Niederer, S., & van Dijck, J. (2010). Wisdom of the crowd or technicity of content? Wikipedia as a sociotechnical system. *New Media & Society*, 12, 1368–1387. <http://dx.doi.org/10.1177/1461444810365297>.
- Viégas, F. B., Wattenberg, M., & McKeon, M. M. (2007). The hidden order of Wikipedia. *Online Communities and Social Computing*, 445–454.
- Voss, J. (2005). Measuring Wikipedia. In *Proceedings 10th international conference of the International Society for Scientometrics and Informetrics* (pp. 221–231). Stockholm, Sweden.
- Yang, W., Liu, W., Viña, A., Tuanmu, M. N., He, G., et al. (2013). Nonlinear effects of group size on collective action and resource outcomes. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 10916–10921. <http://dx.doi.org/10.1073/pnas.1301733110>.
- Yasseri, T., Sumi, R., Rung, A., Kornai, A., & Kertész, J. (2012). Dynamics of conflicts in Wikipedia. *PLoS One*, 7, 1–12. <http://dx.doi.org/10.1371/journal.pone.0038869>.
- Zlatic, V., Bozicevic, M., Stefancic, H., & Domazet, M. (2006). Wikipedias: Collaborative web-based encyclopedias as complex networks. *Physical Review E*, 74, 016115.

Bibliographie

- Aaltonen, A. et J. Kallinikos. 2013, « Coordination and Learning in Wikipedia : Revisiting the Dynamics of Exploitation AND Exploration », dans *Managing 'Human Ressources' by Exploiting and Exploring People's Potentials*, vol. 37, research in the sociology of organizations éd., Emerald Group Publishing Limited, p. 161–192, doi :10.1108/S0733-558X(2013)0000037010. URL [http://dx.doi.org/10.1108/S0733-558X\(2013\)0000037010](http://dx.doi.org/10.1108/S0733-558X(2013)0000037010).
- Adamic, L. A. et B. A. Huberman. Mars 2000, « Power-Law Distribution of the World Wide Web », *Science*, vol. 287, n° 5461, doi :10.1126/science.287.5461.2115a, p. 2115–2116, ISSN 00368075. URL <http://www.sciencemag.org/content/287/5461/2115.full>.
- Adler, B. T., L. de Alfaro, I. Pye et V. Raman. 2008, « Measuring author contributions to the Wikipedia », dans *WikiSym '08*, ACM Press, Porto, Portugal, ISBN 9781605581286, p. 1, doi :10.1145/1822258.1822279. URL <http://portal.acm.org/citation.cfm?doid=1822258.1822279>.
- Aharoni, E. et A. J. Fridlund. 2007, « Social reactions toward people vs. computers : How mere lables shape interactions », *Computers in Human Behavior*, vol. 23, n° 5, doi :10.1016/j.chb.2006.02.019, p. 2175–2189, ISSN 07475632. URL <http://linkinghub.elsevier.com/retrieve/pii/S0747563206000409>.
- Aiello, L. M., M. Deplano, R. Schifanella et G. Ruffo. 2012, « People Are Strange When You're a Stranger : Impact and Influence of Bots on Social Networks », dans *Links*, vol. 697, p. 1–10. URL <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4523/4961>.
- Albert, R., H. Jeong et A.-L. Barabasi. 1999, « Diameter of the World-Wide Web », *Nature*, vol. 401, n° 9, p. 130.
- Alluvatti, G. M., A. Capiluppi, G. D. Ruvo et M. Molfetta. 2011, « User Generated (Web) Content : Trash or Treasure », dans *IWPSE_EVOL'11*, ACM Press, Szeged, Hungary, ISBN 9781450308489, p. 81–90.
- Almeida, R. B., B. Mozafari et J. Cho. 2007, « On the Evolution of Wikipedia », dans *ICWSM'2007*, Boulder, Colorado, p. 1–8.

