

Viable Wikis

Struggle for Life in the Wikisphere

Camille Roth *

CRESS / Department of Sociology, University of Surrey, Guildford, GU2 7XH, UK
c.roth@surrey.ac.uk

Abstract

Wikis are collaborative platforms enabling collective elaboration of knowledge, the most famous and possibly the most successful thereof being the Wikipedia. There are currently plenty of other active open-access wikis, with varying success: some recruit many users and achieve sustainability, while others strive to attract sufficient active contributors, irrespective of the topic of the wiki. We make an exploratory investigation of some factors likely to account for these various destinies (such as distinct policies, norms, user incentives, technical and structural features), examining the demographics of a portion of the wikisphere. We underline the intertwining of population and content dynamics and emphasize the existence of different periods of development of a wiki-based community, from bootstrapping by founders with a pre-established set of rules, to more stable regimes where constant enrollment and training of new users balances out the occasional departure of more advanced users.

Categories and Subject Descriptors H.3.3, H.3.4 [*Information storage and retrieval*]: systems and software, online information services; K.4.3 [*Computers and society*]: collaborative work

General Terms Human Factors, Management, Reliability

Keywords Wikis, online communities, viability, wikisphere, collaborative work, user incentives, Wikipedia.

1. Introduction

Wikis are websites whose content can be elaborated by generally any user in a collective and collaborative fashion — the most famous and possibly the most successful of these platforms being Wikipedia, a multi-lingual encyclopedia which also attracted a substantial academic interest recently [1–

13].¹ There are currently, however, plenty of other active wikis, with varying success: some recruit many users, achieving sustainability with established role distributions, frequent updating and efficient fight against vandalism, while others strive to attract contributors. These many projects either have distinct policies or scope but are still sustainable, or have identical policies but die; all endeavoring to survive within what may be called the “wikisphere”. Examples include topical wikis, such as `ExampleProblems`, a repository of math example problems, or geographical wikis, such as `DavisWiki` which aims at creating Davis, CA-related content.

In this paper, we make an exploratory investigation of some factors likely to account for these various destinies, in terms of policies, norms, user incentives, as well as technical and structural features, so as to describe the pillars of the viability of wiki-based communities. Viability should be understood as dynamic stability of both population and quality content: in other words, a viable wiki should be able to survive large variations in the stock of users and articles so that the whole content can be maintained by a sufficient number of users. As our main interest is in online communities, we do not wish to consider wikis primarily supported by (and whose viability is thus linked to) pre-existing offline workgroups, where casual visitors would be unable to join as (full) contributors. We thus focus on communities emanating from the wiki platform, even if boundaries become relatively fuzzy when non-wiki groups open a wiki and let online visitors contribute (e.g., open-source developers of `iPodLinux` opening their documentation-writing process on a wiki) or when wiki groups partially transform into “real-life” communities (e.g., Wikipedians attending WikiMania conferences).

We first examine the current state of the wikisphere and (dis)similarities between wikis, supported by quantitative analysis of their demographics. In a second part, we qualitatively describe several processes which may account for their dynamics and survivability, and eventually outline some basic ingredients of a model of a viable wiki community.

2. Many wikis... many communities

In this section, we empirically investigate how homogeneous and heterogeneous the wikisphere may be, focusing on two sets of wiki communities: **(I)** for a qualitative insight on the

* I am grateful for enlightening talk with Julien Levrel, and discussions with Nigel Gilbert, Chris Goldspink, Bastien Guerry, Dario Taraborelli and Lu Yang. Research partly funded by EC project “PATRES”, contract #043268.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WikiSym’07, October 21–23, 2007, Montréal, Québec, Canada.
Copyright © 2007 ACM 978-1-59593-861-9/07/0010...\$5.00

¹ Online collaborative knowledge creation is not a brand-new phenomenon: discussion forums with experts tutoring novices, open-source software development groups are of the same sort. By contrast, sharp computer-related skills are not needed to participate in most wikis, whose ease of use have pulled in a larger range of users.

