The evolution of online co-production groups and its effects on content quality

Sorin Matei, Nicolas Jullien, Amira Rezgui, Diane Jackson

The evolution of online co-production groups and its effects on content quality

Diane L. Jackson[1], Nicolas Jullien[2], Sorin A. Matei[1], Amira Rezgui[2,3]

1: Purdue University; 2: IMT Atlantique, LEGO-Marsouin; 3: iSchool, Syracuse University

Abstract

What types of social processes generate strong online co-production groups? How do these groups evolve to reach peak performance? How does the quality of the products generated by the groups co-vary with the evolution of their social systems over time? The paper analyzes the entire English and French Wikipedia editorial histories, from their inception until 2015, identifying the specific phases through which two different massive online production systems grew. By tracking the emergence of the high contribution group across two different online spaces on a fine-grained level, the paper uncovers their temporal evolutions and impacts on the organization of social systems. Furthermore, the paper reveals how the quality of the content co-evolves with the emergence of the production groups through each growth phase.

Introduction

What types of social processes generate strong online co-production groups? How do these groups evolve to reach peak performance? How does the quality of the products generated by the groups co-vary with the evolu-

tion of their social systems over time? At the dawn of the social media era, peer-production was seen as a type of egalitarian collaboration (Benkler, 2007; Kelly, 1995; Surowiecki, 2004), a necessary component of successful online groups and a direct precondition for generating content of high quality (Kittur, Lee, & Kraut, 2009). Since then, the emergence of contradictory evidence and the success of non-egalitarian online platforms (Matei, Abu Jabal, & Bertino, 2018) have necessitated deeper analysis of this phenomenon.

These non-egalitarian groups are dominated by a small number of members, whose contributions are both more frequent and more extensive. Effort in many of these groups follows a power law distribution, wherein a few members contribute most of the effort (Huberman, 2001; Matei & Britt, 2017). Such online groups, commanded by a select few members, have proven their capacity to create vast, cohesive social aggregates with little supervision or planning and in spite of the skew in contributions made to them (Johnson, Faraj, & Kudaravalli, 2014).

Olson’s (1971) theoretical proposition that selective incentives maximize collective action might explain this phenomenon within online collaborative projects. Contribution leaders could get an extra dose of motivation from their leadership, which serves as a form of selective incentive. Selective incentive creates a virtuous circle: the greater the contribution, the greater the rewards, which further stimulate contributions. With this framework in mind, it is possible that uneven distribution of effort may be a condition sine qua non for the success of online collaborative projects.

However, there is much to be uncovered about the evolution of non-egalitarian online groups. Returning to Olson's (1971) argument, as groups become larger, the proportion of effort for each individual user should diminish and the rewards for his or her participation should lessen. By this logic, even the most engaged and prolific contributors may theoretically grow discouraged and in time abandon the project, causing collective action to decrease. Although these non-egalitarian online groups do frequently expand, many of them never experience this critical turn. What explains the fact that some online voluntary production groups survive for longer periods of time, where an uneven division of labor prevails over shared participation?

One possibility to consider is the fact that online social systems generate compensatory mechanisms at the macro-level. Research on collective production identified alternative macro-mechanisms that could overcome the diminution of motivation at larger scales (Amrit & Van Hillegersberg, 2010). Some research on online group production showed that in open source software projects, the most prolific users are at the same time "core" producers (Giuri, Rullani, & Torrisi, 2008; Rullani & Haefliger, 2013). They do more than add additional content. They shape the project and, by extension, provide structure and direction for future contributions. Contributing to the strategic direction becomes a type of incentive, which is not sensitive to the relative proportion of content produced by each individual member, and therefore, becomes more important than the

proportion of their contribution relative to the entire production process (Borzillo, Aznar, & Schmitt, 2011).

In other words, the amount of work matters not only by sheer mass, but by impact. The higher output of some participants increases their responsibilities and their role in animating the teams or in energizing new sub-projects (Borzillo et al., 2011). In time, the strategic leadership group also establishes stable interactions and positive feedback within itself, which serves as additional work incentives (Matei & Britt, 2017). Therefore, a virtuous cycle of intrinsic rewards and peer support is reinforced and the uneven distributions of contributions are, in turn, accentuated.

These processes may explain the mechanism of social differentiation and ultimately group resilience. However, the way in which these mechanisms emerge has yet to be elucidated. In conjunction, the explanations mentioned above do not provide sufficient insight to determine if there are some constants in the relative proportion between core members and other members, or if such proportion may vary in time (Matei & Britt, 2017). Determining these ratios over time and identifying the discrete phases in the growth of online communities and in the growth of the elite groups of these online communities are still works in progress (Matei et al., 2018).

The present paper aims to contribute to this line of research. Specifically, the present study aims to fill in gaps in research by identifying the specific phases through which two different massive online production

systems grow. By tracking the emergence of the high contribution group across two different online spaces on a fine-grained level, we aim to uncover their temporal evolutions and impacts on the organization of social systems. Furthermore, we will look at how the quality of the content co-evolves with the emergence of the production groups through each growth phase. Wikipedia, the site of investigation, is one of the most prolific peer-production production sites in the world, as the sixth most visited site in the world (Alexa.com) and the default source of information for many Google searches, even for common nouns, such as “sun,” “car,” or “man.”

We will analyze and compare two Wikipedia projects, both of which are large and global in scope and anchored in a specific geo-cultural space: English, anchored in the North Atlantic; and French, anchored in the West European. This will provide the context needed to understand how macro-structural factors across socio-cultural boundaries may impact the growth phases and the relationship between structural factors and quality.

These questions are not of mere theoretical importance. Connecting the evolution of high-production groups and quality across several macro-contexts may help explain the overall success of online collaborative production groups generally. Most importantly, a significant point of the present study is that while the global quality of the products generated by these groups varies, the mere fact that this quality is mostly usable and occasionally indistinguishable from that of content produced through a far more structured process (Jullien, 2012) demands a novel explanation.

Literature Review

The present study examines how contribution unevenness evolves and how inequality translates into discrete impacts of the quality of content at each phase of evolution. We consider how the impact of uneven contributions on content vary across two significant global knowledge construction projects. In this section, we propose several explanatory mechanisms that support an evolutionary approach to understanding the fluctuation of quality in wiki projects and we use these mechanisms to generate research questions that directed our empirical study.

The first theoretical concern of the present paper is the self-selection, out of which process groups of highly productive members emerge. Members of these groups invest significant amounts of time and energy in voluntary projects (Matei & Britt, 2017). For this reason, understanding the choices that lead to the emergence of an “elite of active members” demands that we uncover the discrete phases in growth of any online production system as well as the trade-offs between investment and rewards within these phases.

Knowledge Production Projects, Phases, and Critical Mass.

