

User Engagement on Wikipedia: A Review of Studies of Readers and Editors

Marc Miquel Ribé

Universitat Pompeu Fabra marc.miquel@upf.edu

Abstract

Is it an encyclopedia or a social network? Without considering both aspects it would not be possible to understand how a worldwide army of editors created the largest online knowledge repository. Wikipedia has a consistent set of rules and it responds to many of the User Engagement Framework attributes, and this is why it works. In this paper, we identify these confirmed attributes as well as those presenting problems. We explain that although having a strong editor base Wikipedia is finding it challenging to maintain this base or increase its size. In order to understand this, scholars have analyzed Wikipedia using current metrics like user session and activity. We conclude there exist opportunities to analyze engagement in new aspects in order to understand its success, as well as to redesign mechanisms to improve the system and help the transition between reader and editor.

1. Introduction

Wikipedia has become the paradigm of collaborative creation success as well as an icon of Internet possibilities. Since 2001, it has grown to 4,5 M articles in the English edition and 34 M in total counting the 288 languages in which it is available. However, the most surprising is that this process has been made a reality by thousands of editors who have devoted their free time, converting it into a free product for mass consumption, while aiming at “gathering the sum of all human knowledge”. Moreover, it is used all over the world use it, and this is confirmed by its position in the top 10 Alexa rank¹ of most visited sites.

Many researchers have tried to understand how the system works, or in other words, what the pillars of its success are. User Engagement framework defines the attributes which constitute an engaging experience (O’Brien and Toms 2008). Reliability, trust and expectation, richness and control are some which have been studied by scholars

on Wikipedia. Their results help in explaining how engagement occurs on both the reader and editor sides. Yet, there are some attributes presenting room for improvement detected almost five years ago; overall usability and the design of particular communication channels could be revised in order to mitigate frustration.

The duality between the two groups of users, readers and editors, has acted similarly to a feedback loop system; new content availability helped to popularize the encyclopedia and improved the position for searchers, which in turn increased its use and its editing base in order to create new articles. Suh et al. (2009) explains this growth as a self-reinforcing mechanism, the more valuable Wikipedia became the more contributors joined it and gave value to it. Although during the 2015 first quarter the number of editors increased², the general trend during the last few years has been a soft decline. Ortega, Gonzalez-Barahona and Robles (2008) found that a very engaged minority of editors was responsible for most of the activity, and Stuart and Halfaker (2013) verified that those who joined in 2006 are still the most active group.

Thus, questions like how readers become editors or how to raise writing activity have become relevant to the community, Wikimedia Foundation and scholars (Okoli 2014). Triggered by them, researchers applied some of the most usual metrics in user experience, such as session analysis, in order to analyze the different types of editors. Wikipedia is a very suitable object for analysis with longitudinal data and every action performed tracked in its databases.

The aim of this study is to put together all the studies from the perspective of readers and editors perspective to give an integrated overview; we want to understand the uses and difficulties users encounter during their experiences. We want to see how they engage and disengage, as stated by Attfield et al. (2011) in the definition “the relationship they establish at a behavioral, emotional, and cognitive level”.

Copyright © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹ <http://www.alexa.com/topsites>

² https://meta.wikimedia.org/wiki/Research:Active_editor_spike_2015

We begin by examining studies concerning some aspects of engagement in Wikipedia; then we review their attributes and how they have been described. We assess the state of the community with its social and structural characteristics; we want to see its composition, current engagement and its growth possibilities for growth by attracting new editors. Next we review all the available literature that analyzes readers and editors, both by using metrics based on user activity, community and topics. Finally, we argue for some improvements proposed to current research and outline future engagement research.

2. How Wikipedia engages

Wikipedia has been approached in numerous peer-reviewed academic articles both as a data source and as a study object to understand how it works (Okoli 2009). Although many facets of the encyclopedia have been studied profoundly, no research covers directly its quality of engagement as a whole, either for readers or editors, to understand how a combination of factors contributes to its quality and repeatable experience.

2.1 Wikipedia as an everyday tool

Engagement studies have been applied to many different contexts, from games and educational sites to e-commerce and online news (O'Brien 2008). The broad application of the concept has brought very different outputs; literature has differentiated everyday engagement from games and other kinds in the sense that it is a less immersive experience (Lalmas and O'Brien 2014). Wikipedia is this kind of engagement; the encyclopedia has become an object integrated into our everyday work and personal lives. Wikipedia usage has been reported in foundation studies³ as divided into different devices including phones, tablets and personal computers. Wikipedians often multitask and edit Wikipedia while watching TV, even chatting in IRC Wikipedia dedicated channels or any other social network.