name	platform	license	language	scope	from	users	pages	governance	main website
Wikipedia	MediaWiki	GFDL	<i>multi</i>	all-purpose	2001	4,540k	9,047k	admins	wikipedia.org
EnciclopediaLibre	MediaWiki	GFDL	Spanish	all-purpose	2002	<2k	56k	admins	enciclopedia.us.es
Citizendium	MediaWiki	GFDL	<i>multi</i>	all-purpose	2006	31k	18k	RR, moderated	en.citizendium.org
Ekopedia	MediaWiki	L.A.L.	<i>multi</i>	topical	2004	<2k	6k	admins	ekopedia.org
WikiIndex.org	MediaWiki	CC	<i>multi</i>	topical	2005	<2k	12k	admins	wikiindex.org
iPodLinux	MediaWiki	GFDL	English	topical	2004	7k	3k	RR, admins	ipodlinux.org
ExampleProblems	MediaWiki	pd	English	topical	2005	<2k	27k	RR, admins	exampleproblems.com
Wikinfo	MediaWiki	GFDL	<i>multi</i>	all-purpose	2003	<2k	52k	admins	internet-encyclopedia.org
WikiTravel	MediaWiki	CC	<i>multi</i>	topical	2003	16k	43k	admins	wikitavel.org
World66	<i>proprietary</i>	CC	English	topical	2001	n.a.	131k	admins	world66.com
WikiWikiWeb	<i>proprietary</i>	pd	English	topical	1995	n.a.	32k	no admins	http://c2.com/cgi/wiki
DavisWiki	<i>proprietary</i>	CC	English	topical	2004	4k	10k	RR, no admins	daviswiki.org

Table 1. Some features of a few wikis. *Platform*: in this small sample, at least, MediaWiki is predominant. *License*: CC stands for “Creative Commons”, GFDL “Gnu Free Document License”, L.A.L. “License Art Libre”, pd “public domain”. *Language*: set of languages in use (actual or tentative), “*multi*” means multi-lingual (in which case statistics correspond to the main/most important language). *Governance*: “RR” means “Registration Required to contribute”; “admins” denotes the enforcement of community decisions by administrators, while “moderated” means that editions will be reviewed before being published. Data retrieved in May 2007.

wikisphere, we use an *arbitrary* selection of a dozen wikis; we gather some very basic features for these sites on Tab. 1 (although this random set may seem rather partial, we deem it sufficiently large and diverse to provide an intuition of both similarity and variety among wikis) (II) for a quantitative insight, we use a larger set of 7649 wikis selected from `s23.org`, a website providing statistics on MediaWiki-based wikis² — see Sec. 2.3 for detailed statistics.

2.1 Functional regularity

Some features are shared by many wikis: tautologically, these communities are all based on “wiki” platforms — **software** that generally allows modification of content by anyone, to at least some extent: despite a large variety of wiki interfaces³, key functional features enabling easy collective and collaborative content edition are eventually quite homogenous, including notably the possibility of creating an account and online identity, immediate inline content edition, comprehensive revision history and, often, discussion or “talk” pages [9]. Also determinant is the **licensing** mode of contributed content: regardless of the profusion of licenses in use on wikis, from the GNU Free Documentation License to Creative Commons and Licence Art Libre, *inter alia*, their function is essentially identical — allowing contributions from everyone while providing some protection on the content for future use, without the rigidity of classical copyrighted work. In some instances, no license at all is used and contributions enter in the public domain (e.g., WikiWikiWeb or ExampleProblems). From a **structural** point of view, a technical study of article networks in Wikipedias in several languages [11] suggested that wiki sites share some topological features with WWW networks, notably identical distributions of local patterns, scale-free degree distribution, unassortative linking (highly linked pages are connected to weakly linked articles) and high clustering/transitivity.⁴

² Retrieved on 5/5/07. `wikia` wikis were excluded as artifacts, with roughly the same number of users (~135k) and very few pages (≤ 10), suggesting an automatic process.

³ The Wikipedia page on wikis denombates at least 60 of them. Among them, MediaWiki appears to be prominent, possibly because of its exemplar use in Wikipedia — it also makes the large majority of our arbitrary selection (Tab. 1) and supports our quantitative case study, with several thousands of wikis based on it. To our knowledge, there is however no easy way to assess its exact usage.