- Anderson, B. 2006, *Imagined communities : Reflections on the origin and spread of nationalism.*, Verso Books, New York, New York, ISBN 9781844674848, 247 p..
- Antin, J. et C. Cheshire. 2010, « Readers are Not Free-Riders : readers are Not Free-Riders : Reading as a Form of Participation on Wikipedia », dans *CSCW 2010*, ACM Press, Savannah, Georgia, USA, ISBN 9781605587950, p. 127–130.
- Apic, G., M. J. Betts et R. B. Russell. 2011, « Content disputes in Wikipedia reflect geopolitical instability. », *PLoS ONE*, vol. 6, n° 6, doi :10.1371/journal.pone.0020902, p. 1–3, ISSN 1932-6203. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3120813&tool=pmcentrez&rendertype=abstract>.
- Arazy, O., O. Nov et F. Ortega. 2014, « The [Wikipedia] World is not Flat : on the Organizational Structure of Online Production Communities », dans *Twenty Second European Conference on Information Systems*, Tel Aviv, Israel, p. 1–15.
- Arazy, O., L. Yeo et O. Nov. 2013, « Stay on the Wikipedia Task : When Task-Related Disagreements Slip Into Personal and », *Journal of the American Society for Information Science and Technology*, vol. 64, n° 6, doi :10.1002/asi, p. 1634–1648.
- Armitage, P. et G. Berry. 1994, *Statistical Methods in Medical Research*, Blackwell, Oxford, United Kingdom, ISBN 9780470775349, 832 p..
- Bar, A. R. et M. Maheswaran. 2014, *Confidentiality and Integrity in Crowdsourcing Systems*, springerbr éd., Springer, New York, ISBN 9783319027166, 79 p..
- Bock, G.-w., R. W. Zmud, Y.-g. Kim et J.-n. Lee. 2005, « Behavioral Intention Formation in Knowledge Sharing : Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate », *MIS Quarterly*, vol. 29, n° 1, p. 87–111.
- Breazeal, C. L. 2002, *Designing Sociable Robots*, MIT Press éd., A Bradford Book, Cambridge, ISBN 0262025108, 282 p..
- Brossard, D. 2013, « New media landscapes and the science information consumer. », *PNAS*, vol. 110, n° 3, doi :10.1073/pnas.1212744110, p. 14096–101, ISSN 1091-6490. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3752175&tool=pmcentrez&rendertype=abstract>.
- Butler, B., E. Joyce et J. Pike. 2008, « Don't Look Now, But We've Created a Bureaucracy : The Nature and Roles of Policies and Rules in Wikipedia », dans *CHI 2008 Proceedings*, ACM Press, Florence, Italy, ISBN 9781605580111, p. 1101–1110.
- Canavan, J. 2005, « The Evolution of Malicious IRC Bots », dans *VB2005 Conference*, Symantec, Cupertino, CA, USA, p. 1–37.

- Carr, N. 2008, « A Spider's Web », dans *The Big Switch : rewiring the world, from Edison to Google*, chap. 10, WW Norton & Compagny, p. 185–209.
- Cassiman, B., M. Cohen, J. King, W. Powell, M. Prato et M. Ven. 2007, « The emergence of governance in an open source community », *Academy of Management Journal*, vol. 50, n° 5, p. 1079–1106.
- Chang, H. H. et S.-S. Chuang. 2011, « Social capital and individual motivations on knowledge sharing : Participant involvement as a moderator », *Information & Management*, vol. 48, n° 1, doi :10.1016/j.im.2010.11.001, p. 9–18, ISSN 03787206. URL <http://linkinghub.elsevier.com/retrieve/pii/S0378720610000844>.
- Chen, S.-l. 2010, « Self-Governing Online Communities in Web 2.0 : Privacy, Anonymity and Accountability in Wikipedia », *Alb. LJ Sci. & Tech*, vol. 20, p. 421–432.
- Chesney, T. 2006, « An empirical examination of Wikipedia's credibility », *First Monday*, vol. 11, n° 11. URL http://firstmonday.org/issues/issue11_11/chesney/index.html.
- Chiu, C.-M., M.-H. Hsu et E. T. Wang. 2006, « Understanding knowledge sharing in virtual communities : An integration of social capital and social cognitive theories », *Decision Support Systems*, vol. 42, n° 3, doi :10.1016/j.dss.2006.04.001, p. 1872–1888, ISSN 01679236. URL <http://linkinghub.elsevier.com/retrieve/pii/S0167923606000583>.
- Ciampaglia, G. L. 2011, *User Participation and Community Formation in Peer Production Systems*, thèse de doctorat, Università della Svizzera Italiana.
- Ciampaglia, G. L. et A. Vancheri. 2010, « Empirical Analysis of User Participation in Online Communities : The Case of Wikipedia », dans *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, Washington, DC, p. 219–222.
- Colón-Aguirre, M. et R. A. Fleming-May. 2012, « “You Just Type in What You Are Looking For” : Undergraduates' Use of Library Resources vs. Wikipedia », *The Journal of Academic Librarianship*, vol. 38, n° 6, doi :10.1016/j.acalib.2012.09.013, p. 391–399, ISSN 00991333. URL <http://linkinghub.elsevier.com/retrieve/pii/S0099133312001462>.
- Cunningham, W. et B. Leuf. 2001, *The Wiki Way : Quick Collaboration on the Web*, Addison-Wesley Professional, Boston, US, 464 p..
- Diderot, D. 1755, « Encyclopedia », .
- Ehmann, K., A. Large et J. Beheshti. 2008, « Collaboration in context : Comparing article evolution among subject disciplines in Wikipedia », *First Monday*, vol. 13, n° 10, p. 1–23.