Although editors and readers behave differently depending on the role they are taking, their behavior repeats over (Stuart and Halfaker 2013). Surveys and analytical studies identify some users who spend several hours a day. A usual behavior is to switch from article to article using their wiki hypertextual structure. Explorative navigation allows us to frame Wikipedia use as an inter-site engagement, the kind of experience in a network of pages in the same site (Yom-Tov et al. 2013).

³ http://upload.wikimedia.org/wikipedia/Editor_Survey_Report_-_April_2011.pdf

2.2 User Engagement Framework

User Engagement has been an evasive concept to define, for it has been used to express qualities similar to attracting, captivating or enticing depending on the study (Lalmas and O'Brien 2014; Chapman, Selvarajah and Webster 1999). Some researchers grounded it to psychological theories like Flow, Play and Aesthetics, which can explain different aspects of it (O'Brien and Toms 2008). Flow, for instance, describes the state in which there is attention, control and an intrinsic interest (Csikszentmihalyi 1997). Many studies presented engagement as a quality of user experience with certain attributes that influence or compound it. O'Brien and Toms (2008), in their User Engagement Framework, listed them as challenge, aesthetic, sensory appeal, feedback, novelty, interactivity, awareness, motivation, interest and affect. Additionally, the duration of the experience or its repetition has been considered a clear indicator of engagement (Attfield et al. 2011).

From the site manager perspective, creating engagement is positive as it gives continuity to its users. Wikipedia aims at creating high-quality information through the engagement of a broad and multilingual community. As a website its success depends on its qualities to captivate editors as well as to give content to fulfill readers' expectations. Engagement attributes like reputation and trust, novelty and attention are more linked to the readers, while user context, motivation, usability and positive affect in addition influence editors more directly.

I discuss each of the most common attributes in engagement research and their relation to Wikipedia studies.

Reputation, trust and expectation

Wikipedia's reputation has always been questioned along with its reliability. Readers need to know if the content they are reading is trustable; an Encyclopedia created by a group of anonymous people initially seemed audacious and doomed to failure. However, in 2006 a study compared it with the Encyclopedia Britannica and showed that it had fewer errors (Giles 2005). One study showed that the better the coordination between editors, the higher the quality of the articles (Kittur and Kraut 2008). Wikipedia's main editing rule is to reach a Neutral Point of View (NPOV) within each article. Instead of objectivity, a contrast of different positions and their representation in the text is required to editors for an article to have quality.

An issue threatening content quality is vandalism. It is confronted with policies and bots, which can restore old content and ban the user who is misbehaving. Generally, authors work on articles providing all the available data and references. Lucassen and Schraagen (2010) in a study developed the features by which an article is considered valuable (text, images, references). Likewise, the commu-

nity considers which articles have quality characteristics of quality and list them in 'featured articles'. In the end of the day, expectation depends on the use intended and Wikipedia is often reported as a background information resource (Head and Eisenberg 2010)

Novelty

Wikipedia's welcoming message is "the free encyclopedia that anyone can edit", and it should add "at any time". Immediacy is one of the virtues of the site, which translates into almost a real-time event conversion into a Wikipedia article – any political elections or celebrity deaths are updated by a hectic group of editors. This coverage is not as instantaneous as in the social network Twitter, but it provides a rich information background to understand reality that fulfills readers' expectations (Osborne et al. 2012). Complementary to the sense of familiarity or met expectations being met, novelty can also promote re-engagement as it calls us to something new and unexpected (O'Brien and Toms 2010). In Wikipedia the sense of novelty and surprise appears when readers find updated information and additionally do 'hypertextual reading', that is jumping from one page into another in the search of understanding and exploring new concepts (Zhang 2006; Lehmann et al. 2014).

Attention

Attention is a characteristic of engagement, which is defined in detriment of any other activity. The state of flow, which implies attention, has been reported both while reading or writing (Csikszentmihalyi 1997). Lehmann et al. (2014) studied reading patterns and detected four main behaviors, which were named: focus, trending, exploration and passing. She found that the state of concentration on reading is very related to the kind of content and task users are performing - that can be learning about a topic or checking a specific date. Also, Wikipedia is accessed in very different circumstances, for instance the scenario of multi device access by readers and editors. Focused attention may appear in some behaviors while using Wikipedia, although this is not the unique type of engagement generated by the encyclopedia.