⁴ The fact that article networks in wikis and WWW are similar should not be surprising *per se*, yet other networks could be defined in wikis (such as social and socio-semantic

2.2 Organizational variety

In contrast to this functional homogeneity, several semantic and organizational dimensions strongly differentiate wikis. Foremost is the **scope**: there are all-purpose projects aiming at building broad encyclopedias, vs. topical projects based on narrower matters, e.g. precise technical or geographical areas, particular political flavors. Scope includes **language(s)**, as wikis may aim at building multi-language content. In this case, language may or may not define distinct communities, depending on the existence of bridges between various language sub-wikis, possibly typical of distinct culturally-biased behaviors or representing various instances of a same system at distinct development stages, the latter being more a matter of knowledge diffusion than cultural relativism.

Editing content within the scope of the wiki project and using its prescribed language(s) may be among the most basic policies a benevolent user has to respect. Wikis generally obey to several other types of **policies** [3, 7, 9], which may be explicit (rules, usually presented in dedicated meta-articles) or implicit (norms), some of the latter potentially becoming explicit as crystalized norms — all such policies depending on the particular, historical development of each wiki. For instance, the so-called rule of “Neutral of Point of View” (NPOV) in encyclopedia projects such as Wikipedia or Citizendium is indeed not common to all wikis: it does not exist in some topical wikis focused on travel (such as WikiTravel or World66), environmental or alternative issues (like Ekopedia), and even in some encyclopedia projects such as for instance Wikinfo which precisely encourages diverse points of views for a same article.

Traditional rules include editorial scope policies or behavioral prescriptions, determining e.g. how to deal with conflicts. Some wikis have policy pages, enabling users to themselves evolve the wiki rules. On the same level, some rules essentially relate to governance issues and have a direct influence on how the wiki is managed: non-open and/or asynchronous encyclopedia such as Citizendium obviously require a different institutional structure and administration than more “anarchic” wikis, such as WikiWikiWeb.

networks: interactions between authors or around articles) and conceivably reveal patterns typical of stable wikis or functionally helpful towards their viability.

Eventually, policy variations extend to strict **technical policies**, directly opposed to the low-level technical similarities emphasized in Sec. 2.1. In particular, some criteria relevant for appraising technical policies in other types of collaborative online systems (such as tagging systems [14]) and pertaining to contributor rights, access rights to other users' information and social interaction opportunities, *inter alia*, may be valid for wikis as well. Drawing a comprehensive typology of technical divergences between wikis could start with questions such as: who can contribute? (anonymous visitors; registered users; approved contributors), how are interaction and social identity building concretely organized? (user pages; discussion pages; off-wiki media such as IRC channels or mailing-lists), what kind of technical features are available? (existence of special pages such as category, project, policy pages; possibility of uploading multimedia content). In turn, these diverse features are likely to fundamentally influence wiki viability, plausibly providing large-scale levers to shape their evolution.

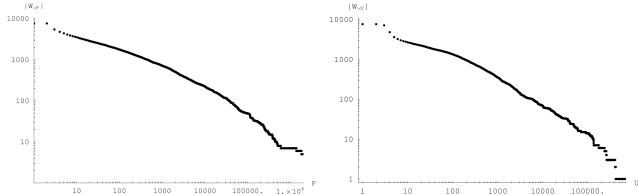


Figure 1. Cumulative numbers of wikis having at least P pages ($W_{>P}$, *left*) and U users ($W_{>U}$, *right*).

2.3 Demographics

To shed some light on how this diversity translates in quantitative terms, we examine basic statistical and demographic parameters on the above-mentioned dataset of 7,649 MediaWiki wikis: (i) population size, or number of registered users U ; (ii) content size, or number of so-called “good pages” P (i.e. excluding technical pages such as redirection and talk pages); (iii) activity measures, or number of edits E .