In the growth of Free, Libre, Open Source projects (FLOS projects), among which we should include wiki projects, Kane, Johnson, & Majchrzak (2014) identify three distinct phases: inception, coalescence, and

stabilization. For each phase, there is a core process: commitment, innovation, and maintenance. For each process, there is specific type of interaction: non-directed contributions for the inception period, joint shaping for the coalescence phase, and defensive filtering for the stabilization phase. During phase one, inception, members self-select their level of commitment, adding content in a non-directive, uncoordinated manner. After a period of wild and wide variation in contribution, some members start to progressively produce more. During coalescence and once the commitment of members is obtained, the rate of innovation increases and committed members focus on adding new content and features, while contributions become more coordinated. Socio-technical mechanisms help the members see the impact of the work, which leads to “joint shaping,” or the tacit process of collaboration by that considers what has been produced before. This ensures growth and creates the premises for the transition to the third phase. Maintenance practices such defensive filtering emerge in this final stage, by which new content is retained selectively by the members that have contributed the most to the process of joint shaping. Defensive filtering aims to prevent a decline in coordination by keeping the contributions focused on a core direction and by communicating to the members who generated the most that their work matters. Thus, in the third phase, the activity focuses on preservation of key contributors by stimulating coordination among them.

This three-phase model is well-known in the software development and collaborative group action literature not only for these specific phases, but

also for its unique S-shape. The curve exhibits two distinct slower growth slopes: one at the beginning and one at the end of the process, which are bridged by a phase of rapid growth in the middle. The S-curve is also known in software development as the Putnam-Rayleigh-Norden curve or model (Pressman, 2005), describing the timing of software production and delivery. These three phases map in software development processes onto the investment, feature enrichment, and gradual improvement activities.

As an open source project, Wikipedia fits well within this model of development. Matei and Britt (2017) have shown that over the first ten years of its existence (2001-2010), the Wikpedian collaboration process followed the three-phase approach very closely. The three-phase model can be further applied to collective action projects fueled by voluntary contributions (Marwell & Oliver, 2007). However, a phased approach to studying the emergence of collaborative projects demands an empirical method to specify the nature of the interaction dynamics within each phase.

Engagement, Core Members, and Phases.

Thus far, we have proposed three phases in the development of online production systems, known as inception, coalescence and stabilization, each of which are characterized by a core process: commitment, innovation, and maintenance. Let us look at the interaction mechanisms specific to each of them.

In the inception/commitment phase, individuals join a project at various levels of interest. Each level is moderated by the context of their work and by their own estimation of the effort they intend to invest into the project. Their effort will vary with the perceived reward generated by their work. When engagement costs are lower, accomplishment is reached sooner, thereby enhancing the perception of contribution as being more valuable. By this, a dynamic of involvement emerges (Marwell & Oliver, 2007; Oliver, Marwell, & Teixeira, 1985), which favors those with lower engagement costs. Engagement costs are lower than other costs due to a variety of reasons, as Marwell and Oliver (2007) emphasize, including more and better skills and knowledge, higher production abilities, more free time, or opportunity costs to engage in other activities. In essence, commitment is the product of an initial cost-benefits analysis, which stimulate those that have more to get out of the project. *We expect a relative increase in unevenness during this first phase, although contributions might be chaotic and non-directed and commitment varying from period to period.*

The coalescence/innovation phase of the project is characterized by the project generating a certain amount of usable product and structuring itself to diminish the cost of contribution. Many other participants that possess different abilities and knowledge sets join the project and the newcomers bring innovative ideas and skills. Given the growing number of competing ideas during this phase, cooperation for innovation needs to be carefully calibrated. Safner (2016) building on Hess and Ostrom (Hess, 2011),

proposed that cooperation between contributors of various abilities and rates of production is enforced through rules, norms, and socio-technical organization that result in what Kane et al. (2014) called the “joint shaping” of products. Joint shaping regulates the innovative process through a quasi-administration process which assigns de facto rights and obligations to the members that aim to ensure for each significant voice time continuity. *Joint shaping leads to motivation of many new members, which leads to an increase in contribution evenness. Simultaneously, we may also notice an increase in the likelihood of those members that are more productive to “hang out” longer in the contribution elite. Thus, there could be a positive association between contribution evenness and “stickiness,” or elite temporal stability (likelihood to be a member of the contribution elite).*

When the project reaches a level of saturation by inclusion of many possible new types of members and content, joint shaping is complemented and subsequently superseded by a new implicit and emergent rule, known as “defensive filtering” (Kane et al., 2014), thereby opening the stabilization/maintenance phase. Most new content is accepted in the inception (commitment) phases and jointly shaped in the coalescence (innovation) phase, while the third phase contributions are filtered in order to maintain the achievements gained thus far. By this, the collaborative system enters a “maintenance mode,” where the collaborative organization takes the shape of a quasi-bureaucracy (Butler, Joyce, & Pike, 2008) or adhocracy (Matei & Britt, 2017). The hallmark of this period is a certain stability in

groups processes, which is not on the surface inherently negative. The controls put in place by the project veterans to filter newcomers is justified by the fact that the accretion of new content can damage the quality of the content. *In this phase, we expect the evenness of the contribution process to stabilize and decline, even while elite temporal stability, or "stickiness," increases.*

Thus, the success of self-regulating, uneven communities depends both on their capacity to initially attract new members and to subsequently filter content while maintaining a level of optimal unevenness. Group stability rests on emerging rules, codes of conduct, or regulative cooperative behavior (Butler et al., 2008; Safner, 2016), which should be enforced by people who are to be selected from the old-timers already on the project. These rules and organizational arrangements necessitate time and deliberation, which impose an energy and time investment that further favors those who invest in learning them. Moreover, as these over-users become adept at a task, they begin using the rules as means of protecting their own production, and thus serve to continue stimulating their own contributions. Mastering the administrative process becomes a type of selective incentive which works in tandem with accumulating work capital to further cultivate uneven contributions across the board (Jullien, Roudaut, & Squin, 2011).

To simplify the narrative outlined thus far, we propose that because of macro and micro motivational constraints, peer-production projects grow

through three phases of production. In the inception phase, contributions are diverse and chaotic. Soon, however, some contributors emerge as highly productive members who engage in rapid production of massive amounts of content. In consequence and as a mark of entering into the second phase, the project coalesces through “joint shaping” and becomes mature enough to attract a new wave of more diverse contributors. At this stage, the old-timers become the “core members” or “sticky elites” (Matei, Jabal, & Bertino, 2017). These members are contribution elites that tend to be stable in time, particularly as the distribution/heterogeneity of contribution effort level increases, and that tend to work together in the creation of content. In the third phase of stabilization and maintenance, sticky elites turn into enforcers and content filters, due to the lack of need for new contributions as the project enters a process of defensive filtering (Kane et al., 2014).

In our study, mapping the discrete phases of the co-production project by observing the distribution of contribution across members and of the evolution of the core contributor group is only a means toward an end. While useful in itself, this mapping exercise should be complemented by a more in-depth look into how changes in the interactional dynamics affect the outcomes of the collaboration process—namely, the quality of the content. How do changes in unevenness of contributions and elite stickiness co-evolve with the quality of the products generated through co-production? This is a core question and major goal of our paper. To elucidate it, we

consider the evolution of production quality, as discussed in previous research.