User context and motivation

The continual use of a system depends on multiple contextual factors that can facilitate it or totally discourage it. Wikipedia is accessed multi device, which enhances task division, like saving information for later, as well as providing an adaptation to multiple momentary informational needs (Pande 2011). In order to contribute, editors' behavior is often explained by motivation studies, which try to understand what are the reasons that push somebody into an action. Nov (2007) found by surveying editors that

fun, ideology and community values were significant in a non-reward scenario like Wikipedia. Other important reasons for editors were the same process of learning skills, socializing with other peers and developing a writing career. Complementary studies (Yang and Lai 2010) proved that a self-concept based motivation is the most important in Wikipedia; reputation within a community after accomplishment, gaining autonomy and even experiencing reciprocity were found relevant.

Richness and control

Play theory explains that an activity involving creativity or learning satisfy social and psychological needs, and aspects like competition and collaboration are promoted (Rieber 1996). The sense of learning through different levels until achieving user's expertise has been defined as "richness and control". In Wikipedia, editors go through a process of learning to edit according to policies and by means of several tools. For instance, user pages act like a sort of personal profile, where editors leave messages to other editors. Another interesting tool is the 'watchlist', which enables following several Wikipedia articles. The progression within the community, in order to become a fully operative editor, can be considered challenging and stimulating. Although the increased complexity has been also reported as a cost with negative impact on production (Suh et al. 2009).

At the very first stage, editing has been considered difficult due to aspects like the wiki-markup, a similar language to html with specific tags native to the MediaWiki system which Wikipedia uses. Several issues regarding site usability ("poor interface", "cluttered") have been reported in the context of using Wikipedia as a pedagogical tool (Raitman, Augar and Zhou 2005). In 2012 a MediaWiki extension, VisualEditor provided a What You See Is What You Get Editor⁴ in the same way as editing in a word processor to solve some of this issues. Wikimedia Foundation has released tools for translation and language switching, mainly solving initiated editors' needs in their usual writing tasks.

Positive affect

Positive affect has been reported as an attribute of engagement that can improve task involvement at an early phase (O'Brien and Toms, 2008). When users are still discovering how to operate within a system, receiving positive emotions helps in building the relationship and creating loyalty. In Wikipedia editors learn to edit by trial and error; sometimes when they contribute to an article, their edits are eliminated by more experienced editors - they are 'reverted'. This is considered a necessary behavior, although sometimes due to the lack of communication it has had very negative effects on new contributors

⁴ <http://en.wikipedia.org/wiki/Wikipedia:VisualEditor>

(Halfaker, Kittur and Riedl 2012). Reducing conflict is a way of improving user experience, since negative emotions have been associated to disengagement eventually causing lack of involvement (O'Brien and Toms 2008). Laniado et al. (2012) found that emotions and dialogue in a peer-production community like Wikipedia were mainly found in article discussion pages; sentiment-analysis showed that emotions took forms of emoticons or virtual gifts. In the list of recommendations they there is encouragement of positive tone, appropriate wording, as well as providing new ways of channeling negative feelings.

Aesthetic appeal

Another common characteristic of engaging systems is their aesthetic appeal (O'Brien and Toms 2008), in the visual aspects of the interface as well as in other graphic elements. Wikipedia has evolved since its very beginning in 2001, always along the visual standards of an encyclopedia with a clear and plain interface, but including a great assortment of images. Some of them are included in initiatives like the most beautiful picture of the year. Images are uploaded in the commons repository and are free to use under this license. However, Wikipedia is timidly approaching the audiovisual, with very few videos illustrating articles. So far, no research has studied Wikipedia in this particular aspect. Aesthetics is considered an important aspect of engagement but clearly only one aspect (O'Brien and Toms 2008), and not the most remarkable.

Most of the User Engagement Framework attributes find some representation in the scholarly studies based on the online encyclopedia. The completeness of mechanisms, contexts of use and applications help in providing an experience that millions of readers and editors are repeating daily. Endurability or the capacity to create a memorable and worth sharing experience has also been linked to engagement (O'Brien and Toms, 2008). The fun while creating articles was reported an important motivation to continue on their activity (Nov 2006). But in the beginning an editor starts with the joy of a reader, then after understanding and knowing the system better in aspects like reliability and novelty, decides to progress with learning the rules and tools, and finally engages in a contributing activity without never stopping being a reader.