- Epley, N., A. Waytz et J. T. Cacioppo. 2007, « On Seeing Human : A Three-On Seeing Human : A Three-Factor Theory of Anthropomorphism », *Psychological Review*, vol. 114, n° 4, doi :10.1037/0033-295X.114.4.864, p. 864–886.
- Fallis, D. 2008, « Toward an Epistemology of Wikipedia », *Journal of the American Society for Information Science and Technology*, vol. 59, n° 5, doi :10.1002/asi, p. 1662–1674.
- Faraj, S., S. L. Jarvenpaa et A. Majchrzak. 2011, « Knowledge Collaboration in Online Communities », *Organization Science*, vol. 22, n° 5, doi :10.1287/orsc.1100.0614, p. 1224–1239, ISSN 1047-7039. URL <http://orgsci.journal.informs.org/cgi/doi/10.1287/orsc.1100.0614>.
- Ferron, M. et P. Massa. 2013, « Beyond the encyclopedia : Collective memories in Wikipedia », *Memory Studies*, vol. 7, n° 1, doi :10.1177/1750698013490590, p. 22–45, ISSN 1750-6980. URL <http://mss.sagepub.com/cgi/doi/10.1177/1750698013490590>.
- Forte, A., N. Kittur, H. Zhu, V. Larco, A. Bruckman et R. E. Kraut. 2012, « Coordination and Beyond : Social Functions of Groups in Open Content Production », dans *CSCW '12*, ACM Press, Seattle, Washington, USA, ISBN 9781450310864, p. 1–10.
- Freitag, M. 2011, *Introduction à une théorie générale du symbolisme*, Liber éd., Diffusion Dimedia, Montréal, Québec, Canada, 481 p..
- Geiger, R. S. 2009, « The social roles of bots and assisted editing programs », dans *Proceedings of the 5th International Symposium on Wikis and Open Collaboration - WikiSym '09*, ACM Press, Orlando, Florida, USA, ISBN 9781605587301, p. 1–2, doi :10.1145/1641309.1641351. URL <http://portal.acm.org/citation.cfm?doid=1641309.1641351>.
- Geiger, R. S. 2011, « The Lives of Bots », dans *Critical point of view : A Wikipedia Reader*, édité par G. Lovink et N. Tkacz, INC READER éd., Institute of Network Cultures, Amsterdam, The Netherlands, ISBN 9789078146131, p. 78–93.
- Geiger, R. S. 2013, « Are Computers Merely “ Supporting ” Cooperative Work ? Towards an Ethnography of Bot Development », dans *CSCW '13 Companion*, ACM Press, San Antonio, Texas, USA, ISBN 9781450313322, p. 51–55.
- Geiger, R. S. 2014, « Bots, bespoke code, and the materiality of software platforms », *Information, Communication & Society*, vol. 17, n° 3, doi :10.1080/1369118X.2013.873069, p. 342–356, ISSN 1369-118X. URL <http://www.tandfonline.com/doi/abs/10.1080/1369118X.2013.873069>.
- Geiger, R. S., A. Halfaker, M. Pinchuk et S. Walling. 2012, « Defense Mechanism or Socialization Tactic ? Improving Wikipedia’s Notifications to Rejected Contributors », dans *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, Palo Alto, CA, USA, p. 122–129.

- Geiger, R. S., S. Hall, B. Ca et H. Ford. 2011, « Participation in Wikipedia’s Article Deletion Processes », dans *WikiSym’11*, ACM Press, Mountain View, CA, USA, ISBN 9781450309097, p. 201–202.
- Geiger, R. S., S. Hall et A. Halfaker. 2013, « When the Levee Breaks : Without Bots, What Happens to Wikipedia’s Quality Control Processes? », dans *WikiSym 2013*, ACM Press, Hong Kong, China, p. 0–5.
- Geiger, R. S. et D. Ribes. 2010, « The Work of Sustaining Order in Wikipedia : The Banning of a Vandal », dans *CSCW 2010*, ACM Press, Savannah, Georgia, USA, ISBN 9781605587950, p. 117–126.
- Gelley, B. S. et T. Suel. 2013, « Automated Decision Support for Human Tasks in a Collaborative System : The Case of Deletion in Wikipedia », dans *WikiSym 2013*, ACM Press, Hong Kong, China, p. 1–12.
- Gerardo, I., T. Yasseri, M. S. Miguel et K. Kaski. 2012, « Opinions, Conflicts and Consensus : Modeling Social Dynamics in a Collaborative Environment », *Physical Review Letters*, vol. 110, n° 8, p. 088 701.
- Giles, J. 2005, « Internet encyclopaedias go head to head. », *Nature*, vol. 438, n° 7070, doi :10.1038/438900a, p. 900–1, ISSN 1476-4687. URL <http://www.ncbi.nlm.nih.gov/pubmed/16355180>.
- Glaser, B. G. et A. L. Strauss. 1967, *The Discovery of Grounded Theory : Strategies for Qualitative Research*, reprint éd., Aldine, 271 p..
- Golbeck, J. et P. Mutton. 2004, « Semantic Web Interaction on Internet Relay Chat », dans *Proceedings of Interaction Design on the Semantic Web*, Springer Science & Business Media, p. 1–5.
- Goodrich, M. A. et A. C. Schultz. 2007, « Human–Robot Interaction : A Survey », *Human-Computer Interaction*, vol. 1, n° 3, doi :10.1561/1100000005, p. 203–275.
- Gorgeon, A. et E. B. Swanson. 2011, « Web 2.0 According to Wikipedia : Capturing an Organizing Vision », *Journal of the American Society for Information Science and Technology*, vol. 62, n° 10, doi :10.1002/asi, p. 1916–1932.
- Guitton, M. J. 2012, « The immersive impact of meta-media in a virtual world », *Computers in Human Behavior*, vol. 28, n° 2, doi :10.1016/j.chb.2011.10.016, p. 450–455, ISSN 07475632. URL <http://linkinghub.elsevier.com/retrieve/pii/S0747563211002330>.
- Halfaker, A., R. S. Geiger, J. T. Morgan et J. Riedl. 2012, « The Rise and Decline of an Open Collaboration System : How Wikipedia’s Reaction to Popularity Is Causing Its Decline »,