Research questions

Starting on the production side, we need to measure the emergence of a core or “elite” group of contributors in terms of its temporal stickiness and in terms of its share in total knowledge production. Formal research questions that guided our study are:

RQ1a: Are there discrete phases in the evolution of contribution evenness in wiki projects?

RQ1b: Are there discrete phases in the evolution of elite stickiness?

RQ1c: How do the phases differ across co-production wiki systems?

Following the line of reasoning that the phases move from non-direct contributions to joint shaping and defensive filtering, we expect that unevenness and stickiness should behave differentially within each phase. For example, in phase one, both elite stickiness and unevenness should start low and slowly increase. In phase two, characterized by fast, strong, and effervescent contributions, unevenness should go up and reflect the increasing involvement of certain users in the production system. Yet, stickiness should partially decline, as the role of the elite members changes from editing to editing supervision and from contributor to policy-makers, with an undefined impact on their share of the total contribution (as there are a lot of other

contributors, but not all of them succeed). The following questions explore these issues:

RQ2a: Does the relationship between contribution unevenness and elite stickiness behave differentially across phases?

RQ2b: How does the relationship between contribution unevenness and elite stickiness behave differentially across co-production system scales?

As mentioned above, uncovering phases in structural evolution is a means toward the end of better understanding the production of quality content. In consequence, the relationship between unevenness and stickiness should be projected against an outcome variable, which can tangibly capture their impact. The final four research questions provide guidance in the exploration process regarding the production of quality content.

RQ3a: Does elite stickiness impact content quality in the various phases of wiki development?

RQ3b: Does contribution unevenness impact content quality in the various phases of wiki development?

RQ3c: How does the impact of stickiness on content quality vary across co-production system scales?

RQ3d: How does the impact of contribution unevenness on content quality vary across co-production system scales?

We will explore this phased approach and its impact on the knowledge co-production of two language Wikipedia projects. After having explained why we have chosen Wikipedia and these two languages, we will describe

our operationalization of the following variables: the users' involvements, their effort distributions, the production elite time resilience (i.e., stickiness: the quality of the knowledge generated by the wiki projects), and finally, the relationship between all of these variables.

Methods

Choosing Wikipedia

As previously indicated in the literature review, Wikipedia is a particularly interesting project for studying the manner in which online groups that are dominated by several voices evolve, self-regulate, and continue producing even as some members grow more motivated through increased contribution.

We have the record of all editorial interactions that have transpired on Wikipedia, totaling a decade and a half of information. Published research indicates that these interactions are highly, unevenly distributed across collaborators in the English Wikipedia (Matei et al., 2018) and that the project has reached a "maturation phase," where inequality has stabilized. We also know that the phases come with different types of interaction patterns. Matei and Britt have shown that during the period of inception, inequality is lower but also varies widely (Matei & Britt, 2017). Over time, the amount of effort needed to add new content increases as new edits are more likely to be rejected, making the work less (Aaltonen & Seiler, 2015; Ransbotham & Kane, 2011).

We should thus expect that the online organization will start with one type of interactional arrangement and transition to one or more subsequent arrangements that consider the variable costs of contribution and barriers to entry. During the maturation phase, for example, inequality and the tendency of the top contributors to be in the top contribution group stabilizes. At this stage and as the organization becomes more bureaucratic (Butler et al., 2008), contributions increasingly focus on improving the existing articles, rather than creating new ones (Safner, 2016).

While English Wikipedia has been studied in depth, we do not have a good comparative investigation of the issues mentioned above, especially of structural evolution and its effect on content. Are the processes gleaned from the previous literature, especially phased development, specific to English Wikipedia or do they occur across wiki co-production systems? We propose a comparative investigation of evolutionary phases and the impact of structural factors on quality by looking at another major Wikipedia project—namely, that in the French language. French Wikipedia is among the largest Wikipedia projects. It has both a local (French) and global editorial base, including contributors from Europe (especially Belgium), Africa, and the Middle East. The distribution of effort across local and global spaces make it a good comparison site to English Wikipedia, which is similarly locally (in the US, Canada, and UK) and globally located (including contributions from India, South East Asia, and generally from most English educated elites in the world).

Datasets

The questions were explored using the Wikipedia database "dump files." These contain precise information about all actions performed in any Wikipedia language. Dump files are available from Wikimedia Foundation and can be retrieved from the Wikimedia Downloads center[1]. We downloaded data from English and French Wikipedia. For each language, we first retrieved Wikipedia XML database dump file "pages-meta-history.xml.7z" from the set of available dump files. Since this is part of a longer time project initiated several years ago, we used data extracted on December 12, 2015. The data contains the complete metadata of every version of all articles from the beginning of the online encyclopedia (January 2001) to December 2015.

Due to size and complexity, English Wikipedia data was processed slightly different from French Wikipedia. These data were extracted using a parallel computing system, which computed a weekly value of contribution for each registered user. We evaluated the amount of user contribution by considering the number of characters added (abbreviated as *A*), deleted (abbreviated as *D*), or modified (abbreviated as *M*) compared with the preceding edit (abbreviated as *v*). The number of modified characters is calculated using the edit distance (Adler, De Alfaro, & Raman, 2008) to measure the total amount of relative change in text position and structure. As a result, the user contribution is formally defined through the contribution delta formula:

$$d(u, v) = \max(A, D) - 0.5 * \min(A, D) + M$$

1 https://dumps.wikimedia.org/

For French Wikipedia, we used WikiDAT for data extraction. WikiDAT is a tool for Wikipedia data analytics, based on Python and R and using MySQL database. It aims to create an extensible toolkit for Wikipedia using Python and R to automate the extraction of Wikipedia data into 5 different tables of the MySQL database (page, people, revision, revision hash, logging). In this study, we were interested in the following tables:

- Table page stores information about all pages in a Wikipedia language.
- Table revision contains metadata about all revisions performed in a Wikipedia language.
- Table people lists the numerical identifier and nickname of all registered users in a Wikipedia language.

Once tabulated, user-specific contribution scores were calculated in a manner similar to the one described for English Wikipedia.

Variables

The variables used to measure the evolution of each Wikipedia project capture structural and quality changes. Some of them are purely descriptive (number of edits), while some are the result of data transformation. The directly descriptive variables include number of edits and elite stickiness. Calculated variables include our core social structuration variable, entropy, and quality. Entropy is a measure of content distribution evenness across participant members. It reaches zero when one member contributes all the content and a maximum value, which increases with group size, when contributions

are perfectly equal. It was calculated using only registered users, which generate more than half the cotent on both projects. Data was summarized at the monthly level to provide sufficient time to compensate for contribution volatility from day to day.