First, distributions of wiki populations and content sizes are very heterogeneous, with many wikis being weakly populated or having few pages, and a handful of wikis having a large number of users or articles — the cumulative distributions, presented on Fig. 1, exhibit power-law shapes. However, this statistical parameter hides the much deeper diversity of the wikisphere, which displays a rather unclear relationship between the number of articles and the population size. Indeed, for the 95% least populated wikis ($U \leq 898$), content sizes appear to be spread over several orders of magnitude (roughly, P is essentially included in $[1; 2 \cdot 10^3]$), with best linear fits being generally weakly accurate and without any obvious scale for any of these variables, as the density plot emphasizes (Fig. 3, top and left). Most of these relatively weakly populated wikis, with less than a thousand users and thus perhaps generally younger, may correspond to diverse growth stages with significantly varied success. Conversely, for larger numbers of users and, more precisely, for the first percentile of population size ($U \geq 8111$), the spectrum is narrower with a more monotonous relationship between both variables (Fig. 3, right).

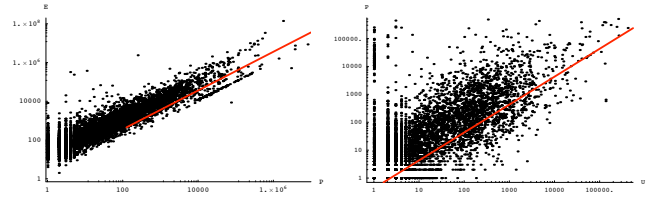


Figure 2. *Left*: activity, or number of edits E as a function of “good pages” P . *Right*: scatterplot of P vs. number of users U . Each dot is a wiki, the thick line shows the best linear fit.

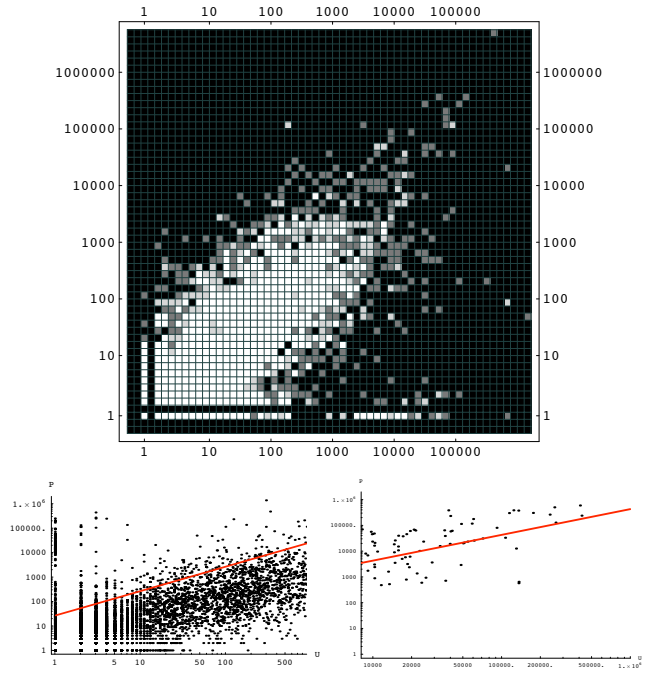


Figure 3. Density plot of P vs. U (*top*), and scatterplots, as on Fig.2-right, but plotted for the 95% least populated wikis only (*left*) and the 1% most populated ones only (*right*).

A plausible interpretation of the “fat” area on Fig. 3 could be that it captures both starting wikis (few users, few to many pages) and prospering wikis (more users, more pages) with all intermediate configurations possible and apparently equally likely. Nonetheless, activity exhibits a much more regular pattern and appears to be roughly proportional to the number of pages (Fig. 2): hence, activity directly scales with content size, whereas users do not seem to.

3. Life and death of wikis

In this context, what makes a wiki a lively, or at least viable, community? Technically, it consists of just two things: (i) a group of users, which may or may not become an active community, and (ii) a set of pages or articles, which may or may not become esteemed content. The growth and dynamic stability of a wiki community should thus helpfully be measured through two variables, the number of users U and of articles P . These variables, in turn, are linked in a subtle way: if more users allow for more articles to be created and cared of, more articles both require more users to maintain them and attract more users because of a better content. In short, population is dependent on content, and vice versa. The viability of a

wiki is thus a dual issue, in terms of **population dynamics** (recruitment, retention, and exclusion or leave) and **content dynamics** (growth as well as stabilization, with quality articles), which co-evolve. We hereafter review various factors that may impact one or both of these dynamics.