- American Behavioral Scientist*, vol. 57, n° 5, doi :10.1177/0002764212469365, p. 664–688, ISSN 0002-7642. URL <http://abs.sagepub.com/cgi/doi/10.1177/0002764212469365>.
- Halfaker, A., A. Kittur et J. Riedl. 2011, « Don't Bite the Newbies : How Reverts Affect the Quantity and Quality of Wikipedia Work », dans *WikiSym'11*, ACM Press, Mountain View, CA, USA, ISBN 9781450309097, p. 163–172.
- Halfaker, A., R. Kraut et J. Riedl. 2009, « A Jury of Your Peers : Quality, Experience and Ownership in Wikipedia », dans *WikiSym'09*, ACM Press, Orlando, Florida, USA, ISBN 9781605587301, p. 1–10.
- Halfaker, A. et J. Riedl. 2012, « Bots and Cyborgs : Wikipedia's Immune System », *Social Computing*, vol. 12, n° 3, p. 79–82.
- Hinds, P., T. Roberts et H. Jones. 2004, « Whose Job Is It Anyway? A Study of Human-Robot Interaction in a Collaborative Task », *Human-Computer Interaction*, vol. 19, n° 1, doi :10.1207/s15327051hci1901\&2_7, p. 151–181, ISSN 0737-0024. URL http://www.informaworld.com/openurl?genre=article&doi=10.1207/s15327051hci1901&2_7&magic=crossref||D404A21C5BB053405B1A640AFFD44AE3.
- Holtgraves, T., S. Ross, C. Weywadt et T. Han. 2007, « Perceiving artificial social agents », *Computers in Human Behavior*, vol. 23, n° 5, doi :10.1016/j.chb.2006.02.017, p. 2163–2174, ISSN 07475632. URL <http://linkinghub.elsevier.com/retrieve/pii/S0747563206000392>.
- Hsu, M.-H., T. L. Ju, C.-H. Yen et C.-M. Chang. 2007, « Knowledge sharing behavior in virtual communities : The relationship between trust, self-efficacy, and outcome expectations », *International Journal of Human-Computer Studies*, vol. 65, n° 2, doi :10.1016/j.ijhcs.2006.09.003, p. 153–169, ISSN 10715819. URL <http://linkinghub.elsevier.com/retrieve/pii/S1071581906001431>.
- Jahnke, I. 2010, « Dynamics of social roles in a knowledge management community », *Computers in Human Behavior*, vol. 26, n° 4, doi :10.1016/j.chb.2009.08.010, p. 533–546, ISSN 07475632. URL <http://linkinghub.elsevier.com/retrieve/pii/S0747563209001411>.
- Joyce, E., J. C. Pike et B. S. Butler. 2012, « Rules and Roles vs. Consensus : Self-Governed Deliberative Mass Collaboration Bureaucracies », *American Behavioral Scientist*, vol. 20, n° 10, doi :10.1177/0002764212469366, p. 1–19, ISSN 0002-7642. URL <http://abs.sagepub.com/cgi/doi/10.1177/0002764212469366>.
- Judd, T. et G. Kennedy. 2011, « Expediency-based practice? Medical students' reliance on Google and Wikipedia for biomedical inquiries », *British Journal of Educational Technology*, vol. 42, n° 2, doi :10.1111/j.1467-8535.2009.01019.x, p. 351–360, ISSN 00071013. URL <http://doi.wiley.com/10.1111/j.1467-8535.2009.01019.x>.