Variables	Description	In this paper
Contribution amount	Amount of contribution	Amount of contribution takes into account, as mentioned above, the nature of each edit, additions counting more than deletions or changes.
Elite stickiness	Percent users that are “sticky” from period to period	Each period, we compared the composition of the most productive 1% user group against the similar group from the previous month. Stickiness represents the percent of top 1% individuals in the previous month were still in the top 1% during the current month.
Contribution	Shannon’s entropy	We use entropy to evaluate

entropy	index $H(X) = -\sum_{i=1}^{m} p(x) \log_2 p(x)\,.$	the dispersion of the contributors according to their amount of contribution per unit of time. Entropy takes a value of 0 when there is absolute order in the system (one element is prevalent at the expense of all others) and a maximum value (which varies from system to system), when there is perfect disorder and diversity (all elements are equally present). Entropy is a synthetic measure that tells us at a glance how well represented are the different components of a social or communicative space.

Normalized entropy	Normalized entropy enables us to compare the evenness of two communication spaces, including over time, by controlling for the number of elements that compose each of them. Normalization can be obtained by dividing the raw entropy score by its maximum log(m), which limits its range from 0 to 1 $H_o = \frac{H}{H_{max}}, \quad 0 \leq H_o \leq 1, \text{where } H_{max} = \log_2 m.$	As the number of contributors is not stable (it is growing) we have to normalize the entropy to be able to make comparison between the distributions from one period to another
Article quality	Formal quality of information	Information quality is inferred from formal article features (such as information richness,

		length, and sourcing) and measured via a machine learning algorithm that uses articles evaluated as high quality as ground truth. Description follows.

Quality

Content quality values at article and period (month) level are obtained from the Wikimedia Foundation ORES service, which releases an inferred level of quality for each article (citation). The dataset predicts the quality of each article created on Wikipedia since 2001 at a monthly level. Objective quality features, such as length, the number of references, the number of headings, information richness, and the number of functional links, are used to predict the quality of all articles. The validity of the features and their relative importance were trained via machine learning. The machine learning algorithm was validated on a sample of high-quality (featured) articles (Halfaker, Sarabadani, & Taraborelli, 2016). Featured status was assigned by human editors. Quality values go up to 5. The highest score indicates the best-quality article.

Figure 1 Quality evolution on English Wikipedia

Figure 2 Quality evolution on French Wikipedia

Final Datasets

Two final datasets were constructed, one for French and the other for English Wikipedia. Each dataset includes periodic data points for the variables of interests (i.e., entropy, stickiness, and quality). Each data point was assigned a sequential number by period number, which will be reflected in the time (x) variable in all charts.

Analysis and Results

Data was analyzed using three procedures, one for each clusters of questions: breakpoint detection, within-phase correlation, and regression analysis. The first two analyses and question clusters were exploratory, while the last one was inferential.

RQ1a: Are there discrete phases in the evolution of contribution evenness across wiki projects?

Yes, there are. Specifically, we identified three phases in each project, each falling within similar temporal boundaries.

To identify the points where collaboration evenness changes rates of increase/decrease significantly. The method was a simple second derivative method. The inflection points were detected where the function curve changes concavity (R code and documentation available in Appendix A). We

used a window of 150 periods to find at what points the rate of increase or decrease in entropy changed at a faster or lower rate than in the previous 150 periods. Overall, entropy starts low, indicating the dominance of a smaller group of people. This increases and decreases dramatically during the first phase, only to reach a minimum, marking the end of the period. The rest of the evolution follows an “increase/decrease/stabilization” trajectory. The analysis showed for English Wikipedia, three change ranges with midpoints during weeks 107 (month 26), 240 (month 60), and 594 (month 148). Visual inspection indicates that the last breakpoint captures a change in entropy that almost negligible, so we chose to ignore it. Thus, we obtained the expected three phase segmentation, corresponding with the three conceptual periods described in the literature review. The two change points correspond with dramatic changes in entropy, with the second one representing the “take off point” for article production.