3.1 Bootstrapping and founding principles

Founders usually design both the wiki purpose and initial policies, determinant for bootstrapping the community and attracting its pioneer contributors. Sanger [15] notes that designing Wikipedia's founding principles had been a centralized process — only applying them has later been decentralized, through “administrators”. Further, the birth of a new wiki may precisely be triggered by the desire of introducing mutations on the core set of rules governing an already-existing wiki, and therefore experiment the effect of new bootstrapping rules [16]. The myriad of encyclopedia projects inspired by Wikipedia are all such instances: for example, while *Citizendium* removed the possibility of immediately editing articles, *Wikinfo* introduced the SPOV.

Various policies, as a result, have *various authoritative-ness*: while some of the founding principles appear to be non-negotiable — they make the core spirit of the wiki, if not part of their purpose — other rules are more permissive and prone to discussion. In particular, if the rule that there should not be any copyrighted material on the site is generally unchallenged on the Wikipedia, on the other hand, the NPOV rule is rather open and often re-interpreted. Old policies can generally be modified, while new rules can also be proposed, such as categories and related pages which only emerged in 2004 in the Wikipedia [9, esp. Fig. 5].

As such, various policies also obey to *various timescales*: while some rules are discussed often and are potentially modifiable by new users, founding principles are much more stable. Wikipedia, again, makes a distinction between “policies” and “guidelines”⁵; in this sense, rules may have distinct legal statuses, with the core set of policies playing the role of a constitution [16]. Still, any Wikipedia policy page has the following banner: “When editing this page, please ensure that your revision reflects consensus. When in doubt, discuss first on the talk page” suggesting indeed that *all* policies are modifiable, including the most fundamental ones like NPOV — they may just have a very long-lasting effect and thereby provide the impression that they are absolutely irremovable and/or controlled by a small set of established users.

3.2 Population dynamics

Enabling user incentives. While bootstrapping policies define a context of primary incentives, there are more broadly general incentives for participation in online or open communities, which are likely to apply to wikis as well. Several previous works have proposed typologies of the ingredients making such communities attractive, including studies on the Wikipedia [3, 5], tagging systems (e.g. Flickr, CiteYouLike, etc.) [14] or open-source software development [17], or reviews of past literature on web-based community survivability [18], to cite a few. On the whole, these studies may offer

a relevant guide for analyzing wikis in the general case; for instance and in particular, users may generally be interested in building one's own identity, elaborating a common good useful to oneself or to others, improving and displaying one's creativity, *inter alia*. Not all users should display sensibility to incentives of some kind, nor be driven by only a certain kind of incentives. Broadly speaking, however, it seems to be possible to distinguish three classes of incentives:

- “altruistic”, by benevolently contributing to a public good (ensured by the type of licenses used by wikis);
- “socially concerned”, such as influencing others or a broader, less visible audience; sharing with others, belonging to a community, meeting/interacting with people, being seen by others and building an online reputation and identity to climb an online social hierarchy;
- “selfish”, such as improving one's skills, or enabling the future retrieval of information for one's own usage.

The variety of combinations of participation, edition and technical policies creates diverse incentive landscapes, plausibly determinant for the viability profile of each wiki. Of course, a minimal requirement would be that at least the topic is appealing to some users — possibly providing an advantage to broad, all-purpose projects such as the Wikipedia, where anyone could think of being an expert in something.

Enrollment and leadership. Incentives are, especially at the beginning, crucial ingredients for retaining contributors and recruiting new users. Bryant et al. [3] carried one of the first ethnographic inquiries on a wiki-based community by describing the whole involvement process and distribution of labor on the Wikipedia, from initial discovery of the wiki and its rules, to first contributions such as correcting typos, account creation, up to becoming administrators. Eventually, this process compares with how apprentices progressively integrate into a community of practice, being granted knowledge and use of more and more skills, with older members lenient towards newbies. As for an explicit distribution of roles, hierarchies in wikis often reduce to a simple dichotomy between contributors as suppliers of content and administrators as enforcers of community decisions. Still, this rudimentary institution not only gives administrators legal *and* actual power over less established users, but also broadly relates to periphery–center roles. Kittur et al. [6] exhibit quantitative differences between Wikipedia administrators and users: the former are often among the most active contributors; they informally owe their status to remarkable involvement. In addition, their edits themselves seem to be essentially different. Normal users indeed appear to make less changes than administrators, deleting more words than they add, thus suggesting that elite users contribute more to the wiki growth.

