DISA

Dipartimento di Informatica
e Studi Aziendali

2009/4

Wikibugs: the practice of template messages in open content collections

Loris Gaio, Matthijs den Besten, Alessandro Rossi, Jean-Michel Dalle

UNIVERSITÀ DEGLI STUDI
DI TRENTO

DISA WORKING PAPER

DISA

Dipartimento di Informatica
e Studi Aziendali

2009/4

Wikibugs: the practice of template messages in open content collections

Loris Gaio, Matthijs den Besten, Alessandro Rossi, Jean-Michel Dalle

UNIVERSITÀ DEGLI STUDI
DI TRENTO

DISA WORKING PAPER

Guidelines for authors
Papers may be written in English or Italian but authors should provide title, abstract, and keywords in both languages. Manuscripts should be submitted (in pdf format) by the corresponding author to the appropriate Associate Editor, who will ask a member of DISA for a short written review within two weeks. The revised version of the manuscript, together with the author's response to the reviewer, should again be sent to the Associate Editor for his consideration. Finally the Associate Editor sends all the material (original and final version, review and response, plus his own recommendation) to the Editor, who authorizes the publication and assigns it a serial number.

The Managing Editor and the Technical Officer ensure that all published papers are uploaded in the international RepEc publicaction database. On the other hand, it is up to the corresponding author to make direct contact with the Departmental Secretary regarding the offprint order and the research fund which it should refer to.

Ricardo Alberto MARQUES PEREIRA
Dipartimento di Informatica e Studi Aziendali
Università degli Studi di Trento
Via Inama 5, TN 38100 Trento ITALIA
Tel +39-0461-882147 Fax +39-0461-882124
E-mail: ricalb.marper@unitn.it

Wikibugs: the practice of template messages in open content collections

Loris Gaio
University of Trento
Department of Computer and Management Science
Via Inama 5
I-38100, