

Démographie des communautés en ligne: le cas des wikis

Camille Roth*[†]
roth@ehess.fr

Dario Taraborelli[†]
d.taraborelli@surrey.ac.uk

Nigel Gilbert[†]
n.gilbert@surrey.ac.uk

31 octobre 2008

Résumé

Les communautés en ligne collaboratives ont connu un succès massif avec l'émergence des services et des plates-formes *Web 2.0*. Les wikis, et notamment la *Wikipedia* sont un des exemples les plus saillants de ce type de communautés de construction collective de contenus. La *Wikipedia* a à cet égard jusqu'ici concentré l'essentiel des efforts de recherche au sujet de ces communautés, même si l'ensemble des wikis constitue un écosystème possédant une très grande diversité de contenus, de populations, d'usages, de systèmes de gouvernance. Au contraire de la *Wikipedia* qui a probablement atteint la masse critique lui permettant d'être viable, la plupart des wikis luttent pour survivre et sont en compétition afin d'attirer contributeurs et articles de qualité, connaissant ainsi des destinées variées, vertueuses – croissance en population et en contenu – ou fatales – inactivité et vandalisme.

Aucune étude exhaustive de la structure et de l'évolution de l'écosystème des wikis n'a pu jusqu'ici être menée du fait, partiellement, du manque de données concernant d'autres wikis que la *Wikipedia*. Notre étude constitue ainsi la première analyse quantitative et dynamique d'un grand échantillon de wikis. Nous nous intéressons en particulier aux corrélations potentielles entre la dynamique démographique de ces wikis et certaines de leurs propriétés liées à leur structure et leur mode de gouvernance. Ceci nous permet, de manière plus globale, d'introduire et de discuter plusieurs axes d'étude de la viabilité des communautés en ligne.

1 Introduction

Les communautés en ligne ont démontré leur capacité à permettre et encourager la production collaborative de grandes quantités de contenu. Les projets de développement de logiciels libres tels que Mozilla Firefox ou Linux [Godfrey & Tu, 2000] ou bien de création d'encyclopédies comme la *Wikipedia*¹ [Giles, 2005] en sont de célèbres exemples. Toutefois, les projets de ce type ne connaissent pas tous un succès aussi franc : leur destin dépend notamment de la capacité des administrateurs à attirer de nouveaux membres, à mettre en place des règles qui entretiennent l'implication des contributeurs, et à garantir un certain niveau de qualité.

Les *wikis* sont, au sens large, des sites Web dont le contenu peut être produit et modifié par n'importe quel utilisateur, de manière collaborative et collective. En tant que tels, ils constituent un des exemples les plus saillants des systèmes collaboratifs basés sur le Web. Le plus célèbre et probablement le plus utilisé des wikis, la *Wikipedia*, a fait l'objet ces dernières années de nombreuses recherches portant sur divers aspects du site liés notamment à la sociologie des usages, l'épistémologie sociale, ou l'étude des activités de cognition distribuée [Lih, 2004; Anthony *et al.*, 2005; Bryant *et al.*, 2005]. Le Web comprend cependant des milliers d'autres wikis au succès très variable : si certains recrutent de nombreux utilisateurs et réussissent à être viables en assurant une mise à jour régulière, des mesures efficaces contre le vandalisme et une bonne répartition des tâches et des rôles entre les membres, d'autres wikis luttent considérablement pour attirer des contributeurs. Dans l'ensemble, les communautés "wikis" peuvent

*CAMS, CNRS/EHESS, 96, Bd Raspail, F-75006 Paris, France.

[†]CRESS, Department of Sociology, University of Surrey, Guildford GU2 7XH, United Kingdom.

¹Accessible à l'adresse <http://wikipedia.org>

aussi bien être viables en suivant une variété de règles et de principes, ou bien périlcliter malgré l'adoption de règles fructueuses ailleurs ; tous s'efforçant de survivre au sein de ce que l'on pourrait désigner par le terme de "wikisphère".

La présente étude vise à fournir un premier aperçu sur la problématique de la viabilité des communautés en ligne, en analysant en particulier les facteurs pouvant rendre compte des diverses destinées d'un wiki — facteurs liés à la gouvernance et aux normes structurant les actions et motivations des utilisateurs, ou bien liés aux propriétés structurelles et techniques des wikis. Nous entendons ici la notion de "viabilité" au sens large, en tant que survie de la population comme du contenu : autrement dit, un wiki "viable" doit pouvoir croître en termes d'articles et de membres de telle manière que le contenu puisse être entretenu par un nombre suffisant d'utilisateurs. Notre but n'est pas de fournir une définition formelle de la viabilité d'un wiki, mais plutôt de présenter une analyse descriptive détaillée des propriétés et de la dynamique d'un grand échantillon de wikis, dans la perspective d'ouvrir sur de futures études empiriques.

Puisque nous nous intéressons aux communautés en ligne, nous ne prenons pas en compte les wikis émergeant principalement de communautés pré-existantes (comme les wikis d'entreprise) car leur viabilité a de fortes chances de dépendre de celle des communautés sous-jacentes. Au contraire, nous nous focalisons sur les communautés existant en vertu de la plate-forme "wiki", c'est-à-dire telles que n'importe quel visiteur ait la possibilité de s'inscrire en tant que contributeur plein.²

En outre, décrire le paysage de la wikisphère requiert de dépasser le cas de la *Wikipedia* et d'analyser les similitudes et dissimilitudes entre wikis, à la fois de manière qualitative et quantitative. Dans une première partie, notre étude vise ainsi à proposer un typologie succincte des wikis constituant la wikisphère. Dans les sections suivantes, nous donnons un aperçu de la structure et de l'évolution d'un large échantillon de wikis, au travers d'une sélection de mesures démographiques et dynamiques simples. Nous concluons l'article par une discussion de ces résultats en suggérant plusieurs axes d'étude de la viabilité des wikis.

2 Similarité et diversité dans la wikisphère

Si les communautés "wikis" varient sensiblement quant à leur structure, organisation interne et mode de gouvernance, la plupart se reconnaissent fondamentalement par un certain nombre de propriétés communes.

Similarité fonctionnelle

Par définition, les communautés wikis sont toutes fondées sur une "plate-forme wiki" ou un "moteur wiki", c'est-à-dire un **logiciel** fournissant un service Web qui permette la modification collective de son propre contenu, tout du moins à un certain degré.³ En dépit de l'existence d'une grande variété d'interfaces⁴, les caractéristiques fonctionnelles facilitant l'édition collaborative de contenu restent relativement homogènes, comme par exemple la possibilité de créer un compte, de construire une identité en ligne, de modifier les articles de façon immédiate, de suivre l'historique des révisions et même, souvent, de dialoguer et argumenter avec les autres contributeurs par l'intermédiaire de pages de discussion [Viegas *et al.*, 2007].

²La distinction entre, d'une part, les communautés purement "en ligne" et, d'autre part, les communautés "hors ligne" avec une contrepartie "en ligne", peut sembler relativement floue lorsque des utilisateurs d'un wiki décident de se rencontrer dans le "monde réel" (par exemple, les membres de la *Wikipedia* qui assistent aux conférences annuelles "*WikiMania*") ou bien lorsque des individus décident de créer un wiki ouvert au public (comme dans le cas d'*iPodLinux*, wiki ouvert par les développeurs du système d'exploitation éponyme et dédié à l'écriture de sa documentation — accessible à l'adresse <http://ipodlinux.org>).

³La définition traditionnelle d'un wiki comme un site Web que "tout le monde peut éditer" ("*anyone can edit*") n'est plus rigoureusement exacte depuis l'apparition de wikis dont l'accès et/ou l'édition peuvent être restreints.

⁴Pour une présentation exhaustive des moteurs de wikis existants, voir <http://www.wikimatrix.org>

Une seconde propriété commune à tous les wikis tient au mode de mise à disposition du contenu — le **licensing**. Malgré la profusion des licences utilisées (de la *GNU Free Documentation License* aux licences *Creative Commons*, en passant par la licence *Art Libre*, entre autres), leur fonction reste en effet essentiellement identique : il s’agit de permettre la contribution de tout participant potentiel, tout en fournissant certaines garanties concernant l’utilisation future du contenu, sans néanmoins faire preuve d’autant de rigidité que le droit d’auteur classique. À l’extrême, certains wikis n’utilisent même pas de licence : les contributions font alors directement partie du domaine public (c’est le cas de *WikiWikiWeb*⁵ ou bien d’*ExampleProblems*⁶).

Enfin, une étude quantitative et technique menée par *Zlatic et al.* [2006] sur les réseaux d’articles des diverses *Wikipedias* (en plusieurs langues) a montré que ces wikis partagent aussi un certain nombre de **propriétés structurelles** liées à la connectivité des pages⁷ ; propriétés qui sont, plus largement, communes aux réseaux de pages Web.

Variété organisationnelle

Au-delà de cette homogénéité fonctionnelle, de nombreuses dimensions sémantiques et organisationnelles permettent de différencier les wikis. Les wikis varient en effet principalement en fonction de leur **champ**, et les projets généralistes visant à construire des encyclopédies diversifiées côtoient des wikis thématiques traitant d’objets plus spécialisés : par exemple, des zones géographiques particulières (comme *DavisWiki*⁸ qui rassemble du contenu concernant la ville de Davis en Californie), des domaines techniques précis (comme *ExampleProblems*⁶), ou bien encore du contenu politiquement orienté (comme “*Conservapedia*”⁹, qui affiche explicitement un ton conservateur). Le **langue** en vigueur fait aussi partie du champ d’une communauté wiki, quoique certains wikis visent à produire du contenu multilingue (auquel cas il est difficile de déterminer si la langue définit une unique communauté multilingue ou bien plusieurs communautés “monolingues”).