- Jullien, N. 2012, « What We Know About Wikipedia : A Review of the Literature Analyzing the Project(s) », *SSRN Electronic Journal*, doi :10.2139/ssrn.2053597, p. 1–86, ISSN 1556-5068. URL <http://www.ssrn.com/abstract=2053597>.
- Kennedy, K. 2009, « Textual Machinery : Authorial Agency and Bot-Written Texts in Wikipedia », *The Responsibilities of Rhetoric : Proceedings of thw 2008 Rhetoric Society of America Conference*, p. 1–7.
- Kittur, A., E. Chi, B. A. Pendleton, B. Suh et T. Mytkowicz. 2007a, « Power of the Few vs. Wisdom of the Crowd : Wikipedia and the Rise of the Bourgeoisie », *World Wide Web*, vol. 1, n° 2, p. 1–9.
- Kittur, A. et R. E. Kraut. 2008, « Harnessing the Wisdom of Crowds in Wikipedia : Quality Through Coordination », dans *CSCW'08*, ACM Press, San Diego, California, USA, ISBN 9781605580074, p. 37–46.
- Kittur, A., B. Suh, B. A. Pendleton, E. H. Chi, L. Angeles et P. Alto. 2007b, « He Says, She Says : Conflict and Coordination in Wikipedia + », dans *CHI 2007*, ACM Press, San Jose, CA, USA, ISBN 9781595935939, p. 453–462.
- Knight, C. et S. Pryke. 2012, « Wikipedia and the University, a case study », *Teaching in Higher Education*, vol. 17, n° 6, p. 649–659.
- Konieczny, P. 2014, « Rethinking Wikipedia for the Classroom », *American Sociological Association*, vol. 13, n° 1, doi :10.1177/1536504214522017, p. 80–84.
- Kriplean, T., I. Beschastnikh et D. W. Mcdonald. 2008, « Articulations of WikiWork : Uncovering Valued Work in Wikipedia through Barnstars », dans *CSCW'08*, ACM Press, San Diego, California, USA, ISBN 9781605580074, p. 1–10.
- Kriplean, T., I. Beschastnikh, D. W. Mcdonald et S. A. Golder. 2007, « Community, Consensus, Coercion, Control : CS*W or How Policy Mediates Mass Participation », dans *GROUP'07*, ACM Press, Sanibel Island, Florida, USA, ISBN 9781595938459, p. 1–10.
- Kuznetsov, S. 2006, « Motivations of contributors to Wikipedia », *ACM SIGCAS Computers and Society*, vol. 36, n° 2, doi :10.1145/1215942.1215943, p. 1–7, ISSN 00952737. URL <http://portal.acm.org/citation.cfm?doid=1215942.1215943>.
- Laniado, D., C. Castillo, A. Kaltenbrunner et M. F. Morell. 2012, « Emotions and dialogue in a peer-production community : the case of Wikipedia », dans *WikiSym'12*, ACM Press, Linz, Austria, ISBN 9781450316057, p. 1–10.
- Laniado, D. et R. Tasso. 2011, « Co-authorship 2.0 : Patterns of collaboration in Wikipedia », dans *HT'11*, ACM Press, Eindhoven, The Netherlands, ISBN 9781450302562, p. 201–210.

- Latour, B. 1991, *Nous n'avons jamais été modernes. Essai d'anthropologie symétrique.*, l'ar-millière éd., La Découverte, Paris, France, 211 p..
- Latour, B. 1992, « “Where Are the Missing Masses? The Sociology of a Few Mundane Ar-tifacts” », dans *Shaping Technology/Building Society : Studies in Sociotechnical Change*, chap. 10, MIT Press, Cambridge, MA, USA, p. 225–258.
- Laurent, M. R. et T. J. Vickers. 2009, « Seeking health information online : does Wikipedia matter ? », *Journal of the American Medical Informatics Association*, vol. 16, n° 4, doi : 10.1197/jamia.M3059, p. 471–9, ISSN 1067-5027. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2705249&tool=pmcentrez&rendertype=abstract>.
- Lawler, C. 2006, « A ‘resource review’ of Wikipedia », *Counselling and Psychotherapy Re-search*, vol. 6, n° 3, doi :10.1080/14733140600862279, p. 215–216, ISSN 1473-3145. URL <http://www.tandfonline.com/doi/abs/10.1080/14733140600862279>.
- Lessig, L. 1999, *Code and Other Laws of Cyberspace*, Basic Book, ISBN 9780465039135, 297 p..
- Lortie, C. L. et M. J. Guitton. 2011, « Judgment of the humanness of an interlocutor is in the eye of the beholder. », *PLoS ONE*, vol. 6, n° 9, doi :10.1371/journal.pone.0025085, p. 1–7, ISSN 1932-6203. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3178592&tool=pmcentrez&rendertype=abstract>.
- Loveland, J. et J. Reagle. 2013, « Wikipedia and encyclopedic production », *New Media & Society*, vol. 15, n° 8, doi :10.1177/1461444812470428, p. 1294–1311, ISSN 1461-4448. URL <http://nms.sagepub.com/cgi/doi/10.1177/1461444812470428>.
- M. Livingstone, R. 2012, *Network of knowledge : Wikipedia as a sociotechnical system of intelligence*, thèse de doctorat, PHD Thesis - University of Oregon.
- Martin, O. S. 2010, « A Wikipedia Literature Review », *eprint arXiv :1110.5863*.
- Mateas, M. et P. Sengers. 1998, *Narrative Intelligence*, John Benjamins, Californie, 340 p..
- May, C. et S. K. Sell. 2005, *Intellectual Property Rights : A Critical History*, Lynne Rienner Pub, 253 p..
- Mesgari, M., C. Okoli, M. Mehdi, F. A. r. Nielsen et A. Lanamaki. 2014, « “ The sum of all hu-man knowledge ” : A systematic review of scholarly research on the content of Wikipedia », *Journal of American Society for Information Science and Technology*, p. pre-print.
- Metcalfe, B. 1995, « Metcalfe’s law : A network becomes more valuable as it reaches more users », *Infoworld*, vol. 17, n° 40, p. 53–54.