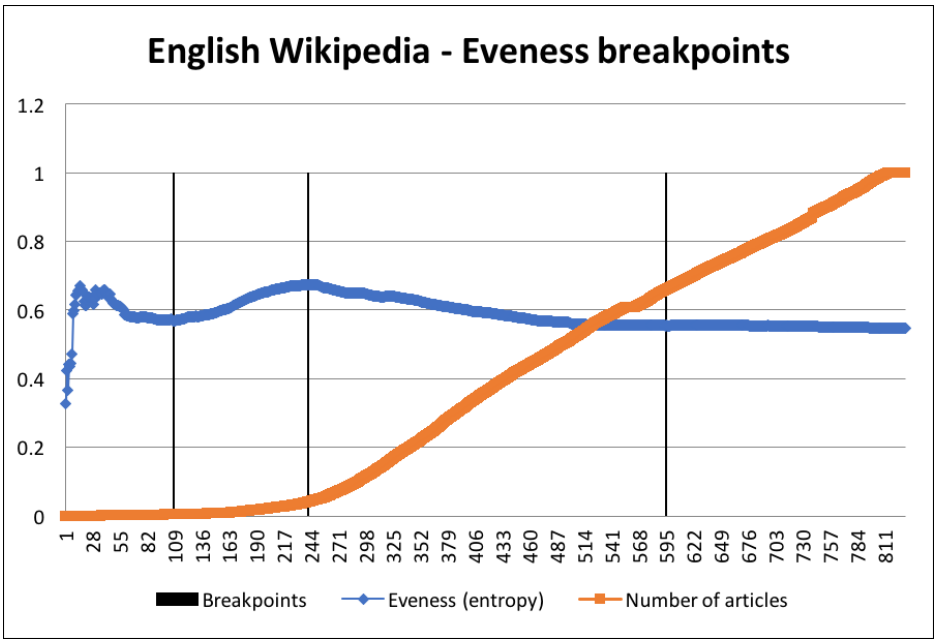

Figure 3 English Wikipedia Growth Phases

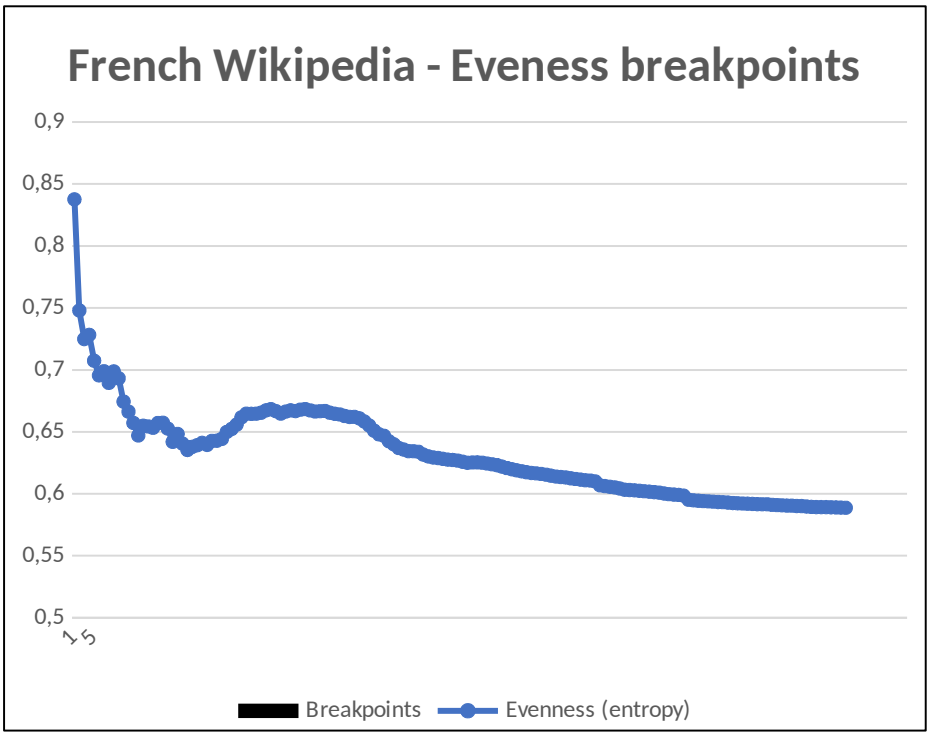

Figure 4 French Wikipedia Growth Phases

The evolution of evenness on French Wikipedia is characterized by a trajectory very similar to that identified on English Wikipedia. In purely descriptive terms, there is an initial dip, followed by a maximum and then by a long, flat, and slightly declining trend. Breakpoint analysis also identified two core break points: one during month 24 and the other during month 47. Both breakpoints were also similar in location to those identified on English Wikipedia: at the 2- and 4-year marks, respectively. More importantly, the same tripartite division of collaborative evenness into discrete phases existed. For the remainder of the paper, we will use the two pairs of breakpoints for segmenting the two collaborative processes and for analyzing the relationship between stickiness, evenness, and quality within each stage.

RQ1b: Are there discrete phases in the evolution of elite stickiness?

No, there are no significant breakpoints in stickiness for either project.

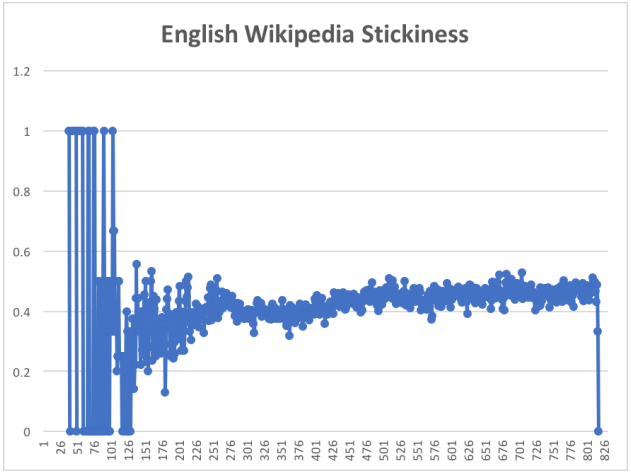

Figure English Wikipedia Stickiness

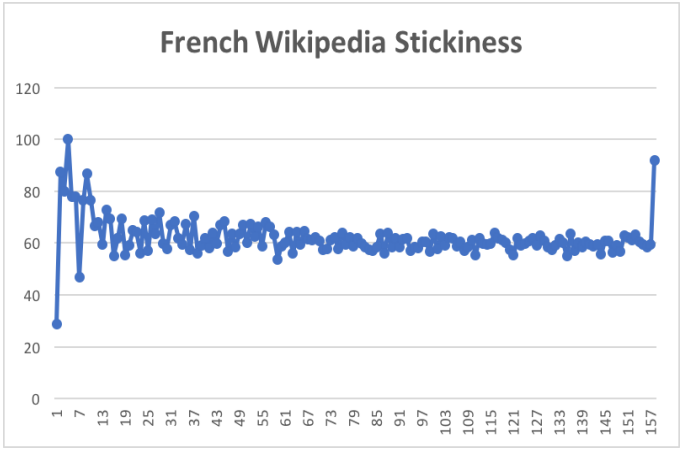

Figure 6 French Wikipedia Stickiness

Breakpoint analysis did not indicate any major discrete breakpoints that would divide the time series of stickiness for either Wikipedia in major phases. Although breakpoints were identified, these were very numerous and defined by the irregular, or noisy, and flat nature of stickiness across periods. The group of big contributors, even in big Wikipedias, is relatively small, and even small changes in activity visibly affects the variable.

Beyond this noise, stickiness remains remarkably stable for both languages after a starting period. The only notable difference was that the average period to period stickiness was substantially higher for French Wikipedia. Sixty percent of the top 1% contributors on French Wikipedia were among the top 1% on a month-to-month basis. By contrast, although not as low as expected, the top English Wikipedia editors stuck at the top of the contribution elite in proportion of 30-40%.

RQ1c: How do the phases differ across various co-production system scales?

The two differences that were noted include: first, entropy started higher on French Wikipedia and lower on English Wikipedia; and second, that stickiness decreased on French Wikipedia, while it increased on English Wikipedia. Yet, overall, the trajectory of the entropy phases was very similar.

Although entropy presented the same evolution overall, with a middle "hump" and long tail that indicated an increase in evenness during the first phase of life (years 1-4), there was a noticeable difference in the starting points. While English Wikipedia was started by a small group of individuals

who monopolized the contributions and who became more numerous as their contributions grew relatively more diverse over time, the French Wikipedia started with a group of individuals who shared significantly more effort. Although we are talking about a very small number of people (several dozen for the first year), it is important to note the distinction as the two projects converged on the same trend of decrease in entropy over the last few years of the project's existence. No matter where the two projects started, the overall trajectory, decline in entropy, and increase in uneven distribution of effort was shared.

The second difference between datasets can be identified in terms of stickiness. English Wikipedia presented more volatility and steadier growth in stickiness after the first two years of life. After the first two-year period of intense volatility, stickiness increased from an average of 30% during years 2-4 to an average of 45% during the last two years. In contrast, French Wikipedia presented a slow, but noticeable drift toward lower stickiness. The values went from an average of 70% during the first two years to an average of 50% during the last two years.

The first research questions explored the presence and differences in the development phases of the two projects. The next few questions explored the relationship between unevenness and stickiness and its impact on content quality. The relationship between evenness and stickiness was meant to determine whether an increase in unevenness also means a true increase in elite emergence over time. This can be captured by a positive

relationship between unevenness and stickiness. Additionally, we are interested in whether trends in the relationship between unevenness and stickiness at phase level cut across projects (English vs. French). Thus, these two questions were explored next:

RQ2a: Does the relationship between contribution unevenness and elite stickiness behave differentially across phases in each project?

Generally, stickiness and evenness go hand in hand across periods when they are significantly correlated. There is no significant relationship reversal across periods. However, the two processes are not correlated during the first phase on English Wikipedia and during the second phase on French Wikipedia.

RQ2b: How does the relationship between contribution unevenness and elite stickiness behave differentially across co-production system?

The relationship between the two variables was remarkably similar across projects.