Le respect du champ d’application et l’usage d’une langue particulière font partie des principes les plus simples que doit respecter un contributeur bienveillant. En général toutefois, les wikis obéissent à de nombreux autres types de règles d’édition [*Bryant et al.*, 2005; *Levrel*, 2006; *Viegas et al.*, 2007] qui peuvent être explicites, et ainsi traditionnellement présentées dans des méta-articles dédiés, ou bien implicites, et de fait davantage assimilables à des normes. Certaines de ces normes peuvent par ailleurs se cristalliser et devenir ensuite des règles explicites ; l’agencement des principes au sein d’un wiki dépendant finalement des aléas particuliers de son développement historique. Par exemple, la règle dénommée “neutralité de point de vue” (“*Neutral of Point of View*”, NPOV) dans les projets encyclopédiques comme la *Wikipedia* ou le *Citizendium*¹⁰ n’est pas commune à tous les wikis : elle n’existe pas dans certains wikis thématiques focalisés par exemple sur le voyage (comme *WikiTravel*¹¹ ou *World66*¹²), sur les questions environnementales (comme *Ekopedia*¹³), et même certains projets explicitement encyclopédiques, comme *Wikinfo*¹⁴ qui encourage précisément pour un même article la diversité de points de vue et donc de versions.

⁵Le tout premier wiki, créé en 1995 et toujours accessible à l’adresse <http://c2.com/cgi/wiki>

⁶Un répertoire de problèmes de mathématiques accessible à l’adresse <http://exampleproblems.com>

⁷Ces réseaux exhibent notamment des distributions similaires de motifs topologiques locaux, des distributions de connectivité en loi-puissance, de la disassortativité (i.e., les pages très connectées sont généralement liées à des articles faiblement connectés) et un fort taux de transitivité/*clustering*

⁸Accessible à l’adresse <http://daviswiki.org>

⁹Accessible à l’adresse <http://www.conservapedia.com>

¹⁰*Citizendium* est une encyclopédie initiée en 2006 afin de concurrencer *Wikipedia* avec la particularité que les articles et leurs diverses versions sont soumis à l’approbation des pairs avant publication. Accessible à l’adresse <http://en.citizendium.org>

¹¹Accessible à l’adresse <http://wikitravel.org>

¹²Accessible à l’adresse <http://world66.com>

¹³Accessible a l’adresse <http://ekopedia.org>

¹⁴Accessible à l’adresse <http://www.wikinfo.org>

D'autres règles, enfin, ont trait à la conduite à adopter vis-à-vis des autres membres : il peut s'agir de prescriptions éditoriales par exemple, afin de gérer un conflit d'édition. Certains wikis ont à cet égard des groupes d'articles dédiés exclusivement à leurs propres règles, permettant ainsi aux membres de modifier eux-mêmes le mode de fonctionnement de leur wiki. En toute généralité, ces règles définissent la gouvernance du wiki et ont une influence directe sur la manière dont les interactions et les activités de cognition distribuée y ont lieu : les encyclopédies dont l'édition est asynchrone et à accès restreint, comme *Citizendium*, requièrent de toute évidence une structure institutionnelle différente de celle des wikis plus "anarchiques", comme *WikiWikiWeb*.

Éventuellement, ces différences dans le mode de régulation se retrouvent dans l'**agencement technique** du wiki, c'est-à-dire les particularités du logiciel sous-jacent qui contraignent concrètement les modes d'interaction entre utilisateurs eux-mêmes, et entre utilisateurs et contenus. Un passage en revue exhaustif des divergences techniques entre wikis n'est clairement pas l'objet du présent article mais pourrait être partir de questions telles que : *qui peut contribuer ?* (visiteurs anonymes ; membres enregistrés ; contributeurs approuvés), *comment s'organisent les interactions en ligne et les identités virtuelles ?* (existence de pages de membres ; de pages de discussion ; de médias hors du wiki, comme les canaux de dialogue *IRC* ou les listes de diffusion), *quelles fonctionnalités techniques sont disponibles ?* (existence de pages spéciales pour définir des catégories, des projets, des règles ; possibilité de télécharger du contenu multimédia). Ces différentes configurations, en tant que leviers d'action globaux, sont à même d'orienter l'évolution des wikis et d'avoir en retour un impact sur leur viabilité. Néanmoins, au vu de l'état de l'art de la recherche sur les communautés en ligne et notamment des données accessibles, une analyse de l'impact de ces propriétés précises semble relativement ardue. Dans la section suivante, nous nous focalisons ainsi sur un certain nombre d'indicateurs simples dont nous considérons qu'ils nous permettent déjà de caractériser la vaste majorité des wikis et, spécifiquement, de comprendre ce qui peut influencer leur développement en tant que communautés.

3 Démographie des wikis

Les divers modes de gouvernance et configurations logicielles possibles, c'est-à-dire les contraintes sociales et techniques, définissent un paysage au sein duquel chaque communauté en ligne s'insère, croît et vit. En tant que communautés de production de contenu, les wikis évoluent principalement suivant deux dimensions :

- (I) les **CONTRIBUTEURS**, qui peuvent (ou non) constituer une communauté active ; comme l'ont décrit par exemple [Bryant et al. \[2005\]](#) ;
- (II) les **ARTICLES**, qui peuvent (ou non) devenir du contenu crédible, utile et/ou de qualité ; comme le démontre par exemple [Giles \[2005\]](#).

Utilisateurs et pages sont susceptibles d'obéir à une dynamique duale : tandis que la présence de davantage d'utilisateurs permet de contribuer à davantage de pages, la prolifération du contenu semble demander davantage d'attention de la part des membres. À première vue, il peut ainsi sembler judicieux d'appréhender la santé d'un wiki au travers de ces variables, considérées ainsi des indicateurs démographiques de sa croissance et de son activité.¹⁵

¹⁵Soulignons ici le fait que la croissance brute du contenu n'est pas en soi un bon indicateur d'un wiki viable, comme le suggèrent de récentes études sur la "prolifération" au sein des wikis [[Happel & Treitz, 2008](#)].

3.1 Protocole empirique

3.1.1 Données

Nous avons construit une base de données recensant des statistiques simples concernant un grand nombre de wikis basés sur la plate-forme “MediaWiki”¹⁶ et qui nous permet de travailler sur le même ensemble de variables pour tous les wikis et nous assurer qu’elles sont à chaque fois disponibles. Les données ont été collectées entre août 2007 et avril 2008 à partir d’une autre base de données disponible publiquement¹⁷ et rassemblant plus de 11 500 wikis.

Exclusions concernant les plates-formes De nombreux wikis présents dans cette base de données sont sur des “fermes wikis”, i.e. des plates-formes hébergeant de nombreux wikis et fournissant des services facilitant la création et la gestion du wiki, à l’instar des plates-formes de blogs. Or, certaines de ces plates-formes renvoient des statistiques fausses concernant la ferme toute entière plutôt que chaque wiki individuellement. Pour éviter cet effet, nous avons exclu de la base de données tous les wikis hébergés par de telles plates-formes (comme par exemple *Wikia* qui renvoie le même nombre d’utilisateurs (~135k) et très peu de pages (≤ 10) quel que soit le wiki). Cette opération réduit la taille de notre échantillon à 7 412 wikis.

3.1.2 Variables

Nous considérons quatre variables quantitatives :

- la *taille de la population* (U), mesurée par le nombre d’utilisateurs *enregistrés* ;
- la *taille du contenu* (P), mesurée par le nombre de “bonnes pages” (i.e. les articles véritables, à l’exclusion des pages par défaut créées par le moteur wiki), ci-après dénommées indifféremment “pages” ou “articles” ;
- le *nombre d’administrateurs* (A), soit les utilisateurs ayant des droits spéciaux permettant de modifier du contenu sensible et d’effectuer certaines opérations de maintenance ;
- l’*activité d’édition* (E), mesurée par le nombre total d’éditions d’articles par des utilisateurs, membres ou non.

ainsi qu’une variable qualitative indiquant la présence d’un mécanisme de contrôle d’accès :¹⁸

- l’*autorisation d’édition* (R), i.e. la possibilité pour les utilisateurs anonymes et non-enregistrés de créer une page. R vaut soit 1, selon que l’édition est ouverte aux utilisateurs anonymes, ou 0, selon qu’elle est réservée aux utilisateurs enregistrés.

Aussi simplistes que puissent sembler ces variables, celles-ci fournissent d’ores et déjà des indicateurs-clés au sujet de la dynamique globale d’un wiki et mettent en lumière divers aspects de sa structure et de son évolution. Nous avons collecté et observé les valeurs de ces variables pour chaque wiki, chaque jour sur une période de 250 jours, soit approximativement 8 mois.

3.2 Paysage statique

Afin de donner un premier aperçu de la démographie de la wikisphère, nous examinons les distributions des populations des wikis et de la taille de leur contenu : la figure 1 représente le nombre de wikis qui ont un nombre donné d’utilisateurs et de pages. Nous observons ainsi que les wikis sont très

¹⁶Cette base de données inclut initialement, entre autres, un grand nombre de *Wikipedias* dans diverses langues.

¹⁷Accessible à l’adresse <http://s23.org/wikistats/largest.html.php>. Cette base est entretenue par un utilisateur appelé “Mutante” qui nous a gracieusement autorisé à faire usage de ses données.

¹⁸Cet indicateur ne faisait pas partie des données originales fournies par “Mutante”/s23 et a été récupéré au début de l’enquête en interrogeant chaque wiki de la base grâce à un robot permettant de déterminer automatiquement si la création d’article était possible sans s’enregistrer au préalable.

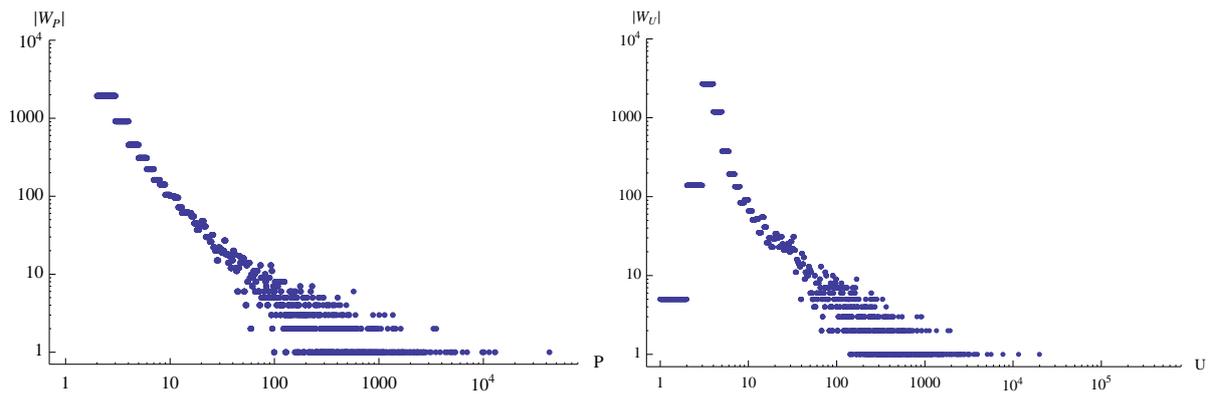


FIG. 1 – Distributions des wikis ayant P pages (W_P , à gauche) et U utilisateurs (W_U , à droite). Les deux axes sont à l'échelle logarithmique.