- Mitter, S., C. Wagner et M. Strohmaier. 2013, « A Categorization Scheme for Socialbot Attacks In Online Social Networks », dans *WebSci'13*, ACM Press, Paris, France, ISBN 9781450318891, p. 1–6.
- Muchnik, L., S. Pei, L. C. Parra, S. D. S. Reis, J. S. Andrade, S. Havlin et H. A. Makse. 2013, « Origins of power-law degree distribution in the heterogeneity of human activity in social networks », *Scientific Reports*, vol. 3, n° 1783, doi :10.1038/srep01783, p. 1–7, ISSN 2045-2322. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3646275&tool=pmcentrez&rendertype=abstract>.
- Müller-Birn, C., J. D. Herbsleb et L. Dobusch. 2013, « Work-to-Rule : The Emergence of Algorithmic Governance in Wikipedia », dans *C&T'13*, ACM Press, Munich, Germany, ISBN 9781450321044, p. 1–10.
- Nass, C., B. J. Fogg et Y. Moon. 1996, « Can computers be teammates ? », *Human-Computer Studies*, vol. 45, p. 669–678.
- Nass, C. et Y. Moon. 2000, « Machines and Mindlessness : Social Responses to Computers », *Journal of Social Issues*, vol. 56, n° 1, doi :10.1111/0022-4537.00153, p. 81–103, ISSN 0022-4537. URL <http://doi.wiley.com/10.1111/0022-4537.00153>.
- Nazir, F. et H. Takeda. 2008, « Extraction and Analysis of Tripartite Relationships from Wikipedia », *Technology and Society*, vol. 8, n° 2008, p. 1–13.
- Niederer, S. et J. van Dijck. 2010, « Wisdom of the crowd or technicity of content? Wikipedia as a sociotechnical system », *New Media & Society*, vol. 12, n° 8, doi : 10.1177/1461444810365297, p. 1368–1387, ISSN 1461-4448. URL <http://nms.sagepub.com/cgi/doi/10.1177/1461444810365297>.
- Okoli, C., M. Mehdi, M. Mesgari, F. A. Nielsen et A. Lanamäki. 2012, « The People's Encyclopedia Under the Gaze of the Sages : A Systematic Review of Scholarly Research on Wikipedia », *SSRN Electronic Journal*, doi :10.2139/ssrn.2021326, p. 1–138, ISSN 1556-5068. URL <http://www.ssrn.com/abstract=2021326>.
- Okoli, C., M. Mehdi, M. Mesgari, F. A. R. Nielsen et A. Lanamaki. 2014, « Wikipedia in the eyes of its beholders : A systematic review of scholarly research on Wikipedia readers and readership », *Journal of the American Society for Information Science and Technology*, p. 1–39.
- Ortega, F., J. M. Gonzalez-barahona et G. Robles. 2008, « On The Inequality of Contributions to Wikipedia », dans *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, IEEE, Hawaii, US, p. 1–7.

- O'Sullivan, D. 2011, « What is an Encyclopedia ? From Pliny to Wikipedia », dans *Critical point of view : A Wikipedia Reader*, édité par G. Lovink et N. Tkacz, INC READER éd., Institute of Network Cultures, Amsterdam, The Netherlands, p. 34–49.
- Paechter, M., B. Maier et D. Macher. 2010, « Students' expectations of, and experiences in e-learning : Their relation to learning achievements and course satisfaction », *Computers & Education*, vol. 54, n° 1, doi :10.1016/j.compedu.2009.08.005, p. 222–229, ISSN 0360-1315. URL <http://dx.doi.org/10.1016/j.compedu.2009.08.005>.
- Paepcke, S. et L. Takayama. 2010, « Judging a Bot By Its Cover : An Experiment on Expectation Setting for Personal Robots », dans *Proceedings of Human-Robot Interaction Conference*, ACM Press, Osaka, Japan, p. 45–52.
- Päivärinta, T. et A. Lanamäki. 2010, « Genre Analysis of Metacommunication in Wikipedia », dans *Iris : Selected Papers of the Information Systems Reserach Seminar in Scandinavia*, The IRIS Association, Denmark, p. 147–162.
- Panciera, K., A. Halfaker et L. Terveen. 2009, « Wikipedians Are Born , Not Made A Study of Power Editors on Wikipedia », dans *GROUP'09*, ACM Press, Sanibel Island, Florida, USA, ISBN 9781605585000, p. 51–60.
- Pastinelli, M. 2007, *Des souris, des hommes et des femmes au village global : parole, pratiques identitaires et lien social dans un espace de bavardage électronique.*, Presses Université Laval, Québec, 322 p..
- Pastinelli, M. 2009, « La mémoire et l'oubli dans l'univers de l'archive totale », URL <http://www.espacestemp.net/en/articles/la-memoire-et-lrsquooubli-dans-lrsquounivers-de-lrsquoarchive-totale-en/>.
- Pastinelli, M. 2014, « Habiter le temps réel : ethnographie des modalités de l'«être ensemble» dans l'espace électronique », *Anthropologie et Sociétés*, vol. 30, n° 2, doi :10.7202/014120ar, p. 199–217.
- Priedhorsky, R., J. Chen, S. T. K. Lam, K. Panciera, L. Terveen et J. Riedl. 2007, « Creating, destroying, and restoring value in wikipedia », dans *Proceedings of the 2007 international ACM conference on Conference on supporting group work - GROUP '07*, ACM Press, Sanibel Island, Florida, USA, ISBN 9781595938459, p. 259, doi :10.1145/1316624.1316663. URL <http://portal.acm.org/citation.cfm?doid=1316624.1316663>.
- Ransbotham, S. et G. C. Kane. 2011, « Membership Turnover and Collaboration Success in Online Communities : Explaining Rises and Falls from Grace in Wikipedia », *MIS Quarterly*, vol. 35, n° 3, p. 1–38.