3. State of the community

The state of the community has been a general concern for scholars ever since Wikipedia attained its sudden success and growth around 2007. The following year, studies started appearing in order to quantify statistically how contributions are divided between editors. Although the results from scholars and the same Wikimedia Foundation were

not alarming⁵, the community stopped growing in absolute numbers - editors were joining and leaving Wikipedia at a very similar rate. This slight decline of editors acted as a trigger for analytical research to work on it during the later years, and for some researchers to develop metrics and hypotheses on how editor retention works (Halfaker et al. 2014; Suh et al. 2009). Here we select the most important studies, paying special attention to those that give explanations on the causes of why new editors are the ones leaving.

3.1 The inequality of contributions: the decline of editors

Perhaps the first article that studies Wikipedia as a community and quantifies user work was from Voss (2005). He examined in German Wikipedia the distribution of distinct authors per article and found that they were following a general power law and the number of distinct articles per author followed a Lotka's Law. These statistical distributions explained that a minority created a great majority of the content. When Wikipedia had already achieved great popularity, Ortega (2008) widely validated these results using the top-ten Wikipedia languages editions. In order to calculate the level of inequality in the contributions, he used the Gini coefficient and found that more than 90% of the content can be attributed to less than 10% of the community. This remained constant for every language history.

Suh et al. (2009) examined different kinds of work and their weight in editors' activity and found that coordination (maintenance and discussion) or bureaucracy (formulating and discussing policies) was taking time that would have been instead dedicated to article creating. As everywhere, they suggested bureaucracy was part of the process of getting to system maturity. In addition, other authors like Butler, Joyce and Pike (2008) had already affirmed that the complexity of Wikipedia with its roles and policies acted as a bureaucracy. In 2012, other studies found that the community had decreased by a third (Halfaker et al. 2013). During those years of impasse, the slowing growth of Wikipedia articles was explained by an even increasing activity by the very active users and a diminished activity from the middle group of editors. One year later, it was demonstrated that editors who joined in 2006 were still more active than any other annual group (Stuart and Halfaker 2013) and the editors who were leaving were the new ones.

3.2 Identified barriers for new editors

In his study, Halfaker et al. (2013) confirmed the hypothesis of a settled bureaucracy acting as a barrier for new edi-

⁵ http://strategy.wikimedia.org/wiki/Editor_Trends_Study

tors. While norms were revised and expanded, new ones did not emerge at the same pace since 2006. Furthermore, newer editors were finding their policy propositions most likely rejected compared to previous editors, and this did not stop them from contributing to essays and community governance, but it was a clear restriction to overall activity.

Pancierera, Halfaker and Terveen (2009) investigated both qualitatively and quantitatively the nature of editors' activity. They found a recurrent pattern in every future high activity editor; when new Wikipedians made a large number of edits initially, the probability of becoming a highly active editor increased by 18%. These results explain why there is a certain kind of new editors who adapts to the bureaucracy with ease, even though these results are contradicted by the difficulties of the remaining editors, of whom unfortunately 60% never made another edit after the 24 hours of registration.

In a similar manner, another recurrent struggle for new editors is receiving reversion or rejection of their edits by experienced editors (Halfaker, Kittur and Riedl 2011; Suh et al. 2009). By returning to previous versions of an article, editors protect content from vandalism. However it has also been proved that when performed to new editors it drastically effects their future activity (Halfaker, Kittur and Riedl 2011; Zhang and Zhu, 2006). This resistance makes it hard for newcomers to penetrate in daily Wikipedia activities. In the worse case scenario reverts do not include feedback but negative comments. From the engagement theory perspective, the consequences of a negative revert were creating an interruption which brought disengagement (O'Brien, 2008). In other cases, when the editor was already familiar with the community he took being reverted as a learning experience (Halfaker, 2011) and increased the quality of their work.

So far, all studies conclude that it is essential to channel better communication in the initial phase for a new editor. Event notification and conversation channels are found to encourage engagement in online communities (Millen and Patterson 2002). Wikipedia has user spaces and discussion pages dedicated to each article, but researchers should study if the asynchronous communication they provide is not enough for current community needs. In this sense, Halfaker et al. (2011) proposed an interface change to inform editors about to revert a newcomer edit. After testing it in a trial group, he found that a simple warning message could improve the involvement and content quality from editors with different degrees of experience.