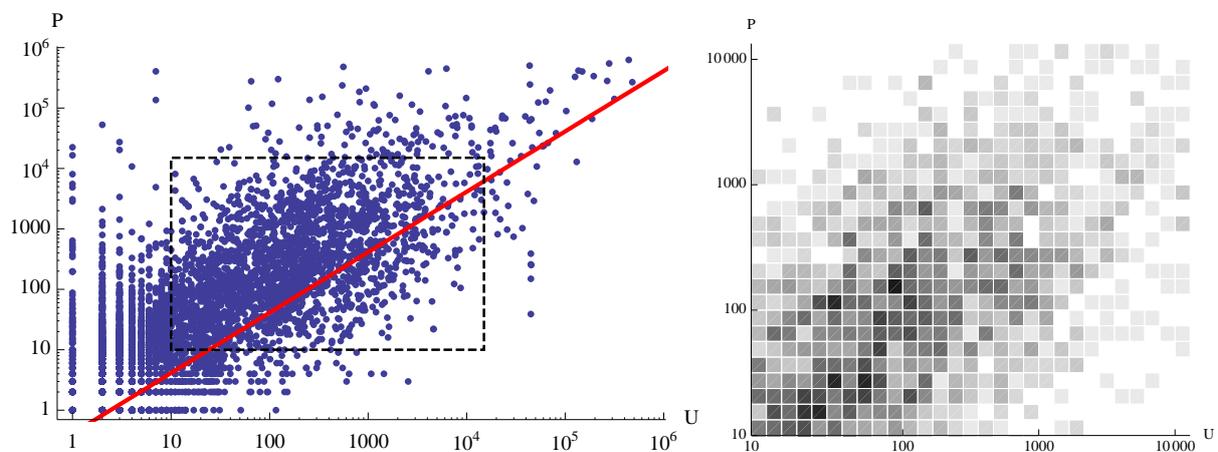


FIG. 2 – À gauche : nuage de points représentant le nombre de pages (P) en fonction du nombre d'utilisateurs (U). Chaque point correspond à un wiki, la ligne épaisse montre la meilleure régression linéaire. À droite : représentation en densité de P en fonction de U pour les wikis provenant de la région encadrée en pointillés dans le diagramme précédent. Les carrés les plus sombres correspondent à une densité de wikis plus élevée.

hétérogènes : beaucoup possèdent peu de pages tandis qu'une poignée ont un grand nombre d'utilisateurs et d'articles (plus précisément et techniquement, ces distributions suivent approximativement une loi "puissance").

Lorsque l'on examine la relation entre le nombre d'utilisateurs et la taille du contenu, nous voyons que pour 95% des wikis les moins peuplés ($U \leq 1\,113$), les valeurs de la taille du contenu sont étalées sur plusieurs ordres de grandeur : P varie entre 1 et environ 2 000. Signe supplémentaire de cette diversité, les régressions linéaires correspondent mal aux données et n'indiquent aucune échelle caractéristique des valeurs, comme le souligne le nuage de points de la figure 2. La plupart de ces wikis faiblement peuplés, ayant moins d'un millier de membres et donc étant probablement plus jeunes, peuvent correspondre à des wikis à divers stades de croissance et de succès. À l'inverse, lorsque le nombre d'utilisateurs est plus conséquent et, plus précisément, pour le premier centile des tailles de population ($U \geq 8\,111$), le spectre des valeurs est plus étroit et exhibe une relation plus monotone entre les variables.

Qualitativement, une interprétation possible de la grande zone dense sur la figure 2 serait qu'elle capture à la fois les wikis débutants (peu d'utilisateurs, peu ou beaucoup de pages) et certains wikis florissants (plus de membres, plus de pages), où toutes les configurations intermédiaires restent possibles

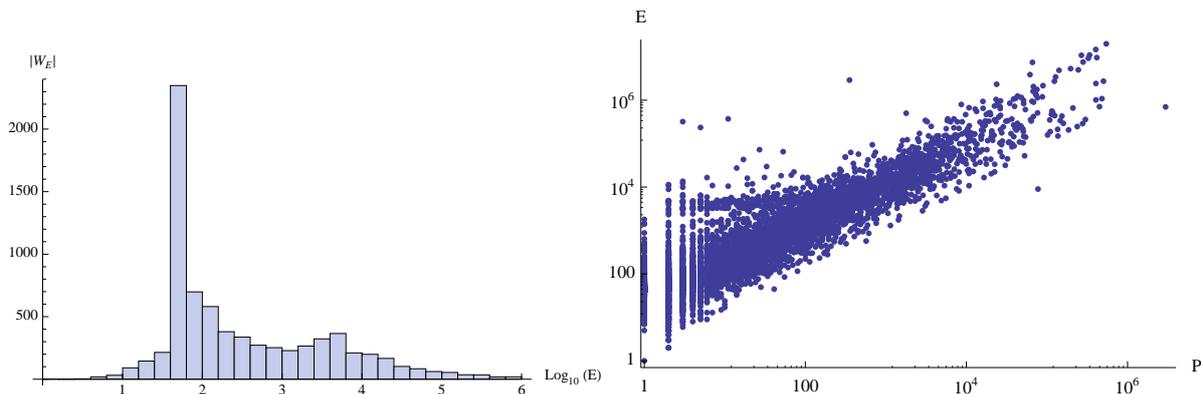


FIG. 3 – Activité, ou nombre d’édérations E : histogramme de la distribution (à gauche) et nuage de points en fonction du nombre de pages P (à droite).

et apparemment similairement probables. La zone correspondant aux wikis les plus peuplés, qui sont donc ceux qui ont eu un certain succès, semble comparativement beaucoup plus disciplinée. Quoiqu’il en soit, l’activité a quant à elle un comportement beaucoup plus régulier puisqu’elle est plus ou moins proportionnelle au nombre de pages (Fig. 3) : de fait, le niveau d’activité correspond très directement à la taille du contenu, ce qui n’est pas le cas pour la taille de la population.

3.3 Dynamique des wikis

La littérature concernant la dynamique des communautés “Web” reste peu fournie [Kumar *et al.*, 2006; Golbeck, 2007] et, à notre connaissance, notre étude représente la première analyse longitudinale du contenu et de la population d’un grand échantillon de wikis. Outre le fait qu’ils se soient essentiellement intéressés à la *Wikipedia*, les travaux quantitatifs précédents menés sur les wikis ont en effet principalement visé à examiner la topologie des réseaux d’interaction ou des liens hypertextes sous-jacents [Capocci *et al.*, 2006; Zlatic *et al.*, 2006], ou bien encore la structure du contenu au niveau des articles [Brandes & Lerner, 2008; Wilkinson & Huberman, 2007], en délaissant fréquemment les déterminants de la dynamique démographique elle-même (à l’exception de Kittur *et al.* [2006] qui compare la démographie des contributeurs réguliers vs. occasionnels sur la *Wikipedia*). Dans les sections suivantes, nous analysons l’évolution de notre échantillon de wikis en montrant de quelle manière certains indicateurs sont corrélés avec différents modes de croissance ; il s’agit en soi d’une étape préliminaire pour comprendre plus globalement les facteurs sous-tendant certains motifs d’évolution dans la wikisphère.

3.3.1 Restrictions supplémentaires sur la base de données

- **Intervalles de population** : Nous avons montré ci-dessus qu’une grande majorité de wikis étaient à la fois faiblement peuplés (i.e. souvent moins de 10 membres) et/ou n’ont pas un nombre significatif d’articles (i.e. généralement moins de 10). Afin de travailler sur un ensemble de wikis relativement homogènes et d’éviter de fonder notre analyse sur des données s’étalant sur de trop nombreux ordres de grandeur, nous avons exclu les wikis dont la population était soit très grande soit très faible : nous considérons finalement les wikis dont la population est incluse dans l’intervalle $[400; 20000]$ au premier jour de la collecte (rappelons que dans cette région, les tailles de contenu sont très étalées, d’une dizaine à plusieurs centaines de milliers d’articles).
- **Discontinuités** : Certains wikis font l’expérience de modifications brutales dans leur population ou leur contenu, parfois d’un facteur dix environ. Plusieurs raisons peuvent être invoquées pour

expliquer ces phénomènes, comme une attaque de *spam* ou bien la décision de la part des administrateurs de transférer, créer, fusionner ou supprimer un grand ensemble de pages d’un coup, ou bien encore d’accueillir ou de bannir brusquement certains types d’utilisateurs.¹⁹ Afin d’exclure les wikis exhibant ce type de comportement “suspect”, suivant l’hypothèse que le recrutement soudain de membres ou la création massive de pages ne peut pas se produire “naturellement” sur une période aussi courte que 24 heures, nous définissons un seuil sur le pourcentage de variation de la population comme du contenu entre deux jours successifs, que nous fixons à 0.05. En d’autres termes, nous excluons les wikis qui subissent une augmentation ou une réduction du nombre de membres ou d’articles de plus de 5% sur 24 heures.

Pour résumer, la base de données “propre” utilisée pour l’étude dynamique est finalement constituée d’environ 360 wikis qui ont tous une population initiale entre 400 et 20000 utilisateurs, qui ne sont *pas* hébergés sur les “fermes à wiki” renvoyant des données inexactes, et qui ne présentent pas de discontinuité majeure dans la variation quotidienne de la population et/ou du contenu.