As a result, administrators in general do not distinguish themselves only on formal powers: they also build their informal authority through humble and sustained efforts towards the project [8], such as being concerned with explaining policies and transmitting norms to newbies [3]. In turn, respected members are likely to gain some command when crises arise. In an alleged egalitarian wikipedia, and very probably in wikis in general as it does in other virtual communities [17],

⁵ http://en.wikipedia.org/wiki/Wikipedia:Policies_and_guidelines

implicit authority commands: as such, leadership substantiates one aspect of user incentives — virtual identity building is not vain — and shed light on wiki viability — in spite of a purported decentralized decision process, leaders may exert some limited *centralized* power to sustain a wiki.⁶

Growth and decline. As shown in Sec. 2.3, wiki population mildly correlates with content size. The exponential growth of the Wikipedia in terms of both articles and users suggests that successful wikis follow a virtuous demographic path, with content and contributors co-evolving and most likely positive feedbacks between both. In the present study we did not gather dynamic data and could not elaborate further on potential development regimes which could be observed in the larger wikisphere. Nevertheless, the dichotomy between simple contributors and more involved administrators appears to induce distinct population dynamics [6]: after a first period where expert users are the largest contributors in terms of edits, and where their influence grows, a second step exhibits a larger overall activity from small contributors. Initially, thus, a small group of people becomes involved with an embryo of content, bootstraps and establishes the foundations of the wiki, as said above; then the wiki should ideally evolve towards a phase of recruitment of more users and/or content of better quality. While there does not seem to be a clear-cut transition between the two modes, one may distinguish the gestation period from the “cruising” mode. This process may support the explanation of the “wide” area at the lower-left of the scatterplot of Fig. 3, with wikis bootstrapping and wikis at a more advanced stage altogether.

Open questions relate to which wiki-related reasons may make people leave. As Levrel [16] notes, holding contributors in an open project is a challenge in itself, notably when the project grows and requires more and more supervising and strongly involved users: in the French-language Wikipedia for example, some administrators required their number to be increased under the threat of leaving the project, as they were at some point too few to maintain the ever-growing content. Another dreadful threat lies in forks, i.e. concurrent wikis emerging from a given project while re-using the same content as a seed, dividing communities and significantly lowering the available workforce for each fork, possibly increasing confusion between them. On the bright side, yet, the open-licensing system makes *coups d'état* pointless, as anyone can create a fork whenever the project appears to be monopolized by a small club. For instance *Enciclopedia Libre*, a fork of the Spanish-language Wikipedia, appeared at the very first rumor that advertising could be used to sustain maintenance costs — the project going “private” would suppress one fundamental user incentive, the contribution to a common good.

3.3 Content dynamics

Some factors underlying population growth are directly linked to content characteristics. Quality content, attesting a successful collaborative project, is likely to have a positive impact on population dynamics. Content growth, on the

⁶ See Wikipedia’s founder, Jim Wales, who could willingly play the role of “a British monarch” [19, 8].

other hand, has a more ambiguous effect, by both improving the completeness of the project (hence the quality) and the number of users needed to maintain it. In particular, the “upper-left” wikis on Fig. 3, with a deficit in users compared to content size, are likely to experience a higher threat of abandon, relinquishment or vandalism than other wikis. Given these constraints, how to ensure sustainability?

Content stabilization. First, “content disasters” [5] are impossible on wikis: the effortless reversibility discourages mass vandalism. More interesting are regular threats towards content consistency, including inaccurate contributions, occasional vandalism, more or less ostensible rule-breaking and edit wars. Whereas the grossest vandalism does not require more than an appropriate number of users watchdogging wiki contents⁷, fact falsification or disputes among editors, e.g., induce more subtle regulation and arbitration procedures.