3.3 The breach between readers and editors

In addition to understand editor retention, an important focus of study has been directed to the transition from readers to editors. For most of the studies, the role of the reader

is considered as a second-class user, a passive user ("lurker"), compared to an active and well-coordinated contributor. However, this division only reinforces the difficulties of becoming an editor, instead of considering that a reader is a possible future editor still in a learning phase. Antin and Cheshire (2010) deployed a survey among 165 participants and found that readers become familiar with functional details and policies from the encyclopedia. Many contributors did not acknowledge the editing button and, instead of being self-interested they could be better defined as cautious. Similar results were found in a laboratory and remote testing organized by the Wikimedia Foundation⁶. Readers often did not notice all interface elements and felt often overwhelmed.

Halfaker, Keyes and Taraborelli (2013) discussed the implications of receiving new contributions and the necessity for moderating them. In his study, to help in bridging the transition between reader to editor, he introduced a new tool called "Article Feedback". It was implemented as a new UI layer on the Wikipedia article interface with the tag "Improve this article". After testing different tag prominence based scenarios, he could see that many readers used it to give their impressions, reflecting that the possibility of editing went unnoticed for many. His conclusions were that although unproductive edits and comments may appear, the proportion of good new edits still benefits the development of Wikipedia.

4. Measuring Wikipedia engagement

Due to its social and technical characteristics, enabling content sharing and interaction, Wikipedia has become a kind of "living laboratory" ideal for research (Suh et al. 2009). Every content change and editor action are stored in databases and XML dump, which are regularly provided by Wikimedia Foundation. From a content perspective, Wikipedia is the output of engagement; from the community, it is a technical artifact where editors perform actions in a journey across topics and articles.

Available engagement studies are primarily interested in measuring users' endurance, how they want to return to Wikipedia. They want to know if it fulfills any of their needs and if it creates a memorable experience (Lalmas and O'Brien 2014). Most of the engagement measurements in Wikipedia do not link to any of the attributes which we previously explained, but study its impact on editor behavior to characterize the state of the community and product development. Likewise, there is scarcely research dedicated to understand reader behavior.

⁶ http://usability.wikimedia.org/wiki/Usability_and_Experience_Study

4.1 Wikipedia editor metrics

Basically, two main types of engagement metrics are used, session and activity. They are divided regarding time and the specific actions performed. The smallest portion of a user activity is an edit, which includes text changes in any page: articles, user pages, and discussions. The rate at which an editor saves revisions to pages can vary substantially based on their wiki-work habits and the kind of activity they are engaged in. Some even work in a text processor and later paste their work before submitting.

In User Engagement, a same type of website (e.g. news) can give very different results when analyzing the same metrics (Lehmann et al. 2012). For instance for a user, time spent in a site in a single session (known as dwell time), can be either indicative of entertainment or not finding what they are looking for. In the case of Wikipedia, editing work can be considered as the longer the session time the more engagement. But what is a session? The concept “session” has several definitions; they share in common the acceptance of a period with queries and a following one dedicated to examination. Stuart and Halfaker (2013) examined sessions in and found a habitual inactivity threshold at one hour in a distribution of actions over time.

Absence time or the time between visits is also important because it allows understanding of recurrent behaviors (Dupret and Lalmas 2013). Intersession analysis explains how this absence time is linked to the kind of action to be performed immediately afterwards, hence the existence of difference behavior patterns. Stuart and Halfaker (2013) studied and detected behavior patterns at the within session and between level in Wikipedia. He suspected that there would exist different session lengths and found three differentiated distributions: within-session (minutes), between-sessions (days) and extended session breaks (months). The advantage of measuring sessions is that they provide unequivocal information on labour. The metric edit count (number of edits) per user, although it may have much importance to some editors, does not give any detail on what kind of actions have been performed - the length of an edit could be an entire article, a comma or a comment in a discussion page. When quantifying wiki changes number of bytes is taken into consideration. However, there is no best metric since it depends on the context of the study and research question.

At a community or a multi-user engagement level, network metrics can explain complex dynamics among editors on a single article or a mesh (Kaltenbrunner and Laniado 2012). Other studies have shown that mixing discussion and edits in network analysis could explain a quality increase by means of editor coordination (Kittur and Kraut 2008). Kaltenbrunner and Laniado (2012) analyzed real-time events, the articles that covered them and their Wikipedia discussions, and found that they reached depth that

would take years for other types of articles. In other words, depending on the topic and the speed of reply engagement varied in developing discussions and further editions.

Taking everything into account, the maturity of the existing metrics applied to Wikipedia is relatively high. They respond to most of the situations that could characterize the community or individual behavior. However, activity can boil down to many differentiated tasks - adding new content, contributing to a discussion, correcting typos or translating articles. There is still a possibility of studying this specific area of task specialization and different types of user since there is no study focused on it.