Figure English Wikipedia Phase 2

Figure English Wikipedia Phase 3

Correlation analysis between unevenness and stickiness within each of the three phases for both groups indicated two major differences. First, the first phase of non-directed contributions indicated no relationship between stickiness and unevenness on English Wikipedia, as expected. In other words, the processes during the inception (commitment) phase were rather chaotic. Variations in unevenness, which tended to decrease, did not translate in variations in stickiness. On French Wikipedia, there was a strong positive correlation (r=.68, p<.01), as both evenness and stickiness declined in parallel. In other words, the process was collaborative and distributed, where contributions were even, so contributors tended to linger less in the contribution

Figure French Wikipedia Phase 1

Figure French Wikipedia Phase 3

elite.

In phase two, however, there was a differential relationship between evenness (entropy) and stickiness. On English Wikipedia, entropy and stickiness were positively correlated (r=.34, p<.01). Even when production became marginally more distributed, some individuals tended to become more stable contributors. On French Wikipedia, however, there was no significant correction between the two. In other words, the coalescence phase (2) of French Wikipedia indicated that members rose to the top of the pyramid in-

dependently regardless of whether the general trend in the community favored evenness.

The most important trend for both Wikipedia spaces, however, was in Phase 3, of maintenance and defensive filtering. In both collaborative groups, evenness and stickiness were positively correlated *(r*=.32, *p*<.01 for French, and *r*=.4, *p*<.01 for English Wikipedia). This goes against the theoretical assumption that in this third phase, work would be less even and the members would simultaneously be stickier. We will explore this finding in more detail in the discussion section.

RQ3a: Does elite stickiness impact content quality in the various phases of wiki development?

Stickiness is a significant predictor only in 2 of the 6 phases (3 for each project), which indicates a rather weak explanatory variable.

RQ3b: Does contribution unevenness impact content quality in the various phases of wiki development?

Unevenness is a stronger predictor than stickiness, being detected in 5 out of the 6 phases (3 for each project). Furthermore, it is more likely to be negatively associated with quality in the later phase (3rd).

RQ3c: How does the impact of stickiness on content quality vary across collaborative projects?

Stickiness matters a significant amount in English Wikipedia and not at all in French Wikipedia.

RQ3d: How does the impact of contribution unevenness on content quality vary across collaborative scale?

Unevenness varies quite a bit across projects. Although it is negatively associated with quality in the third phase across projects, its effects in the other phases are mixed. This demands a differential interpretation, which speaks about differential mechanisms in the two projects (non-directed contributions, joint shaping, or defensive filtering) and their corresponding phases, which we will discuss in the next section.

To answer these questions, we regressed quality on evenness (entropy) and stickiness within each of the three phases. The goal was to detect the relationship between evenness (entropy) or stickiness and quality net of each other and to track these relationships across phases. In tune with our proposed research framework, we expected that at each phase the relationships would be different, where they were sometimes converging and other times diverging. The results presented in the summary table and the two charts below show that the relationship between entropy (evenness) and quality changes from phase to phase. The tables and the charts show the beta values for the effect of entropy or stickiness on quality. Positive values indicate that as entropy or stickiness increase, quality increased as well. Negative values indicate that increases in stickiness or entropy led to a decrease in quality.

	French Unevenness	English Unevenness	French Stickiness	English Stickiness

	(Entropy)	(Entropy)		
Phase 1	**0.443***	**-0.964****	0.256	0.002
Phase 2	-0.36	**0.98**	-0.219	**0.038****
Phase 3	**-0.893****	**-0.855****	0.014	**0.072****

Table 1. Beta values for entropy and stickiness effects on quality. *p<.05, **p<.01

Specifically, the result for entropy indicates that at each phase and on each Wikipedia, the relationship between entropy and quality changed. For English Wikipedia, entropy was negatively related to quality in phases 1 and 3. In other words, in these periods, a decline in entropy and corresponding concentration of effort led to higher quality. In the second phase, however, quality increased as contributions became marginally more even. This hints at the fact that the emergence and stabilization phases in the life-cycle of a project demand control, while the period of fast growth demands a relative loosening of the reins.

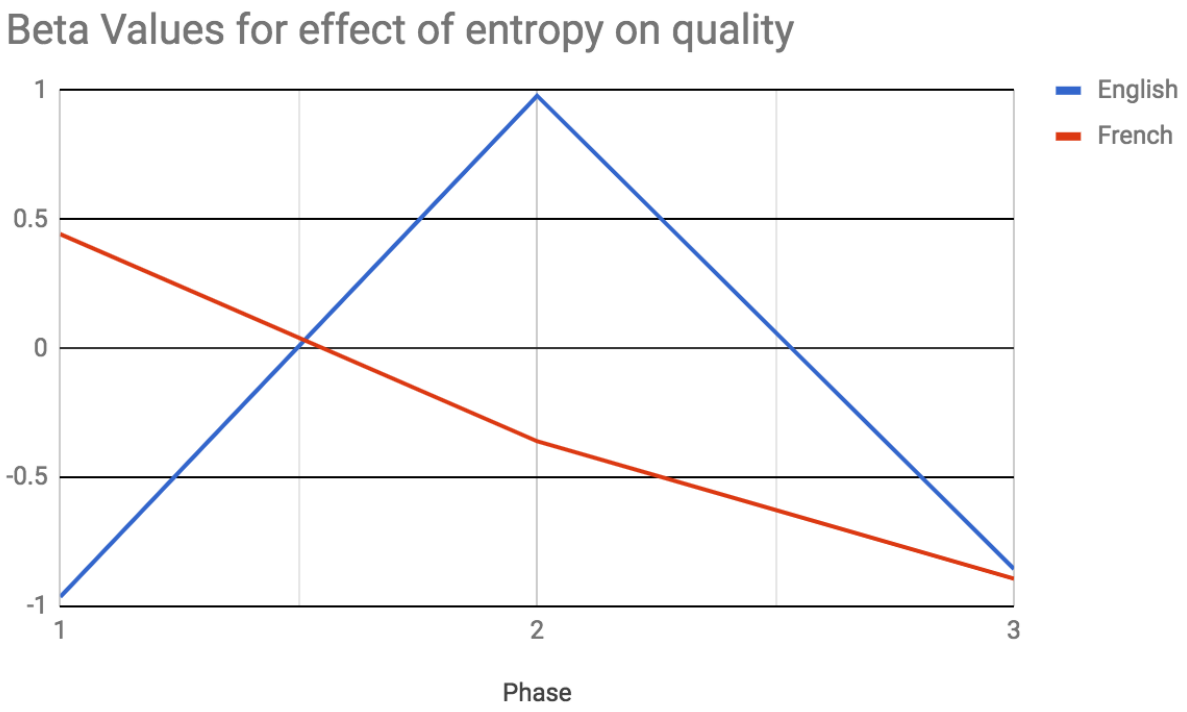

Figure 5. Beta Values for the effect of evenness (entropy) and Quality.

French Wikipedia also presented a linearly-differential story for each phase. First, although nominally higher entropy values are associated with higher quality in the first phase, the relationship did not present statistical significance. We kept the data point in chart 7 to present a fuller picture.