- Reagle, J. M. 2010a, « "Be Nice" : Wikipedia norms for supportive communication », *New Review of Hypermedia and supportive communication*, vol. 16, n° 1-2, doi :10.1080/13614568.2010.498528, p. 161–180.
- Reagle, J. M. 2010b, « "Be Nice" : Wikipedia norms for supportive communication », *New Review of Hypermedia and Multimedia*, vol. 16, n° 1-2, doi :10.1080/13614568.2010.498528, p. 161–180, ISSN 1361-4568. URL <http://www.tandfonline.com/doi/abs/10.1080/13614568.2010.498528>.
- Remy, J. et D. Ruquoy. 1936, *Méthodes d'analyse de contenu et sociologie*, Fac St Louis, Paris, 241 p..
- Restivo, M. et A. van de Rijt. 2012, « Experimental study of informal rewards in peer production. », *PLoS ONE*, vol. 7, n° 3, doi :10.1371/journal.pone.0034358, p. e34358, ISSN 1932-6203. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3315525&tool=pmcentrez&rendertype=abstract>.
- Ridge, M. 2013, « New Challenges in Digital History : Sharing Women's History on Wikipedia », dans *Women's History in the Digital World*, vol. 37, Open University, p. 1–30.
- Rutjes, J. 2013, *Expectations and Human-Robot Interaction : The influence of robot expectations on personality attribution, impressions and anthropomorphism*, thèse de doctorat, University of Twente.
- Saygin, A. P., I. Cicekli et V. Akman. 2000, « Turing Test : 50 Years Later », *Minds and Machines*, vol. 10, n° 4, p. 463–518.
- Schneider, J., A. Passant et J. Breslin. 2010, « A qualitative and quantitative analysis of how wikipedia talk pages are used. », dans *Proceedings of the WebSci10 : Extending the Frontiers of Society*, Raleigh, NC, USA, p. 1–7.
- Schneider, J., A. Passant et S. Decker. 2012, « Deletion Discussions in Wikipedia : Decision Factors and Outcomes », dans *WikiSym'12*, ACM Press, Linz, Austria, ISBN 9781450316057, p. 1–10.
- Shao, G. 2009, « Understanding the appeal of user-generated media : a uses and gratification perspective », *Internet Research*, vol. 19, n° 1, doi :10.1108/10662240910927795, p. 7–25, ISSN 1066-2243. URL <http://www.emeraldinsight.com/10.1108/10662240910927795>.
- Soto, J. F. O. 2009, *Wikipedia : A quantitative analysis*, thèse de doctorat, Universidad Rey Juan Carlos.
- Srivastava, S., S. Guglielmo et J. S. Beer. 2010, « Perceiving Others' Personalities : Examining the Dimensionality, Assumed Similarity to the Self, and Stability of Perceiver Effects »,

- Journal of Personality and Social Psychology*, vol. 98, n° 3, doi :10.1037/a0017057, p. 520–534.
- Steiner, T. 2014, « Bots vs. Wikipedians , Anons vs. Logged-Ins (Redux) : A Global Study of Edit Activity on Wikipedia and Wikidata », dans *OpenSym'14*, ACM Press, Hamburg, Germany, ISBN 9781450330169, p. 1–7.
- Stvilia, B., M. B. Twidale, L. E. S. Gasser et L. C. Smith. 2005, « Information Quality Discussions in Wikipedia », *Journal of the American Society for Information Science and Technology*, vol. 59, n° 6, p. 983–1001.
- Stvilia, B., M. B. Twidale, L. C. Smith et L. Gasser. 2008, « Information Quality Work Organization in Wikipedia », *Journal of the American Society for Information Science and Technology*, vol. 59, n° 6, doi :10.1002/asi, p. 983–1001.
- Suh, B., G. Convertino, E. H. Chi et P. Pirolli. 2009, « The Singularity is Not Near : Slowing Growth of Wikipedia », dans *WikiSym'09*, ACM Press, Orlando, Florida, USA, ISBN 9781605587301, p. 1–10.
- Suzor, N. et D. Woodford. 2013, « Evaluating Consent and Legitimacy amongst Shifting Community Norms : an EVE Online Case Study », *Journal of Virtual Worlds Research*, vol. 6, n° 3, p. 1–14.
- Taraborelli, D. et G. L. Ciampaglia. 2010, « Beyond Notability. Collective Deliberation on Content Inclusion in Wikipedia », *2010 Fourth IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshop*, doi :10.1109/SASOW.2010.26, p. 122–125. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5729608>.
- Thom-santelli, J., D. Cosley et G. Gay. 2009, « What ' s Mine is Mine : Territoriality in Collaborative Authoring », dans *CHI 2009*, ACM Press, Boston, MA, US, ISBN 9781605582467, p. 1–4.
- Thomas, K., C. Grier et V. Paxson. 2012, « Adapting Social Spam Infrastructure for Political Censorship », dans *Proceedings of the 5th USENIX conference on Large-Scale Exploits and Emergent Threats*, USENIX Association, p. 1–8.
- Velasco, S. M. M. 2010, « Wikipedia Vandalism Detection Through Machine Learning : Feature Review and New Proposals », dans *PAN'10*, p. 1–8.
- Viégas, F. B. 2007, « The Visual Side of Wikipedia », dans *Proceedings of the 40th Hawaii International Conference on System Sciences*, Hawaii, p. 1–10.
- Viégas, F. B., M. Wattenberg et M. M. McKeon. 2007, « The Hidden Order of Wikipedia », *Online communities ans social computing*, p. 445–454.