Conflict-solving mechanisms are in this case well-developed, talk pages being a prominent means to achieve this [9], in particular in relation with (self-organized) policies which are interactively set up by Wikipedians, precisely through talk pages themselves — such as, for instance, the “3-revert” rule on Wikipedia,⁸ which in turn induces a certain robustness to vandalism of various kinds. Later, discussion pages are also the place where such policies may be enforced, and eventually, as such, the locus of the coevolution of disputes and policies; Viegas et al. [9] emphasize more broadly that a conventional organization could emerge on wikis from activity patterns in talk pages, rather than from a top-down imposition of norms by some agents. It is therefore a crucial instrument on which contributors may act to restore the viability of a wiki, either at a local level, e.g. introducing page protection to avoid vandalism, or at a more global level, e.g. by defining new ways of dealing with conflicts⁹. Finally “offline”, i.e. off-wiki, social networks offer immediate assistance to more serious conflicts [6], as informal places where a smaller and more involved part of the community as a real social network discusses and acts in common outside of the wiki, including IRC channels and mailing-lists, notably.

Here, it is likely that wikis without a governance structure (i.e., administrators) would have more trouble dealing with conflicts and achieve viability. As a whole, these various kinds of stabilization mechanisms suggest that viability can be enforced by various kinds of actors, from very local, bottom-up interactions to global, top-down constraints imposed by a small group of “experts”. To another extent, policy evolutions have a meta-impact on the content by modifying the landscape of user incentives, inducing a modification on the longer-term dynamics — as such a method to adapt to evolutions in user demand and sustain a wiki.

Quality evaluation. A last feature of content stabilization relates to its quality, as a key criterion to attract users [3]; and

⁷ On Wikipedia for example the “Recent Changes patrol”, a self-organized group, checks recently-modified pages for vandalized content.

⁸ Any user making 3 successive reversions of others’ modifications of an article is banned from further editing this article for some time.

⁹ For instance, to avoid expelling worthy contributors having an occasional issue on some article, Wikipedians replaced their policy for dealing with “editor-with-problems” with “relationship-between-two-people-with-problems”, thus without focusing on a single person [20].

indeed content-stabilizing policies intertwine quality policies: on Wikipedia for instance, Stvilia et al. [12] show that the notion of “featured article” sets a quality target, while “articles for deletion” are a permanent threat establishing a minimal quality threshold. Apart from strict vandalism (deletion, corruption, obvious non-respect of rules) quality remains hard to appraise with objective criteria, some criterias may nonetheless be induced and operationalized from studies focused on content quality on one side, and user quality on the other side. As for content quality, Lih [1] has established criteria correlating quality to the number of edits (“rigor”) and editors (“diversity”), while a correlation has recently been established by Wilkinson & Huberman [10] between Google’s PageRank and edit activity; meanwhile, a hand-made comparison of Wikipedia articles with Encyclopedia Britannica unsuccessfully looked for significant differences [13]. As for user quality, Anthony et al. [2] measure the survivability of edits made by Wikipedians and identify two categories of worthwhile actors: registered users producing many edits (experts) and anonymous users making few edits (passers-by, who are more and more [6]), the importance of this last category being counter-intuitive (esp. with respect to [3]), as novices-with-respect-to-a-wiki appear in this case to produce better content than a whole class of registered users.

4. Conclusion

We exhibited both the current state of the wikisphere in terms of resembling and differentiating features and properties, and a few key mechanisms likely to be determinant in explaining diverse destinies of wiki-based communities: bootstrapping settings, initial recruitment processes, incentives to stay in the community and relationships between contributor and content dynamics. From this, we may now sketch out some of the most salient patterns for modeling wiki viability:

- Population dynamics: **(1) Agents:** growth is directly co-evolving with content size (Sec. 3.2); a proportion of users leave, partially depending on the overall quality. **(2) Status:** registered users improve over time; a small subset of most active users is granted administrator status, making better edits, lowering vandalism; some anonymous visitors are better than recently-registered users [2].
- Content dynamics: **(1) Articles:** their growth is linked to population size and existing articles; quality could be assessed through the number of editors and edits [1], the latter proportional to previous edits [10]. **(2) Meta-content:** recreating the ecology of policies would be limited by ontological uncertainty [21], unless one reproduces observed or predefined norms histories; yet, policies could still be independent variables in order to compare development paths induced by various policy sets.