In both Phases 2 and 3, entropy declined as quality increased, but the process became stronger in Phase 3. In other words, on French Wikipedia, we noticed a process of deepening concentration of effort and of higher impact of such concentration of effort on quality, which thereby increased with the deeper concentration and higher impact.

As far as stickiness is concerned, stickiness did not have any significant effect on quality in any phase on French Wikipedia. However, it did have an

increasing effect on English Wikipedia in Phases 2 and 3, where higher levels of stickiness led to higher levels of quality. The differential effect was not in terms of reversal of direction, but of increases in the strength of the relationship. Furthermore, as we can observe, the increases in Phase 2 effects were much above the small and insignificant effect for Phase 1. What we have here is a significant linear increase in the effect of stickiness on quality.

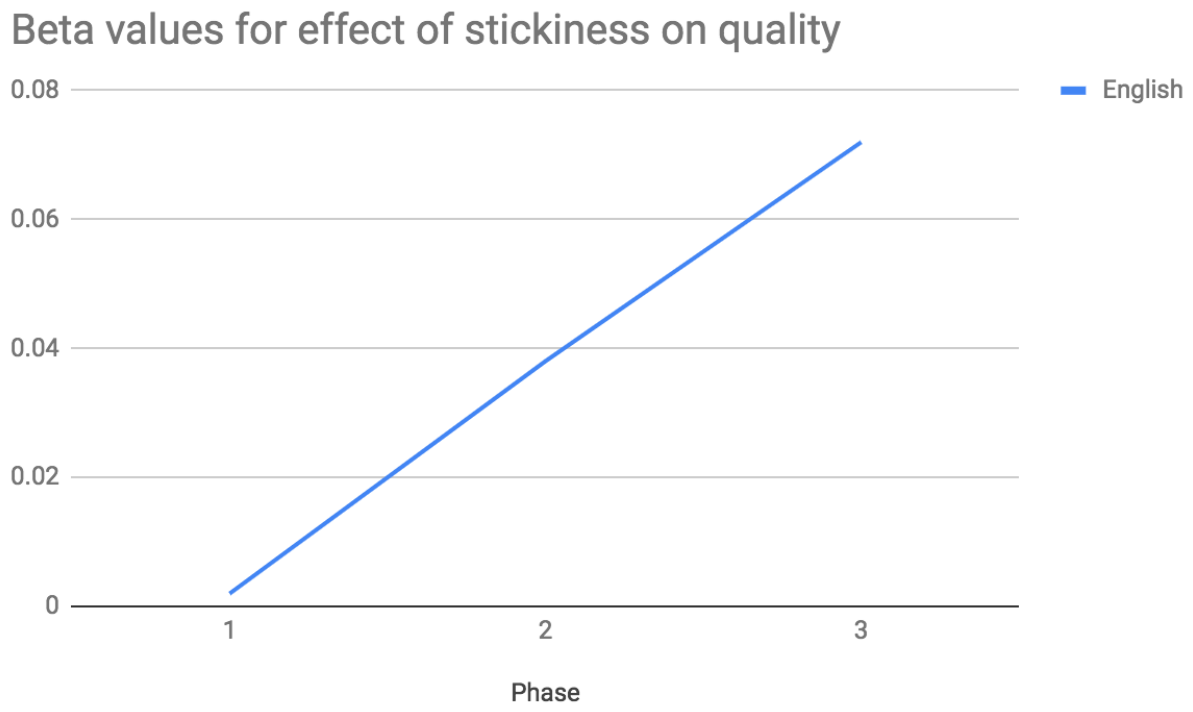

Figure 6 Beta values for the effect of stickiness on quality for English Wikipedia, only. Phase 1 data point is not significant.

Discussion

Our study aimed to detect the effect of structural factors in co-production Wikipedia projects, such as distribution of effort and elite stickiness, on

quality of content. We compared two major Wikipedia projects to detect trends in effects on quality of content that transcend local cultural settings. We proposed a tripartite evolutionary model for the effect of evenness of collaboration and contribution elite time stability (stickiness of top 1% contributors in the 1% group from period to period).

The research questions aimed to address three issues: RQ1) Are there detectable phases across collaborative projects? RQ2) Are there macro-structural relationships within phases and across projects? RQ3) Do macro-structural processes impact quality and do they behave differently across phases and projects? The first two were answered in the affirmative and results are presented in the analysis section. Overall, the conclusion is that co-production sites evolve along a rather traditional path, which includes one period of inception, one of coalescence, and one of the maintenance. Furthermore, we discovered that processes of social growth and the emergence of production elites, which dominated the production on a month to month basis in proportion of 30 – 40%. Furthermore, we noticed that the increase in elite resilience (stickiness) may have been associated with relative dispersion in evenness of contributions. Additionally, we detected the emergence of “adhocracies,” (Matei & Britt, 2017), which thrive by punctual self-organization of sticky elites that emerge from groups that could be quite entropic and diverse.

However, the core issue of our paper is the effect of structural factors (evenness and stickiness) on content quality. Do these processes have any

detectable effect on quality, and if yes, does the impact follow the theoretical model that predicts non-directed contributions during phase one, joint shaping during phase two, and defensive filtering during Phase 3? The review of the literature on the macro-structural factors that may impact quality suggested that during each phase (inception, coalescence, and stabilization), there might be different contribution mechanisms (non-directed contributions, joint shaping, and defensive filtering). We surmised that a relative increase in unevenness would occur in Phase 1 (inception), with the exception of contributions that might be chaotic and commitment varying wildly from period to period (month to month). We also expected that unevenness, or entropy, would impact quality in a non-directed manner. The regression analysis performed to answer research question 3c (Does contribution unevenness impact content quality across wiki projects?) showed that there was variability across the two projects (French and English Wikipedia). While unevenness increased quality on English Wikipedia, unevenness decreased quality the French Wikipedia. In other words, at the beginning of the project, the French Wikipedia distribution of effort created better content, while the reverse was true for the English Wikipedia. More succinctly, a better broth is made by more French cooks, but the same product is yielded by fewer English-speaking ones. This split path suggests that inception phases can be unpredictable. Cultural and social factors might be responsible for these phenomena, such as a more open and contention

cultural style of collaboration in France and a more hierarchical tendency among the mostly male English editors.

On the other hand, we expected an increase in evenness, quality, and stickiness during Phase 2 of joint shaping. Joint shaping leads to motivation of many new members, which leads to an increase in contribution evenness. At the same time, we noticed an increase in the likelihood of those members that were more productive to “hang out” longer in the contribution elite. Thus, there may be a positive association between contribution evenness, elite temporal stability, (or “stickiness”), and quality, which is true for English Wikipedia. In Phase 2, both evenness and stickiness are associated with quality. “The more, the merrier” seems to direct the contribution dynamics at this stage of development on this site. Evenness on French Wikipedia in the second phase, although increasing, was not significantly associated with quality. This could be explained by the simple fact that on French Wikipedia quality increased very slowly.