- Voss, J. 2005, « Measuring Wikipedia », dans *Proceedings 10th International Conference of the International Society for Scientometrics and Informetrics*, Stockholm, Sweden, p. 221–231.
- Wagner, C. et P. Prasarnphanich. 2007, « Innovating Collaborative Content Creation : The Role of Altruism and Wiki Technology », dans *Proceedings of the 40th Hawaii International Conference on System Sciences*, Ieee, Hawaii, p. 1–10.
- Waters, N. L. 2007, « Why You Can't Cite Wikipedia in My Class », *Communications of the ACM*, vol. 50, n° 9, p. 15–17.
- Welser, H. T., D. Cosley, G. Kossinets, A. Lin, F. Dokshin, G. Gay et M. Smith. 2011, « Finding social roles in Wikipedia », dans *Proceedings of the 2011 iConference on - iConference '11*, ACM Press, Seattle, WA, USA, ISBN 9781450301213, p. 122–129, doi :10.1145/1940761.1940778. URL <http://portal.acm.org/citation.cfm?doid=1940761.1940778>.
- Yang, H.-L. et C.-Y. Lai. 2010, « Motivations of Wikipedia content contributors », *Computers in Human Behavior*, vol. 26, n° 6, doi :10.1016/j.chb.2010.04.011, p. 1377–1383, ISSN 07475632. URL <http://linkinghub.elsevier.com/retrieve/pii/S0747563210000877>.
- Yang, W., W. Liu, A. Viña, M.-N. Tuanmu, G. He, T. Dietz et J. Liu. 2013, « Nonlinear effects of group size on collective action and resource outcomes. », *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, n° 27, doi :10.1073/pnas.1301733110, p. 10 916–21, ISSN 1091-6490. URL <http://www.ncbi.nlm.nih.gov/pubmed/23776222>.
- Yasseri, T. et János Kertész. 2013, « Value production in a collaborative environment : socio-physical studies of Wikipedia », *Journal of Statistical Physics, Online First*, p. 1–21.
- Yasseri, T., A. Spoerri, M. Graham et J. Kertész. 2014, « The most controversial topics in Wikipedia : A multilingual and geographical analysis », dans *Global Wikipedia : International and cross-cultural issues in online collaboration.*, Scarecrow Press, p. 1–31.
- Yasseri, T., R. Sumi, A. Rung, A. Kornai et J. Kertész. 2012, « Dynamics of conflicts in Wikipedia. », *PLoS ONE*, vol. 7, n° 6, doi :10.1371/journal.pone.0038869, p. 1–12, ISSN 1932-6203. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3380063&tool=pmcentrez&rendertype=abstract>.
- Zdenek, S. 2003, « Artificial intelligence as a discursive practice : the case of embodied software agent systems », *AI & SOCIETY*, vol. 17, n° 3-4, p. 340–363.
- Zelenkauskaitė, A. et P. Massa. 2012, « Contributing to Wikipedia : Through Content or Social Interaction », *International Journal of Distributed Systems and Technologies*, vol. 3, n° 4, doi :10.4018/jdst.2012100101, p. 1–13, ISSN 1947-3532. URL <http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/jdst.2012100101>.

Zhu, H., R. Kraut et A. Kittur. 2012, « Effectiveness of shared leadership in online communities », dans *CSCW '12*, ACM Press, Seattle, Washington, USA, ISBN 9781450310864, p. 407, doi :10.1145/2145204.2145269. URL <http://dl.acm.org/citation.cfm?doid=2145204.2145269>.

Zhu, H., A. Zhang, J. He, R. E. Kraut et A. Kittur. 2013, « Effects of Peer Feedback on Contribution : A Field Experiment in Wikipedia », dans *CHI 2013*, ACM Press, Paris, France, ISBN 9781450318990, p. 1–10.