Variable	Quantile				
	1	2	3	4	5
activité des utilisateurs (<i>edits per user</i>)	[0.14, 3.67[[3.67, 9.80[[9.80, 24.8[[24.8, 61.2[[61.2, 1903]
ratio d’administration (<i>admins per user</i>)	[.00169, .00192[[.00192, .00347[[.00347, .00576[[.00576, .01[[.01, 1.45]
densité d’édition (<i>edits per page</i>)	[0.12, 9.2[[9.2, 14.3[[14.3, 21.8[[21.8, 35.1[[35.1, 47245]
densité d’administrateurs (<i>admins per page</i>)	[$4.16 \cdot 10^{-6}$, .00103[[.00103, .00309[[.00309, .00926[[.00926, .0299[[.0299, 2.65]
densité d’utilisateurs (<i>users per page</i>)	[.00119, 0.22[[.220, .728[[.728, 2.07[[2.07, 7.05[[7.05, 363]
cas particulier du diagramme de phase Fig. 5: {0.00119497, .116, .307, .619, 1.24, 2.44, 4.99, 12.1, 363}					

quantiles de population	1	2	3	4	
	quantiles moyens:	[400, 615[[615, 1075[[1075, 2407[[2407, 19909]
	quantiles moyens sur le diagramme de phase de la Fig. 5: {400, 504, 676, 949, 1313, 2162, 3787, 19909}				

TAB. 1 – Étendue des quantiles

3.3.2 Évaluation de la dynamique

Nous appréhendons la dynamique des wikis en comparant leurs divers modes de croissance avec un ensemble de variables indépendantes. La ‘croissance’ est d’abord définie en termes de variations dans la population et la taille du contenu : la croissance des utilisateurs G_U (resp. la croissance des pages G_P) est le ratio entre les populations finale et initiale (resp. les tailles du contenu finale et initiale) : $G_U = U_{\text{last}}/U_{\text{first}}$ (resp. $G_P = P_{\text{last}}/P_{\text{first}}$). La dynamique de croissance est ensuite étudiée en fonction de variables indépendantes construites à partir des variables disponibles présentées dans la Section 3.1.2. Il s’agit :

- (1) d’INDICATEURS DESCRIPTIFS, i.e. des variables sur lesquelles les administrateurs n’ont pas de contrôle direct : **(a)** l’*activité des utilisateurs*, i.e. la proportion du nombre d’éditions par utilisateur (E/U , “*edits per user*”), **(b)** la *densité d’utilisateurs*, i.e. la proportion du nombre d’utilisateurs par page (U/P , “*users per page*”), et **(c)** la *densité d’éditions*, i.e. la proportion des éditions par page (E/P , “*edits per page*”).

¹⁹Certes, dans le cas des attaques de *spam*, par exemple, la viabilité du wiki est fortement menacée. Toutefois, sans méthode a priori pour distinguer systématiquement entre les différents cas possibles, nous préférons ignorer dans notre analyse cette portion ambiguë de la base de données incriminée.

(II) de FACTEURS DE GOUVERNANCE, soit des variables que les administrateurs peuvent directement contrôler : **(a)** le *ratio d'administration*, i.e. la proportion d'utilisateurs qui ont le statut d'administrateur (A/U , “*admins per users*”), **(b)** la *densité d'administrateurs*, i.e. la proportion d'administrateurs par page (A/P , “*admins per page*”), et **(c)** l'*autorisation d'édition* (R).

Pour chaque variable, nous avons adopté une approche plus compacte et informative qu'une analyse des nuages de points correspondants, en regroupant les wikis en cinq quantiles, chacun comprenant exactement 20% de tous les wikis de la base de données nettoyée ; la Table 1 décrit l'étendue des quantiles ainsi obtenus. Nous avons ainsi calculé et comparé la moyenne des taux de croissance des wikis d'un même quantile, pour chaque quantile. Nous avons en outre systématiquement distingué les quantiles de nombre d'utilisateurs afin de contrôler parallèlement les effets liés aux tailles de population, en utilisant ici quatre quantiles²⁰. Au final, nous avons tracé un *paysage de croissance* par variable, consistant en une représentation bi-dimensionnelle des divers taux de croissance en fonction, d'une part, des cinq quantiles de la variable indépendante observée et, d'autre part, des quatre quantiles de population. Ce processus a été appliqué à toutes les variables mentionnées ci-dessous, sauf pour R où seulement deux “quantiles” peuvent être distingués (0 ou 1).

Dans les figures qui suivent, pour chacune des variables sauf R , le graphe supérieur indique les valeurs moyennes et les intervalles de confiance ($p < 0.05$) des taux de croissance pour chaque quantile de la variable en question, tandis que le graphe inférieur montre un paysage de croissance sous forme de courbes de niveau : les zones les plus claires correspondent à des zones de plus forte croissance.

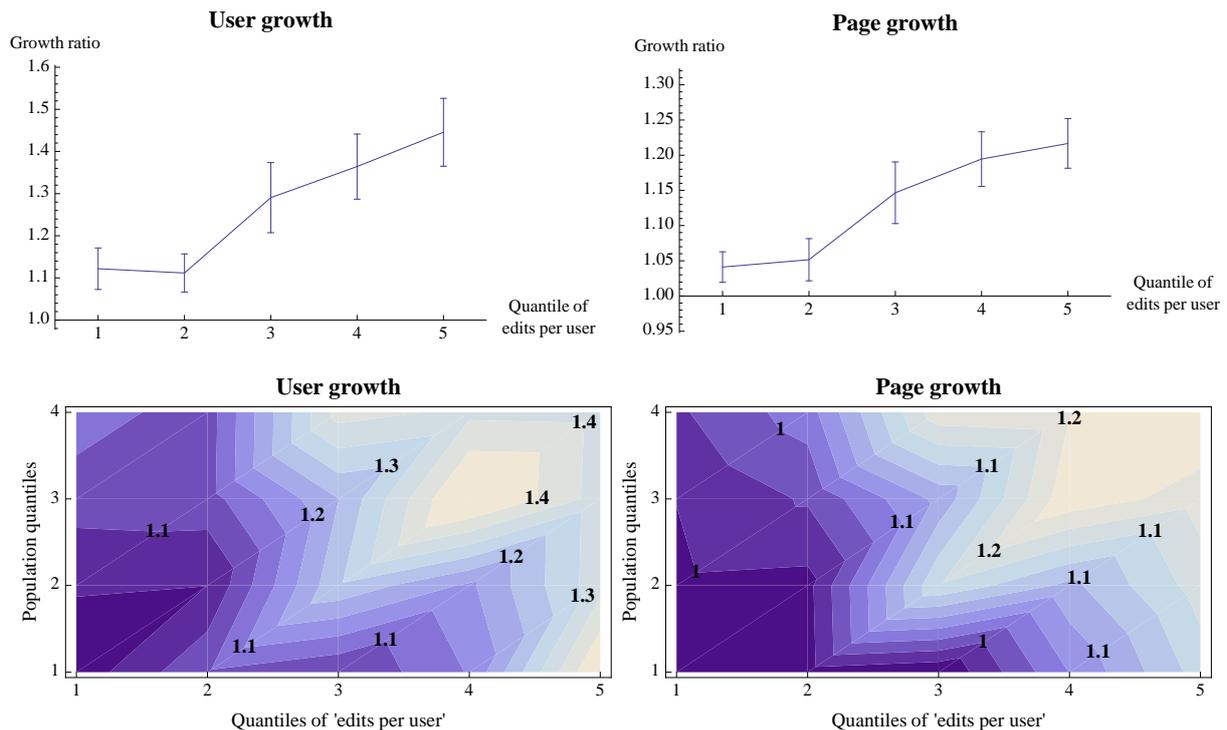


FIG. 4 – Paysage de croissance en fonction de l'activité des utilisateurs, c'est-à-dire la proportion des éditions par utilisateur (E/U) ou “*edits per user*”. Sur le tracé bi-dimensionnel des courbes de niveau (graphes inférieurs), les régions claires indiquent une croissance plus rapide; les valeurs indiquent la croissance moyenne dans les régions correspondantes (“1.4” indique par exemple à une augmentation de 40% sur les 8 mois d'observation).

²⁰À l'exception du diagramme de phase de la Fig. 5 qui utilise huit quantiles.

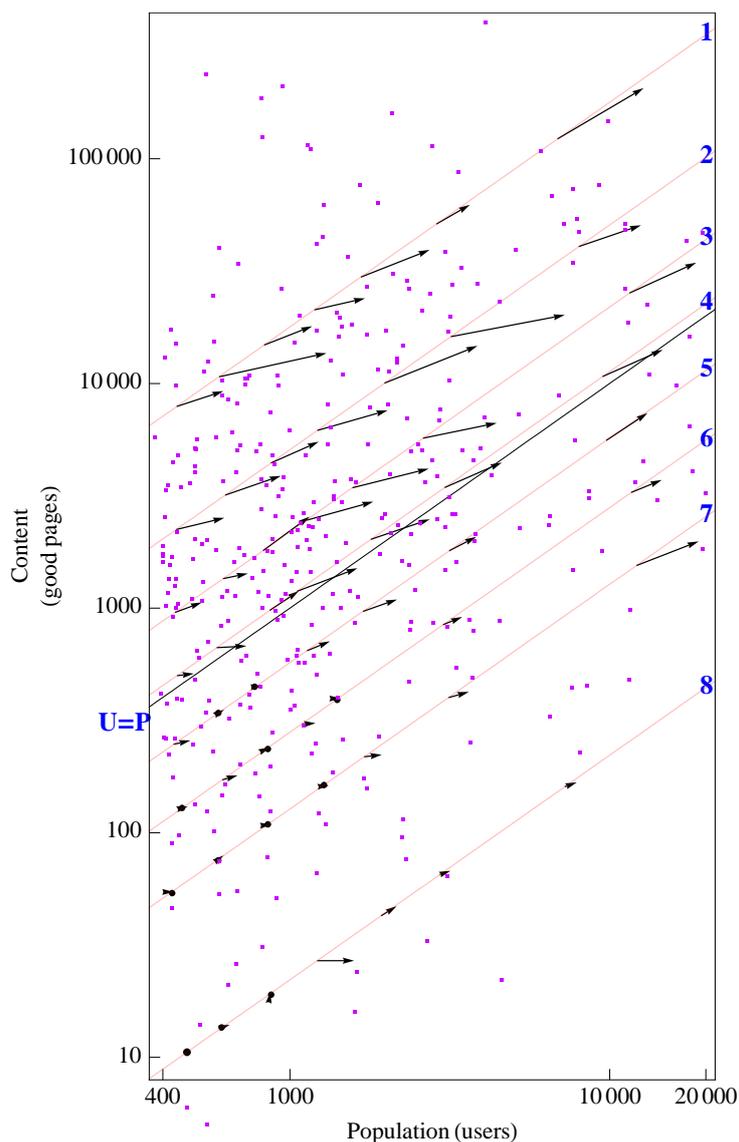


FIG. 5 – “Diagramme de phase” dans l’espace “contenu/population” pour les wikis appartenant à la base de données nettoyée.