4.2 Wikipedia reader metrics

A recent reader behavior study takes popularity, by means of pageviews per article, and reading session characteristics (Janette et al. 2014). They presented different readers' behaviors in a complex method using clusters of activities. Similarly to Halfaker et al. (2014), they separated reading activity by sessions and took the sequence of pages visited before disconnection – considering a session divided by more than 30 minutes elapsed between two successive activities of a user. They characterized the reading behavior of an article by calculating per month the average of Article Views, Reading Time and Session Articles. Articles exhibited different reading patterns that she named as focus, trending, exploration and passing. Interestingly, they could see that reading patterns responded more to topic interest and informational needs like looking for specific data or learning about a subject than the actual article quality. Focus behavior was mainly defined by time spent reading the article, exploration by using related articles, trending had high popularity in the number of views, and passing was exploring articles but not returning to them. Also, they compared reader and editor preferences by measuring correlation between page-views and length and number of edits. With only a 0.22 (article length) and 0.16 (edits), there was a non-alignment between reader and editor preferences. Two main conclusions were found in the article: reading behavior depends less on the article quality but more on the article topic, and editors' interests are often too specific and not aligned to readers' interests.

4.3 Topical coverage and content interest

Identity is a key aspect of any social media site (Kietzmann et al. 2011). The Wikipedia community provides user and discussion pages where editors can express this sense of self, but as a social network we could interpret topic preference as a way of expressing identity. Specially, when different degrees of activity were found depending on the articles subject (Janette et al. 2014; Kaltenbrunner and Lani-

ado 2012). Even though editors are entitled to represent a neutral point of view in each article, the articles they contribute to or discuss might define them as the sum of their personal interests. The contrast between group interest with the Neutral Point of View (NPOV) policy and the personal has been seen as unavoidable (DiStaso 2012). Furthermore, several behaviors indicate that editors might have identification with the content they write. A study by Janette et al' (2014) showed editors' interests did not equal to readers' interests. That is, editors do not write pending on which topics are popular or require improvement but on personal preferences. Also, depending on the article topic, editors discuss with faster replies, sometimes arguing in the middle of a controversy; they write in talk pages at a different speed leading to very different discussion length (Kalterbrunnen and Laniado 2012). Not to mention that a certain sense of content ownership has been also detected as significant (Halfaker et al. 2009). In a repository like Wikipedia where articles are not directly signed, some editors follow the articles they previously edited in order to protect them from changes they might not accept. Wikipedia requires articles not to be signed directly, but users post in their personal pages the articles they have finished and other similar accomplishments. Wikipedia topical coverage aims at gathering all human knowledge in an encyclopedic way, but it is also the sum of all editors' interests. This is reflected in an 80% of content related to social sciences and culture (Kittur and Bongwon 2009), from history events to pop celebrities. Also, different communities have shown different topical coverage distributions (Hecht and Gergle 2010; Miquel and Rodríguez 2011), with content unique to each of them and with around 20% related to events, geography and culture local to each Wikipedia language. All in all, creating meaning has been considered an attribute of engagement in previous research (O'Brien and Toms 2008). Therefore free topic election can be seen in Wikipedia as a way of developing editors' identity and enhancing engagement. Future research should analyze in detail the variability of engagement based on content.

5. Conclusions and Future Lines

In this paper we presented several studies that revised engagement attributes. On the readers' side, the achieved level of content reliability and constant adaptation to new world events are key to explaining Wikipedia's success. On the editors' side, aspects such as intrinsic motivation with the goal of a free encyclopedia and the sense of gaining autonomy in a community are important. Attributes like 'richness and control', 'positive affect' may be contradictory. For instance, usability and the user interface have often been discussed as a source of frustration. The new visual editor WYSIWG might be helpful in this sense, but

finding new ways of improving communication issues and reducing learning curve would be helpful to increase engagement. We suspect attributes like sensory appeal or aesthetics are not studied due to the type of object.

The editing community has been declining in numbers since 2008. The active editors from older generations are mainly in charge for most of the activity nowadays deployed in the encyclopedia. The problem of not retaining new editors has been explained by the calcification of some rules, difficulty in the use of tools and the frustration caused by reverts that are not well communicated. One study showed that changing slightly the feedback to a new user whose edits were reverted could have a positive significant impact in not decreasing their future activity.