Such population and content dynamics might be sufficient to describe and perhaps predict the viability and history of wiki communities. In this respect, the present paper could be seen as a position paper. Since existing work on wikis is essentially focused on the Wikipedia, it could be instructive to carry additional qualitative research on behaviors in other wikis. Besides, further research should endeavor at observing the growth and perhaps decline of a large set of wikis, instead

of focusing on their structure at some static timepoint, so as to distinguish a variety of development paths.

References

- [1] A. Lih. Wikipedia as participatory journalism: Reliable sources? Metrics for evaluating collaborative media as a news resource. In *5th Intl Symp on Online Journalism*, Austin, TX, USA, April 16–17 2004.
- [2] D. Anthony, S. Smith, and T. Williamson. Explaining quality in internet collective goods: Zealots and good samaritans in the case of wikipedia. 2005.
- [3] S. Bryant, A. Forte, and A. Bruckman. Becoming wikipedian: Transformation of participation in a collaborative online encyclopedia. In *Group’05, Sanibel Island, FL*, Nov 2005.
- [4] A. Capocci, V. D. P. Servedio, F. Colaiori, L. S. Burio, D. Donato, S. Leonardi, and G. Caldarelli. Preferential attachment in the growth of social networks: The internet encyclopedia wikipedia. *Phys. Rev. E*, 74(3):036116, 2006.
- [5] A. Cifollilli. Phantom authority, self-selective recruitment and retention of members in virtual communities: The case of wikipedia. *First Monday*, 8(12), 2003.
- [6] A. Kittur, E. Chi, B. Pendleton, B. Suh, and T. Mytkowicz. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie.
- [7] J. Levrel. Wikipédia, un dispositif médiatique de publics participants. *Réseaux*, 24(138):185–218, 2006.
- [8] J. M. Reagle. Do as I do: leadership in the Wikipedia. <http://reagle.org/joseph/2005/ethno/leadership.html>, 2005.
- [9] F. Viegas, M. Wattenberg, J. Kriss, and F. van Ham. Talk before you type: Coordination in Wikipedia. In *Proc 40th Hawaii Intl Conf on System Sciences*, 2007.
- [10] D. Wilkinson and B. Huberman. Assessing the value of cooperation in Wikipedia. *First Monday*, 12(4), 2007.
- [11] V. Zlatic, M. Bozicevic, H. Stefancic, and M. Domazet. Wikipedias: Collaborative web-based encyclopedias as complex networks. *Phys. Rev. E*, 74(1):016115, 2006.
- [12] B. Stvilia, M. Twidale, L. Smith, and L. Gasser. Information quality work organization in Wikipedia. 2007.
- [13] J. Giles. Internet encyclopaedias go head to head. *Nature*, 438(7070):900–901, 2005.
- [14] C. Marlow, M. Naaman, d. boyd, and M. Davis. Position paper, tagging, taxonomy, flickr, article, toread. In *Proc of Collaborative Web Tagging Workshop, WWW 06*, 2006.
- [15] L. Sanger. The early history of nupedia and wikipedia: A memoir. <http://features.slashdot.org/article.pl?sid=05/04/18/164213>, 2005.
- [16] J. Levrel. Le développement des contenus libres sur Internet - Wikipedia. Technical report, FT R&D TECH/SUSI, 2005.
- [17] A. Mockus, R. Fielding, and J. Herbsleb. Two case studies of open source software development: Apache and Mozilla. *ACM Transactions on Software Engineering and Methodology*, 11(3):309–346, 2002.
- [18] L. Gallant, G. Boone, and A. Heap. Five heuristics for designing and evaluating web-based communities. *First Monday*, 12(3), 2007.
- [19] T. McNicho. Wikipedia founder hunts for gold. *Business 2.0*, CNN, February 27, 2007.
- [20] J. Levrel. Personal communication, 2007.
- [21] D. Lane and R. Maxfield. Ontological uncertainty and innovation. *J. Evolutionary Econ.*, 15(1):3–50, 2005.