Indicateurs descriptifs significatifs. On observe premièrement des corrélations significatives entre certains indicateurs descriptifs de la structure des wikis et les taux de croissance du contenu et de la population. La figure 4 montre ainsi l’effet sur les taux de croissance de l’ACTIVITÉ DES UTILISATEURS (mesurée comme le nombre d’*éditions par utilisateur*). Les résultats suggèrent que cette activité est très fortement corrélée avec la croissance, non seulement en termes de production de contenu (ce qui, dans une certaine mesure, n’est pas surprenant) mais aussi en termes de recrutement de nouveaux membres. Cet effet est d’autant plus fort chez les wikis les plus peuplés : plus il y a de membres qui éditent activement, plus le wiki croît en contenu *et* en population.

Plus généralement, le graphe de la figure 5 doit être vu comme une *carte* d’une portion de la wikisphère qui permet de visualiser aisément le destin typique d’un wiki en fonction précisément de sa position initiale dans le même espace. Sur ce diagramme, traditionnellement appelé “diagramme de phase”, une flèche est attachée à chaque région de la wikisphère définie par un couple de quantiles “densité d’utilisateurs, population”. La taille de la flèche représente la force de la croissance moyenne observée pour les wikis de la région correspondante. Plus précisément, les longueurs et hauteurs de ces flèches

sont respectivement proportionnelles aux taux de croissance des nombres d'utilisateurs et d'articles. Par ailleurs, chaque point (couleur claire) correspond à un wiki de la base de données.²¹

En bref, ce diagramme suggère finalement que la position d'un wiki est corrélée avec son profil de croissance. Plus précisément, il illustre bien le fait que les wikis situés dans la portion supérieure/supérieure-droite du diagramme croissent plus rapidement et, de manière plus pragmatique, il fournit un aperçu de la dynamique démographique dans cette région de la wikisphère.

Facteurs de gouvernance significatifs. Examinons à présent la corrélation entre les modes de gouvernance et la dynamique des wikis : nous analysons dans un premier temps l'impact de la DENSITÉ D'ADMINISTRATEURS (nombre d'administrateurs rapporté à la taille du contenu). La figure 6 montre ainsi que la présence d'une proportion élevée d'administrateurs par page tend à réduire fortement la croissance à la fois du contenu et du nombre de membres. Par exemple, alors que le dernier quantile d'admins/page a des taux de croissance quasi-nuls sur les 8 mois, le premier quantile exhibe des taux globalement élevés ($\sim +50\%$ pour les membres, $\sim +25\%$ pour les pages). Cet effet, qui est robuste quelle que soit la population initiale, peut s'interpréter comme l'impact négatif d'une forte activité de gouvernance sur la prolifération des articles et des utilisateurs.

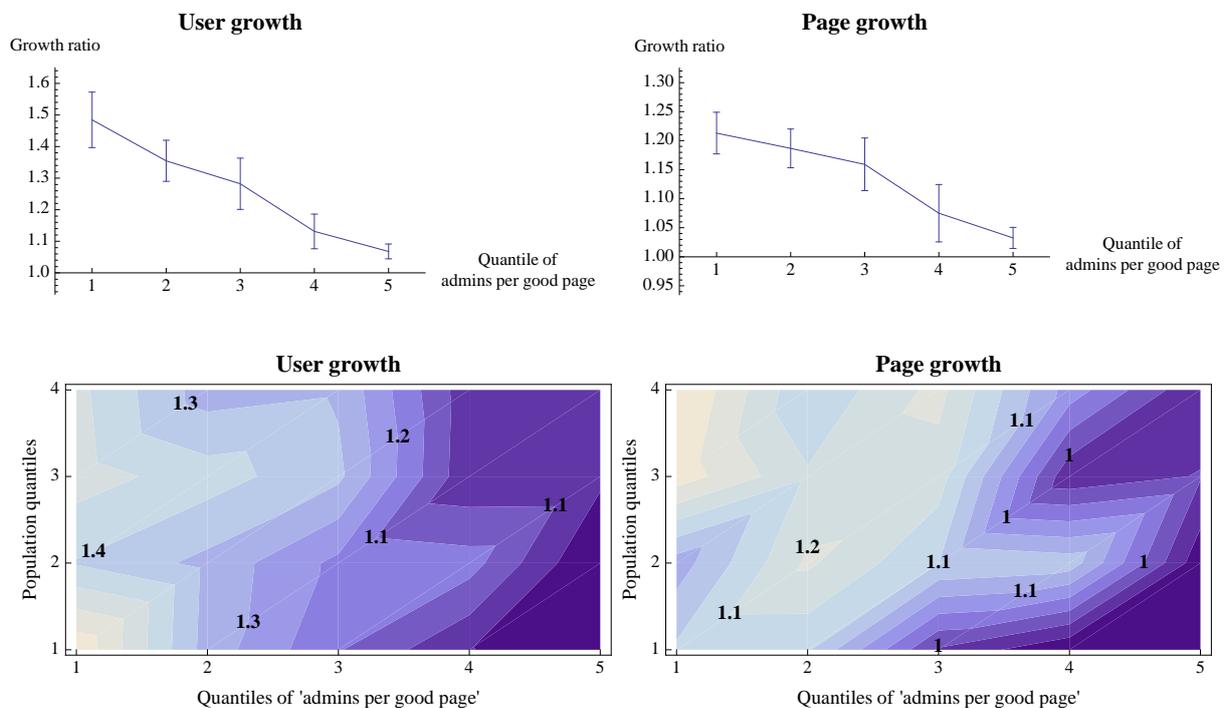


FIG. 6 – Paysage de croissance en fonction de la proportion d'*administrateurs par page* (A/P).

Nous avons pu ensuite observer que l'AUTORISATION D'ÉDITION avait un effet tout aussi sensible. La notion de quantile est caduque pour ce type de variable binaire et nous ne considérons ainsi que deux groupes de wikis : ceux qui permettent la modification du contenu par les visiteurs anonymes et ceux qui l'interdisent. Le paysage de croissance est de fait limité à une comparaison uni-dimensionnelle sur

²¹Techniquement, nous avons représenté sur le même graphe (i) les variables dépendantes G_U et G_P sous forme de flèches, (ii) les quantiles des variables indépendantes U et U/P respectivement en abscisse et sous forme de diagonales, et (iii) les positions initiales des wikis U_{first} and P_{first} sous forme de nuages de points. Ce diagramme nécessitant un plus haut niveau de détail que les figures précédentes, nous avons augmenté la précision de la subdivision en quantiles : il y a ainsi 8×7 quantiles ; les moyennes des quantiles U/P sont ainsi représentées par des diagonales libellées "1-8".

les quantiles de population. Les résultats exposés sur la figure 7 montrent clairement que l'absence de contrôle d'accès est un facteur favorisant significativement la croissance, aussi bien en contenu qu'en population. Ainsi, si la plus forte prolifération du contenu sur les wikis sans contrôle d'accès n'est pas particulièrement surprenant, le fait que le nombre de membres y augmente plus rapidement est bien plus contre-intuitif : l'on pourrait en effet s'attendre à ce que l'absence d'obligation de s'inscrire avant de participer n'incite pas à créer un compte, toutes choses égales par ailleurs. Or, au contraire, l'on voit ici que l'absence de contrôle d'accès favorise davantage la participation et le recrutement de nouveaux membres.

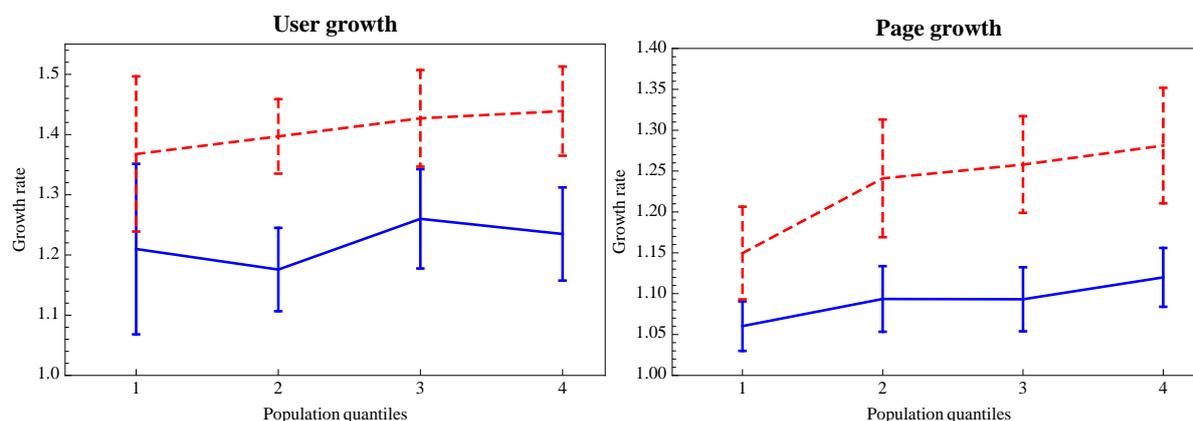


FIG. 7 – Paysage de croissance en fonction de l'autorisation d'édition (R): le trait *pointillé rouge* renvoie aux wikis modifiables anonymement, le trait *continu bleu* correspond aux wikis modifiables par les membres seulement.

Indicateurs neutres. Deux autres indicateurs, enfin, affichent une corrélation beaucoup plus faible avec la dynamique des wikis. D'une part, la DENSITÉ D'ÉDITIONS ("*edits per page*") est seulement corrélée très modérément négativement avec la croissance du nombre de membres—de manière sensiblement plus marquée suivant la taille de la population initiale—alors qu'il n'y étonnamment aucune corrélation avec la croissance du contenu. D'autre part, un RATIO D'ADMINISTRATION plus élevé n'induit aucun impact significatif sur la démographie des wikis ; ce phénomène est à comparer avec l'effet de la densité d'administrateurs mis en évidence ci-dessus.

3.3.3 Récapitulatif des mesures

L'analyse quantitative présentée dans cette section confirme qu'il existe une corrélation sensible entre la destinée d'un wiki et certains facteurs structurels ou institutionnels. Sans extrapoler davantage, il est possible de suggérer que, dans certains cas, ces corrélations témoignent en réalité d'effets causaux sous-jacents. Le tableau 2 et la figure 8 résument le sens de ces corrélations en comparant le gain de population et de contenu observé entre le dernier et le premier quantile, pour chaque variable (sur la figure 8, les variables sont ordonnées de la variable la plus positivement corrélée à celle la plus négativement corrélée). Dans tous les cas mentionnés ci-dessus, nous avons constaté une corrélation frappante entre la croissance de la population et celle du contenu.

Si nous nous intéressons aux caractéristiques structurelles des wikis, nous pouvons noter que plus le nombre d'éditions par utilisateur est élevé, plus le wiki croît rapidement, à la fois en population et en nombre d'articles. Ainsi, les communautés wikis dont les utilisateurs sont très actifs ne sont pas seulement susceptibles de voir leur contenu s'étoffer, elles vont aussi avoir tendance à attirer un plus grand nombre de nouveaux contributeurs. Ce résultat contraste nettement avec le cas apparemment comparable d'une forte densité d'utilisateurs par page, où l'on observe en fait un effet inverse — plus il y a

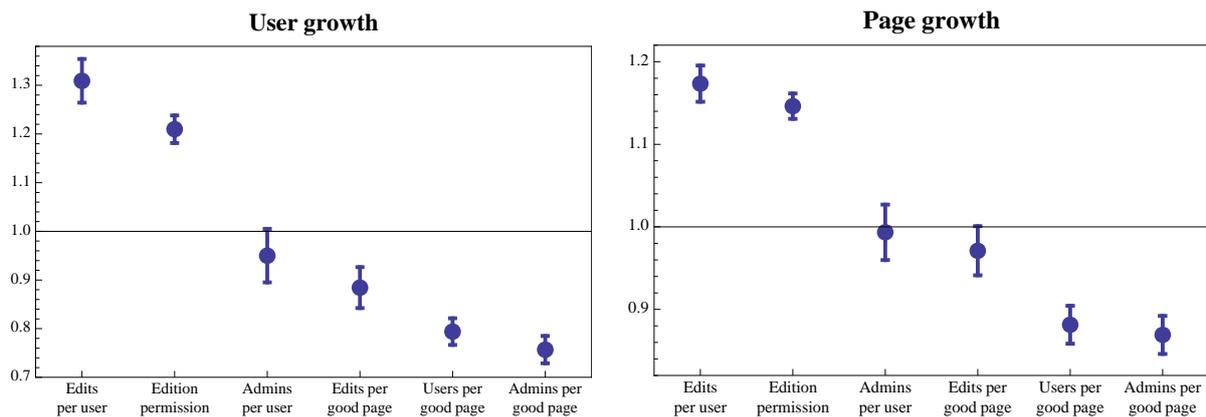


FIG. 8 – Comparaison des taux de croissance entre les derniers et premiers quantiles, pour chaque variable.

	Variable	Taux de croissance	
		Population	Contenu
INDICATEURS STRUCTURELS	Activité des utilisateurs (E/U)	++	++
	Densité d'édition (E/P)	-	—
	Densité d'utilisateurs (U/P)	--	--
FACTEURS DE GOUVERNANCE	Autorisation d'édition (R)	++	++
	Ratio d'administration (A/U)	—	—
	Densité d'administrateurs (A/P)	--	--

TAB. 2 – Effet des différents facteurs sur les taux de croissance des wikis.

d'utilisateurs par page, moins le wiki croît.

En ce qui concerne les modes de gouvernance, nous avons singulièrement constaté que le recrutement est environ 20% plus rapide lorsque les wikis sont éditables par n'importe qui, ce qui semble accréditer l'intuition suivant laquelle l'absence de barrières favorise *in fine* le développement de la communauté. Nous avons pu en outre observer qu'une trop forte densité d'administrateurs *par page* pouvait freiner la croissance d'un wiki, même si le ratio d'administrateurs *par membre* n'apparaît pas avoir d'influence sur la dynamique de croissance.

4 Organisation des communautés wikis et processus démographiques

Les sections précédentes ont permis de dresser un tableau quantitatif des dynamiques démographiques globales des communautés wikis. Cette analyse autorise à son tour la déduction de quelques prédictions simples mais significatives sur la manière dont ces wikis sont susceptibles d'évoluer. Comprendre les processus à l'origine de ces motifs de croissance requiert toutefois une analyse plus fine et plus profonde des liens entre dynamiques observées et comportement des membres de ces communautés en ligne : plus concrètement, comment lier nos observables démographiques aux mécanismes organisationnels à l'œuvre au sein des wikis ? La recherche sur les wikis est encore en ce sens à un stade préliminaire et de futures études empiriques pourraient ainsi viser à, d'une part, intégrer à ce type d'approche des résultats d'enquêtes qualitatives et, d'autre part, collecter des données quantitatives plus fines sur la dynamique microscopique des contenus et de la population, à l'échelle des agents, en vue de les lier aux dynamiques macroscopiques qui ont été mises en évidence ici. À cet égard, nous proposons dans cette section plusieurs axes de recherche potentiels.

4.1 Le lancement d'un wiki : principes fondateurs, normes et règles

Les conditions initiales relatives à la structure et au mode de fonctionnement d'un wiki figurent parmi les déterminants les plus importants de l'évolution de la communauté à ses débuts. Les créateurs d'un wiki en définissent généralement à la fois le champ d'application et les principes fondateurs, qui sont ainsi des ingrédients cruciaux pour attirer les tous premiers contributeurs extérieurs et réussir à développer une communauté de membres. Sanger [2005] remarque ainsi que la définition des principes fondateurs de la *Wikipedia* a été due à un processus essentiellement centralisé — seule l'application de ces règles a pu ensuite être décentralisée, par le biais des administrateurs. Les différentes règles régissant un wiki ont de fait divers degrés d'*autorité* : alors que certains principes fondateurs semblent non-négociables — ils constituent le cœur du wiki, voire sa raison d'être — d'autres règles sont plus permissives et sujettes à discussion. Par exemple, la règle de la *Wikipedia* suivant laquelle aucun contenu sous droit d'auteur ne doit être utilisé est tout à fait admise, tandis que la règle de neutralité de point de vue ("*NPOV*") est plus ouverte et souvent ré-interprétée. De manière générale toutefois, les principes peuvent souvent être discutés puis modifiés, et de nouveaux principes peuvent aussi être proposés à tout moment, comme par exemple ceux liés à la notion de "catégorie" dans la *Wikipedia* qui n'ont pas émergé avant 2004 [Viegas *et al.*, 2007, notamment la figure 5].

En outre, la volonté de faire varier les principes gouvernant un wiki existant peut elle-même être à l'origine de nouveaux wikis visant alors à expérimenter de nouveaux ensembles de règles fondatrices [Levrel, 2005]. La myriade de projets encyclopédiques inspirés par la *Wikipedia* illustre bien ce phénomène : par exemple, *Citizendium* a apporté l'obligation de relecture des articles avant publication, en ôtant la possibilité d'éditer immédiatement les pages accessibles au public ; *Wikinfo*, différemment, a introduit la notion de "point de vue sympathique",²² ou "*SPOV*", qui encourage la multiplicité de points de vue plutôt que le consensus, en remplacement de la règle "*NPOV*" de la *Wikipedia*.

In fine, tous les principes n'obéissent pas aux mêmes échelles de temps : tandis que certaines règles sont discutées souvent, voire sont modifiables par de nouveaux membres, les principes fondateurs semblent beaucoup plus stables. La *Wikipedia*, à nouveau, distingue les "*policies*" des "*guidelines*"²³ ; en ce sens, différentes règles peuvent avoir des statuts légaux distincts ; les principes centraux jouant le rôle d'une constitution [Levrel, 2005]. Toutefois, chaque méta-article concernant les principes de la *Wikipedia* comporte la bannière suivante :

"Any edit to this page should reflect consensus. If in doubt, consider discussing changes on the talk page." (Wikipedia anglophone)

"Bien qu'il soit possible de modifier cette page, il est vivement conseillé d'utiliser la page de discussion ou le Bistro avant toute intervention." (Wikipedia francophone)

ce qui suggère effectivement que tout principe est amendable, jusqu'au plus fondamental, comme *NPOV*. Bien qu'ils semblent *a priori* irrévocables et/ou contrôlés essentiellement par un petit groupe de membres historiques, il est ainsi plausible que ces principes fondateurs évoluent simplement à long terme, lentement.

4.2 Dynamiques de population, dynamiques de peuplement

Motivations des utilisateurs. Nous l'avons vu, les principes fondateurs définissent le contexte initial permettant de motiver l'adhésion d'internautes à un wiki. Les utilisateurs d'un wiki obéissent à des motivations qui sont communes, plus largement, aux communautés en ligne en général. De précédentes études ont ainsi proposé diverses typologies des stratégies permettant d'inciter des individus à prendre part à ce type de communautés [Gallant *et al.*, 2007], qu'il s'agisse en particulier et entre autres d'études

²²Consultable à l'adresse http://www.wikinfo.org/index.php/Wikinfo:Sympathetic_point_of_view

²³Consultable à l'adresse http://en.wikipedia.org/wiki/Wikipedia:Policies_and_guidelines

portant sur la *Wikipedia* [Bryant *et al.*, 2005; Ciffolilli, 2003], sur les *tagging systems* (e.g. Flickr, CiteYouLike, etc.) [Marlow *et al.*, 2006] ou bien encore sur les groupes de développement de logiciels libres [Mockus *et al.*, 2002]. Dans l'ensemble, ces travaux fournissent un guide potentiellement applicable à l'analyse du comportement des membres d'un wiki : ceux-ci peuvent tout autant être motivés, par exemple, par l'idée de créer leur propre identité en ligne, de participer à l'élaboration d'un objet public, d'améliorer leurs propres compétences, d'afficher leur créativité, etc. En toute généralité, il semble possible de distinguer trois types de motivations :

- une motivation *altruiste*, correspondant à la contribution bénévole à un bien public (garanti par le type de licence en vigueur sur les wikis) ;
- une motivation *socialement intéressée*, comme la possibilité d'influencer d'autres membres ou le lectorat d'un wiki au sens large ; de partager et d'interagir avec d'autres individus, en développant le sentiment d'appartenance à une communauté ; ou bien encore de construire une identité virtuelle et de gravir une hiérarchie sociale en ligne ;
- une motivation *égoïste*, visant essentiellement à contribuer pour améliorer ses propres compétences, ou bien à utiliser la communauté strictement comme une ressource d'informations.

Chacun de ces types d'objectif individuel peut être diversement utilisé et mis en avant afin d'attirer certaines catégories d'utilisateurs, tous les utilisateurs n'étant pas identiquement sensibles aux mêmes types d'incitations. En tout état de cause, les multiples combinaisons possibles des règles de participation, d'édition et des configurations techniques définissent un paysage d'incitations précis, à même de déterminer les conditions de recrutement des nouveaux membres et de conservation des anciens, et donc *a fortiori* le profil de viabilité de chaque wiki.

Enrôlement et leadership. Bryant *et al.* [2005] ont mené une des toutes premières études ethnographiques d'une communauté wiki en relatant précisément le processus d'implication de nouveaux membres et la distribution du travail sur la *Wikipedia*, et en décrivant notamment les différentes phases d'acclimatation d'un individu au fonctionnement du wiki : de la découverte initiale du wiki et de ses règles, en passant par les premières contributions, comme la correction de fautes de frappe ou la création d'un compte, jusqu'à jouer un rôle plus prégnant et, parfois, devenir soi-même administrateur du site. Ces auteurs montrent que ce processus est ainsi comparable à la manière dont un apprenti intègre progressivement sa communauté de pratique, en apprenant progressivement à se servir de davantage d'outils, tout en bénéficiant temporairement de la clémence des artisans plus expérimentés.

En ce qui concerne la distribution des rôles entre membres, la hiérarchie au sein d'un wiki se réduit fréquemment à une dichotomie entre contributeurs produisant du contenu et administrateurs appliquant les décisions de la communauté. Implicitement cependant, cette institution rudimentaire ne se limite pas à donner aux administrateurs un simple pouvoir légal sur les utilisateurs plus novices, mais s'interprète plus largement en termes de "centre-périphérie" [Borgatti & Everett, 2000]. Kittur *et al.* [2006] exhibent ainsi des différences quantitatives entre les administrateurs et les autres utilisateurs de la *Wikipedia* : les premiers sont souvent les contributeurs les plus actifs ; ils doivent informellement leur statut à une implication remarquable. De plus, leur travail d'édition semble même être essentiellement différent : les membres normaux effectuent ainsi moins de modifications que les administrateurs, ils effacent davantage de mots qu'ils n'en ajoutent, ce qui laisse penser que les administrateurs sont principalement à l'origine de la croissance des wikis.

Ainsi, les administrateurs ne se distinguent pas seulement par leur pouvoir formel : ils construisent aussi leur autorité informelle à travers des efforts réguliers et généralement humbles [Reagle, 2007], comme prendre le temps d'expliquer et de transmettre les principes aux débutants [Bryant *et al.*, 2005]. En retour, ces membres respectés ont tendance à avoir une influence particulière lors des crises. Malgré une égalité affichée entre membres, au sein de la *Wikipedia* et très probablement au sein des wikis en général comme dans la plupart des autres communautés virtuelles [Mockus *et al.*, 2002], l'autorité implicite est un facteur de gouvernance important : en dépit d'un processus de décision participatif, les

administrateurs bénéficient d'une marge d'action particulière pour orienter significativement la dynamique du wiki.²⁴

Croissance et déclin. La section 3 a par ailleurs montré que la population des wikis était très légèrement corrélée avec la taille du contenu. La croissance exponentielle de la *Wikipedia* en termes d'articles et d'utilisateurs suggère cependant que les wikis "viabiles" pourraient suivre un chemin démographique vertueux, où contenu et contributeurs co-évoluent de manière à renforcer conjointement la croissance de l'un et de l'autre. Cette analyse est en particulier étayée par la carte globale de la Fig. 5 où les flèches décrivent effectivement une dynamique duale.

De manière plus détaillée toutefois, la dichotomie entre les "simples" contributeurs et les administrateurs davantage impliqués semblent induire des dynamiques de population distinctes, au moins sur la *Wikipedia* [Kittur *et al.*, 2006] : après une première période où les membres très actifs sont aussi ceux qui contribuent le plus en nombre d'éditions, et voient ainsi leur influence croître, une deuxième période fait apparaître une activité globale plus importante de la part des petits contributeurs. Plus simplement, ceci signifie que si, au début, il existe un petit groupe d'individus qui s'implique beaucoup dans la création et l'amélioration d'un embryon de contenu afin d'établir les fondations du wiki et d'assurer son développement initial, le wiki ensuite évolue vers une phase de recrutement de nouveaux utilisateurs, probablement moins impliqués en moyenne mais beaucoup plus nombreux. Même s'il ne semble pas y avoir de transition tranchée entre ces deux modes, l'on peut ainsi distinguer la période de gestation de la période "de croisière". À cet égard, il est raisonnable de conjecturer que la période de gestation correspond à la grande zone dense observée en bas à gauche du nuage de points de la Fig. 2, zone dont les wikis peuvent finir par émerger, en cas de succès, et suivre ensuite une dynamique démographique plus régulière, co-évolutionnaire, le long de la diagonale (utilisateurs, contenus).

Les facteurs qui poussent les individus à quitter une communauté wiki ont été peu explorés. Comme le souligne Levrel [2005], retenir des contributeurs dans un projet ouvert est un défi en soi, notamment lorsque le projet croît et requiert de plus en plus de membres impliqués et prêts à prendre des responsabilités. Dans la *Wikipedia* francophone par exemple, certains administrateurs ont exigé à un moment que leur nombre soit sensiblement accru afin de pouvoir poursuivre correctement leur tâche, sous la menace de quitter le projet : ici, la croissance du contenu a constitué un danger pour la stabilité du wiki. Un autre risque sérieux a trait aux embranchements, ou *forks* en anglais, c'est-à-dire l'émergence de wikis concurrents à partir d'un projet existant, dont le contenu est réutilisé comme base de départ — ce qui est précisément rendu possible par les licences ouvertes en vigueur sur les wikis. Ce type d'évènement a fréquemment pour conséquence de diviser les communautés et de diminuer significativement la force de travail disponible pour chacun des projets, en augmentant par ailleurs les chances de confusion entre les différents wikis. De façon plus optimiste, les licences ouvertes rendent aussi toute tentative de "coup d'état" caduque, puisque n'importe qui peut créer un embranchement dès qu'un projet semble monopolisé par un petit groupe. Par exemple, *Enciclopedia Libre*²⁵ est un embranchement de la *Wikipedia* hispanophone et a été créé lorsqu'on a filtré les premières rumeurs d'utilisation de la publicité pour couvrir les frais de fonctionnement : la perspective que le wiki dévie du projet de partage gratuit du savoir ayant suffi à supprimer une source de motivation centrale chez les membres sécessionnistes.

4.3 Dynamique du contenu

Certains facteurs sous-tendant la croissance de la population sont directement liés aux caractéristiques du contenu. Des articles de qualité, témoignant de la réussite du projet collaboratif, sont évidemment susceptibles d'avoir un effet positif sur le recrutement de nouveaux membres. La croissance du contenu,

²⁴Un des fondateurs de la *Wikipedia*, Jimmy Wales, affirme volontiers vouloir jouer le rôle d'"un monarque britannique" [McNicho, 2007; Reagle, 2007].

²⁵Initiée en 2002 et accessible à <http://enciclopedia.us.es>

néanmoins, a un effet plus ambigu, en contribuant à la fois à la complétude du projet (et donc à sa qualité) mais aussi en élevant le nombre de membres actifs nécessaires à son entretien, ce qui n'est pas sans danger comme nous l'avons vu ci-dessus. En particulier, les wikis dans la région supérieure-gauche des figures 2 & 5 sont ceux qui prolifèrent le plus et qui ont simultanément le plus petit nombre d'utilisateurs par rapport au contenu ; en tant que tels, ils courent un risque plus élevé d'abandon ou de vandalisme que les autres wikis. Étant données ces contraintes, quelles sont les meilleures stratégies pour garantir la viabilité ?

Stabilisation du contenu. Premièrement, la possibilité de revenir immédiatement aux versions précédentes inhibe toute velléité de vandalisme de masse, et il est ainsi quasiment impossible de détruire le contenu d'un wiki et d'assister à ce que Ciffolilli [2003] appelle un "*content disaster*". Autrement plus intéressantes sont les menaces régulières à l'intégrité du contenu, comme les contributions inexactes, le vandalisme occasionnel, le non-respect plus ou moins ostensible des règles, ou bien encore les "guerres d'édition", c'est-à-dire un désaccord persistant entre plusieurs éditeurs au sujet d'une idée ou d'une formulation. Tandis que le vandalisme le plus grossier ne nécessite pas davantage qu'une équipe de taille appropriée pour surveiller le contenu d'un wiki²⁶, la falsification de faits ou bien les disputes entre contributeurs requièrent des règles et des procédures d'arbitrage plus subtiles.

Les mécanismes de résolution de ce type de conflits sont en général bien développés — comme, par exemple, la "règle des trois révocations" ("*three-revert rule*") sur la *Wikipedia*,²⁷ qui permet de lutter contre certaines formes discrètes de vandalisme. Lorsqu'elles existent, les pages de discussion jouent ici un rôle central pour faire respecter les règles d'édition [Viegas *et al.*, 2007], et en particulier lorsqu'il s'agit de définir collectivement et consensuellement les procédures de résolution de conflits elles-mêmes : conflits et règles co-évoluent alors essentiellement au travers de ces pages. Viegas *et al.* [2007] soulignent plus largement que les wikis se construisent en tant qu'organisations *conventionnelles* de manière ascendante à travers les pages de discussion, plutôt que de manière descendante du fait de l'imposition de normes par un petit groupe d'agents. Dans l'ensemble, il s'agit là d'un instrument crucial à disposition des membres pour restaurer la viabilité d'un wiki, soit à un niveau local, par exemple en activant la protection d'une page pour éviter des modifications indésirables, soit à un niveau plus global en définissant de nouvelles règles²⁸. Soulignons enfin le rôle des relations sociales "hors-ligne" (c'est-à-dire hors-wiki : canaux IRC, listes de diffusion, messagerie instantanée, etc.) qui permettent aussi une entraide immédiate pour les conflits plus sérieux [Kittur *et al.*, 2006].

Il est probable que les wikis n'ayant pas de structure de gouvernance (i.e., des administrateurs) puissent avoir davantage de mal à gérer les conflits et, *in fine*, à rester viables. En ce sens, la croissance plus faible des wikis avec les plus fortes densités d'administrateurs, illustrée par la figure 6, peut être interprétée comme un effet de régulation. L'existence de tous ces mécanismes de stabilisation suggère en tout cas que la question de la viabilité se pose à diverses échelles — des interactions les plus locales (discussions, négociations entre membres) aux contraintes les plus globales (principes fondateurs, décisions mises en œuvre par un petit groupe d'"experts") — et de manière asynchrone : l'évolution des normes a effectivement un impact de second ordre sur la dynamique du contenu.

Évaluation de la qualité. La qualité du contenu, enfin, semble être un facteur déterminant de la dynamique démographique en ce qu'elle permet d'attirer de nouveaux utilisateurs [Bryant *et al.*, 2005] et de

²⁶Sur la *Wikipedia* par exemple la "Recent Changes patrol" est un groupe auto-organisé qui vérifie si les pages récemment modifiées n'ont pas été vandalisées.

²⁷Un utilisateur annulant successivement trois modifications d'un article par d'autres utilisateurs se voit interdire l'édition de ce même article pour une période donnée.

²⁸Par exemple, afin d'éviter que des contributeurs motivés ayant un conflit occasionnel au sujet d'un article soient trop violemment malmenés et risquent de quitter la communauté, les wikipédiens ont décidé consensuellement de ne plus parler que de "problèmes entre deux personnes" plutôt que de "problèmes avec un éditeur" lors d'un conflit d'édition, sans se focaliser sur un individu en particulier [J. Levrel, communication personnelle].

stabiliser le contenu lui-même. Sur la *Wikipedia*, certaines règles d'édition comportent ainsi de véritables prescriptions concernant la qualité des articles : par exemple, [Stvilia et al. \[2008\]](#) montrent que la notion de “*featured article*”²⁹ fixe un objectif de qualité, tandis que la notion d’“*articles for deletion*” constitue une menace permanente qui fixe un seuil minimum de qualité à respecter.

Cependant, sauf dans les cas de vandalisme strict (suppression, falsification, non-respect évident des règles), la qualité d'un article reste difficile à évaluer avec des critères objectifs. [Lih \[2004\]](#) a établi une liste de critères corrélant la qualité d'une page au nombre d'éditions et d'éditeurs, tandis que [Wilkinson & Huberman \[2007\]](#) ont récemment corrélié l'autorité (“*PageRank*”) d'une page dans Google à l'activité d'édition la concernant ; entre temps, une comparaison manuelle d'articles de la *Wikipedia* avec leur équivalent de l'*Encyclopedia Britannica* n'a révélé aucune différence significative dans la fiabilité du contenu entre les deux encyclopédies [[Giles, 2005](#)]. Plus récemment, des méthodes essentiellement fondées sur la comparaison entre versions d'un même article ont été développées pour évaluer la fiabilité du contenu [[Adler et al., 2008](#); [Kramer et al., 2008](#)]. En ce qui concerne la qualité des contributeurs, [Anthony et al. \[2005\]](#) ont mesuré les divers temps de survie des éditions par les contributeurs de la *Wikipedia* et ont ainsi identifié deux catégories d'acteurs de qualité : les membres qui effectuent beaucoup d'éditions (des “experts”) et des visiteurs anonymes qui font peu d'édition (des “passants”, de plus en plus nombreux [[Kittur et al., 2006](#)]). Ils ont en outre pu montrer que l'importance des passants est proprement contre-intuitive [notamment par rapport à [Bryant et al., 2005](#)], puisque ceux-ci semblent produire du contenu de meilleure qualité que les membres enregistrés.

5 Conclusion

La contribution principale de cette étude tient dans l'analyse démographique, à la fois statique et dynamique, d'un échantillon relativement grand de communautés en ligne, basées sur une plate-forme wiki. À cet égard, nous avons souligné l'imbrication entre croissance de la population et croissance du contenu ; ceci suggère que les modèles de dynamique des communautés wikis devront certainement prendre en compte les corrélations fortes existant entre ces deux variables. Plus largement, nous nous sommes finalement efforcés de rapprocher des variables quantitatives simples à des caractéristiques qualitatives liés à des propriétés organisationnelles, comme la distribution des rôles ou les modes de gouvernance. Bien que s'appuyant sur un ensemble de wikis particulier, ce travail constitue la première contribution empirique à l'étude de la viabilité des communautés wikis, au-delà de l'exemple canonique de la *Wikipedia*. De manière plus pragmatique, ce travail donne un aperçu des indicateurs que les membres d'un wiki peuvent avoir intérêt à prendre en compte pour orienter et planifier leur dynamique démographique.

Remerciements

Nous remercions très chaleureusement “Mutante” et *s23.org* pour nous avoir accordé la permission de collecter leurs statistiques sur les wikis “MediaWiki”. Nous remercions en outre Julien Levrel pour plusieurs discussions fructueuses. Cet article a bénéficié partiellement du soutien du projet PATRES financé par la Commission Européenne (NEST-043268). Des portions de ce travail ont été précédemment présentées aux conférences *WikiSym* en 2007 et en 2008.

Références

B. T. Adler, K. Chatterjee, L. de Alfaro, M. Faella, I. Pye, and V. Raman (2008). Assigning trust to wikipedia content. In : A. Aguiar and M. Bernstein (eds), *WikiSym 4rd Intl Symposium on wikis*.

²⁹Dans la version francophone le terme employé est en fait simplement “article de qualité”.

- D. Anthony, S. Smith, and T. Williamson (2005). *Explaining quality in internet collective goods : Zealots and good samaritans in the case of wikipedia.*
- S. Borgatti and M. Everett (2000). Models of core/periphery structures. *Social networks*, **21**(4), 375–395.
- U. Brandes and J. Lerner (2008). Visual analysis of controversy in user-generated encyclopedias star. *Information visualization*, **7**, 34–48.
- S. L. Bryant, A. Forte, and A. Bruckman (2005, Nov 6-9). Becoming wikipediaian : Transformation of participation in a collaborative online encyclopedia. *In : Group'05, Sanibel Island, FL, USA.*
- A. Capocci, V. D. P. Servedio, F. Colaiori, L. S. Buriol, D. Donato, S. Leonardi, and G. Caldarelli (2006). Preferential attachment in the growth of social networks : The internet encyclopedia wikipedia. *Physical Review E*, **74**(3), 036116.
- A. Cifollilli (2003). Phantom authority, self-selective recruitment and retention of members in virtual communities : The case of wikipedia. *First Monday*, **8**(12).
- L. Gallant, G. Boone, and A. Heap (2007). Five heuristics for designing and evaluating web-based communities. *First Monday*, **12**(3).
- J. Giles (2005). Internet encyclopaedias go head to head. *Nature*, **438**(7070), 900–901.
- M. W. Godfrey and Q. Tu (2000). Evolution in open source software : A case study. *Icsm*, **00**, 131.
- J. Golbeck (2007). The dynamics of web-based social networks : Membership, relationships, and change. *First monday*, **12**(11).
- H.-J. Happel and M. Treitz (2008, May). Proliferation in enterprise wikis. *In : Proceedings of the 8th International Conference on the Design of Cooperative Systems (COOP 08).*
- A. Kittur, E. Chi, B. A. Pendleton, B. Suh, and T. Mytkowicz (2006). Power of the few vs. wisdom of the crowd : Wikipedia and the rise of the bourgeoisie. *In : Sig chi submission ?*
- M. Kramer, A. Gregorowicz, and B. Iyer (2008). Wiki trust metrics based on phrasal analysis. *In : A. Aguiar and M. Bernstein (eds), WikiSym 4rd Intl Symposium on wikis.*
- R. Kumar, J. Novak, and A. Tomkins (2006). Structure and evolution of online social networks. *Pages 611–617 of : Kdd '06 : Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining.* New York, NY, USA : ACM Press.
- J. Levrel (2005). *Le développement des contenus libres sur Internet - Wikipedia.* Tech. rept. FT R&D TECH/SUSI.
- J. Levrel (2006). Wikipédia, un dispositif médiatique de publics participants. *Réseaux*, **24**(138), 185–218.
- A. Lih (2004, April 16–17). Wikipedia as participatory journalism : Reliable sources ? Metrics for evaluating collaborative media as a news resource. *In : 5th intl symp on online journalism.*
- C. Marlow, M. Naaman, D. Boyd, and M. Davis (2006). Ht06, tagging paper, taxonomy, flickr, academic article, to read. *Pages 31–40 of : Hypertext '06 : Proceedings of the seventeenth conference on hypertext and hypermedia.* New York, NY, USA : ACM Press.
- T. McNicho (2007). *Wikipedia founder hunts for gold.* Business 2.0, CNN.

- A. Mockus, R. Fielding, and J. Herbsleb (2002). Two case studies of open source software development : Apache and Mozilla. *Acm transactions on software engineering and methodology*, **11**(3), 309–346.
- J. M. Reagle (2007). Do as I do : leadership in the Wikipedia. *In : Proceedings of WikiSym '07, 3rd Intl Symposium on Wikis*. New York, NY, USA : ACM Press.
- L. Sanger (2005). *The Early History of Nupedia and Wikipedia : A Memoir*.
<http://features.slashdot.org/article.pl?sid=05/04/18/164213>.
- B. Stvilia, M. Twidale, L. Smith, and L. Gasser (2008). Information quality work organization in wikipedia. *Journal of the American society for information science and technology*, **59**(6), 983–1001.
- F. Viegas, M. Wattenberg, J. Kriss, and F. van Ham (2007). Talk before you type : Coordination in Wikipedia. *In : Proc 40th hawaii intl conf on system sciences*.
- D. Wilkinson and B. Huberman (2007). Assessing the value of cooperation in Wikipedia. *First Monday*, **12**(4).
- V. Zlatic, M. Bozicevic, H. Stefancic, and M. Domazet (2006). Wikipedias : Collaborative web-based encyclopedias as complex networks. *Phys. Rev. E*, **74**(1), 016115.