Measuring engagement has been one of the last Wikipedia aspects researched by scholars. While editor metrics can characterize individual and group behavior, reader metrics are relatively unexplored. However, one of the few studies on reading patterns found that content is read differently depending on the topic and this is reflected in the number of page views and the session length. Topic coverage focused on social sciences and popular culture also shows that editors engage differently depending on the content. All in all, this suggests us that a plan on how to meet new editors' interests and improve their welcome with proper communication (Morgan et al. 2013) can be a way to help them overcome the initial learning phase.

References

- Antin, J., & Cheshire, C. 2010. Readers are not free-riders: reading as a form of participation on wikipedia. In *Proceedings of the 2010 CSCW*, 127-130. ACM.
- Attfield, S., Kazai, G., Lalmas, M., & Piwowarski, B. 2011. Towards a science of user engagement (position paper). In *WSDM Workshop on User Modelling for Web Applications*. ACM.
- Butler, B., Joyce, E., & Pike, J. 2008. Don't look now, but we've created a bureaucracy: the nature and roles of policies and rules in wikipedia. In *Proc. of the SIGCHI Conf.* 1101-1110. ACM.
- Chapman, P., Selvarajah, S., & Webster, J. 1999. Engagement in multimedia training systems. In *Proc. of HICSS-32*. IEEE.
- Csikszentmihalyi, M. 1997. *Finding flow: The psychology of engagement with everyday life*. Basic Books.
- DiStaso, W. 2012. Measuring public relations Wikipedia engagement: How bright is the rule. *Public Relations*, 6(2):1-22.
- Dupret, G., & Lalmas, M. 2013. Absence time and user engagement: evaluating ranking functions. In *Proceedings of the Sixth Int. Conf. on Web Search and Data Mining*, 173-182. ACM.
- Giles, J. 2005. Internet Encyclopaedias go head to head. *Nature*, 438 (7070): 900-901.
- Halfaker, A., Geiger, R. S., Morgan, J., & Riedl, J. 2013. The Rise and Decline of an Open Collaboration Community. How Wikipedia's reaction to sudden popularity is causing its decline. *American Behavioral Scientist*.