Finally, the theoretical assumptions of the Norden curve predicted a process of “defensive filtering” in Phase 3 of stabilization. Stickiness was supposed to increase and its effect on quality should have been positive. Similarly, we expected unevenness to also be positively related to quality, as the functional leaders (old timers that generated a lot of content) maintained the quality of the project by defending the gains made up to that point. As Table 1 shows, evenness was indeed negatively correlated to quality. In this last phase, higher quality was obtained both on French and English Wikipedia

when fewer cooks stirred up the pot. Moreover, these cooks were also temporarily resilient (sticky) on English Wikipedia. The data, in effect, broadly validated the model.

Returning to a more theoretical level, the idea that changes in the nature of the content go hand in hand with changes in the interactional dynamics can be directly applied and cross-validated by findings from other domains. Older collaborative groups develop a core group of highly productive individuals. These individuals are instrumental in filtering the new contributions, consolidating the quality gains, and preventing decay through random additions or vandalism. This insight is confirmed by the Apache project, which is a web-server that was built through open-source (collaborative) development over many years and has reached a high level of quality and stability. Predictably, the group of software designers who created the web-server known as the Apache project is small and largely made of a core group of old contributors and newcomers, all of whom are high-output producers[2].

In general, prioritizing quality should automatically lead to a limitation of access to the new content creation at the micro-level, mostly by reversion of new edits/comments by old timers, who have invested more in the articles/subprojects by direct contribution over time. This happens especially during the defensive filtering phase. During this phase, rejecting new

2 See the evolution of the contributions (or “commits”) on the GitHub Web page, https://github.com/apache/httpd/graphs/contributors

additions and defending existing ones also demands more effort, which further motivates older collaborators who want to keep their gateway positions. This phenomenon is also commonly observed in the open source collaborative software production projects.

We believe that these insights provide inspiration for a forward-looking research agenda, which will take into account new mechanisms for explaining the paradoxical positive evolution of online collaborative projects that increasingly rely on smaller groups of top contributors.

References

Aaltonen, A., & Seiler, S. (2015). Cumulative Growth in User-Generated Content Production: Evidence from Wikipedia. *Management Science*, *62*(7), 2054–2069. https://doi.org/10.1287/mnsc.2015.2253

Adler, B., De Alfaro, L., & Raman, P. (2008). Measuring author contribution to the Wikipedia. In *WikiSym 2008 Proceedings*. Association of Computer Machinery. Retrieved from http://users.soe.ucsc.edu/~luca/papers/08/wikisym08-users.pdf

Amrit, C., & Van Hillegersberg, J. (2010). Exploring the impact of socio-technical core-periphery structures in open source software development. *Journal of Information Technology*, *25*(2), 216–229.

Benkler, Y. (2007). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press.

Borzillo, S., Aznar, S., & Schmitt, A. (2011). A journey through communities of practice: How and why members move from the periphery to the core. *European Management Journal*, *29*(1), 25–42. https://doi.org/10.1016/j.emj.2010.08.004

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 *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1101–1110). New York, NY, USA: ACM. https://doi.org/10.1145/1357054.1357227

Giuri, P., Rullani, F., & Torrisi, S. (2008). Explaining leadership in virtual teams: The case of open source software. *Information Economics and Policy*, *20*(4), 305–315.

Halfaker, A., Sarabadani, A., & Taraborelli, D. (2016, October 26). New dataset shows fifteen years of Wikipedia's quality trends – Wikimedia Blog [blog]. Retrieved January 18, 2017, from https://blog.wikimedia.org/2016/10/27/wikipedia-quality-trends-dataset/

Hess, C. (Ed.). (2011). *Understanding Knowledge as a Commons*. Cambridge, Mass: MIT Press.

Huberman, B. A. (2001). *The Laws of the Web: Patterns in the Ecology of Information*. Cambridge, Mass.; London: The MIT Press.

Johnson, S., Faraj, S., & Kudaravalli, S. (2014). Emergence of Power Laws in Online Communities: The Role of Social Mechanisms and Preferential Attachment. *Management Information Systems Quarterly*, *38*(3), 795–808.

Jullien, N. (2012). What we know about Wikipedia: A review of the literature analyzing the project (s). *Available at SSRN*, *2053597*. Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2308346_code728676.pdf?abstractid=2053597&mirid=5&type=2

Jullien, N., Roudaut, K., & Squin, S. L. (2011). L'engagement dans des collectifs de production de connaissance en ligne. *Revue Française de Socio-Économie*, (8), 59–83. https://doi.org/10.3917/rfse.008.0059

Kane, G. C., Johnson, J. X., & Majchrzak, A. (2014). *Emergent Lifecycle: The Tension between Knowledge Change and Knowledge Retention in Open Online Co-Production Communities* (SSRN Scholarly Paper No. ID 2536089). Rochester, NY: Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=2536089

Kelly, K. (1995). *Out of control: The new biology of machines, social systems and the economic world*. Reading, Mass.: Addison-Wesley.

Kittur, A., Lee, B., & Kraut, R. E. (2009). Coordination in collective intelligence: the role of team structure and task interdependence. In *Proceedings of the 27th international conference on Human factors in computing systems* (pp. 1495–1504). Boston, MA, USA: ACM. https://doi.org/10.1145/1518701.1518928

Marwell, G., & Oliver, P. (2007). *The Critical Mass in Collective Action* (Reprint edition). Cambridge: Cambridge University Press.

Matei, S. A., Abu Jabal, A., & Bertino, E. (2018). Social-collaborative determinants of content quality in online knowledge production systems: comparing Wikipedia and Stack Overflow. *Social Network Analysis and Mining*, *8*(1), 36. https://doi.org/10.1007/s13278-018-0512-3

Matei, S. A., & Britt, B. C. (2017). *Structural Differentiation in Social Media: Adhocracy, Entropy, and the "1 % Effect"* (1st ed.). Springer Publishing Company, Incorporated.

Matei, S. A., Jabal, A. A., & Bertino, E. (2017). Do Sticky Elites Produce Online Knowledge of Higher Quality? In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (pp. 72–79). New York, NY, USA: ACM. https://doi.org/10.1145/3110025.3110040

Olson, M. (1971). *The Logic of Collective Action: Public Goods and the Theory of Groups, Second printing with new preface and appendix* (Revised edition). Cambridge, Mass.: Harvard University Press.

Pressman, R. S. (2005). *Software Engineering: A Practitioner's Approach*. Palgrave Macmillan.

Ransbotham, S., & Kane, G. C. (2011). Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in Wikipedia. *Mis Quarterly*, *35*(3), 613–627.

Rullani, F., & Haefliger, S. (2013). The periphery on stage: The intra-organizational dynamics in online communities of creation. *Research Policy*, *42*(4), 941–953.

Safner, R. (2016). Institutional entrepreneurship, wikipedia, and the opportunity of the commons. *Journal of Institutional Economics*, *12*(4), 743–771.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations* (1st ed.). New York: Doubleday.