- Halfaker, A., Kittur, A., & Riedl, J. 2011. Don't bite the newbies: how reverts affect the quantity and quality of Wikipedia work. In *Proceedings of the Seventh International Symposium on Wikis and Open Collaboration*, 163-172. ACM.
- Halfaker, A., Kittur, A., Kraut, R., & Riedl, J. 2009. A jury of your peers: quality, experience and ownership in Wikipedia. In *Proc. of the Fifth International Symposium on Wikis and Open Collaboration*, 15. Orlando, FL, USA. ACM.
- Halfaker, A., Song, B., Stuart, D. A., Kittur, A., & Riedl, J. 2011. NICE: Social translucence through UI intervention. In *Proc. of the Seventh International Symposium on Wikis and Open Collaboration*, 101-104. ACM.
- Halfaker, A., Keyes, O., & Taraborelli, D. 2013. Making peripheral participation legitimate: reader engagement experiments in Wikipedia. In *Proc. of the 2013 Conf. on CSCW*, 849-860. ACM.
- Geiger, R. S., & Halfaker, A. 2013. Using edit sessions to measure participation in Wikipedia. In *Proceedings of the 2013 Conf. on CSCW*, 861-870. ACM.
- Head, A. J., & Eisenberg, B. 2010. How today's college students use Wikipedia for course-related research. *First Monday*, 15(3).
- Hecht, B., & Gergle, D. 2009. Measuring self-focus bias in community-maintained knowledge repositories. In *Proc. of the Fourth Int. Conf. on Communities and Technologies*, 11-20. ACM.
- Kaltenbrunner, A., & Laniado, D. 2012. There is no deadline: time evolution of Wikipedia discussions. In *Proc. of the Eighth Int. Symp. on Wikis and Open Collaboration*, 6. ACM.
- Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. 2011. Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54(3), 241.
- Kittur, A., & Kraut, R. E. 2008. Harnessing the wisdom of crowds in wikipedia: quality through coordination. In *Proc. of the 2008 Conf. on CSCW*, 37-46. ACM.
- Kittur, A., Chi, E. H., & Suh, B. 2009. What's in Wikipedia?: mapping topics and conflict using socially annotated category structure. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, 1509-1512. ACM.
- Lalmas, M., O'Brien, H., & Yom-Tov, E. 2014. *Measuring User Engagement. Synthesis Lectures on Information Concepts, Retrieval, and Services*, 6(4), 1-132. Morgan & Claypool
- Laniado, D., Kaltenbrunner, A., Castillo, C., & Morell, M. F. 2012. Emotions and dialogue in a peer-production community: the case of Wikipedia. In *Proceedings of the Eighth International Symposium on Wikis and Open Collaboration*, 9. ACM.
- Lehmann, J., Müller-Birn C., Laniado D., Lalmas M. and Kaltenbrunner A.. 2014. Reader preferences and behavior on Wikipedia, In *Proc. of the 25th ACM Conference on Hypertext and Social Media*, 88-97. New York. ACM
- Lehmann J., Lalmas M., Yom-Tov E., and Dupret G. 2012. Models of User Engagement. *Springer Berlin Heidelberg*: 164–75
- Lucassen, T., and Schraagen J.M.. 2010. 'Trust in wikipedia: how users trust information from an unknown source.', In *Proc. of the Fourth Workshop on Information Credibility*, 19-26, ACM
- Millen, D. R., & Patterson, J. F. 2002. Stimulating social engagement in a community network. In *Proc. of the 2002 Conf. on CSCW*, 306-313. ACM.
- Miquel Ribé, M., & Rodríguez Hontoria, H. (2011). Cultural configuration of Wikipedia: measuring autoreferentiality in different languages. In *Proceedings of the conference RANLP 2011*.
- Morgan, J. T., Bouterse, S., Walls, H., & Stierch, S. 2013. Tea and sympathy: crafting positive new user experiences on Wikipedia. In *Proc. of the 2013 Conf. on CSCW*, 839-848. ACM.
- Nov, O. 2007. What motivates wikipedians?. *Communications of the ACM*, 50(11): 60-64.
- O'Brien, H. L., & Toms, E. G. 2008. What is User Engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6): 938-955.
- Okoli, C. 2009. A brief review of studies of Wikipedia in peer-reviewed journals. In *Proc. 3rd ICDS*. 155-160. IEEE.
- Okoli, C., Mehdi, M., Mesgari, M., Nielsen, F. Å., & Lanamäki, A. 2014. Wikipedia in the eyes of its beholders: A systematic review of scholarly research on Wikipedia readers and readership." *Journal of the Association for Information Science and Technology*, 65(12): 2381-2403.
- Ortega, F., Gonzalez-Barahona, J. M., & Robles, G. 2008. On the inequality of contributions to Wikipedia. In *Proc. of 41st International Conference on System Sciences*, 304-304, Hawai. IEEE.
- Osborne, M., Petrovic, S., McCreddie, R., Macdonald, C., & Ounis, I. 2012. Bieber no more: First story detection using Twitter and Wikipedia. In *Proceedings of the Workshop on Time-aware Information Access*. TAI (12).
- Panciera, K., Halfaker, A., & Terveen, L. 2009. Wikipedians are born, not made: a study of power editors on Wikipedia. In *Proceedings of the 2009 International Conference on Supporting Group Work*, 51-60. ACM.
- Raitman, R., Augar, N., & Zhou, W. 2005. Employing wikis for online collaboration in the e-learning environment: Case study. In *Proc. of the Information Technology and Applications, International Conference* 142-146. IEEE.
- Rieber, L. P. 1996. Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational technology research and development*, 44(2): 43-58.
- Suh, B., Convertino, G., Chi, E. H., & Pirolli, P. 2009. The singularity is not near: slowing growth of Wikipedia. In *Proc. of the Fifth Int. Symposium on Wikis and Open Collaboration*, 8, ACM.
- Voss, Jakob. Measuring Wikipedia. 2005. In *Proc. Conf. of the Tenth Int. Society for Scientometrics and Informetrics*. Stockholm.
- Yang, H. L., & Lai, Y. 2010. Motivations of Wikipedia Content Contributors. *Computers in Human Behavior*, 26(6), 1377-1383.
- Yom-Tov, E., Lalmas, M., Baeza-Yates, R., Dupret, G., Lehmann, J., & Donmez, P. 2013. Measuring inter-site engagement. In *Proc. of the International Conf. on Big Data*, 228-236. IEEE.
- Zhang, X., & Zhu, F. 2006. Intrinsic motivation of open content contributors: The case of Wikipedia. In *Workshop on Information Systems and Economics*. 1601-1615.
- Zhang, Y. 2006. Wiki means more: Hyperreading in Wikipedia. In *Proc. of the Seventeenth Conf. Hypertext Hypermedia* ACM.

Acknowledgments

I want to thank Mari-Carmen Marcos, David Laniado and Robin Taylor for their advice. Universitat Pompeu Fabra and research center Barcelona Media.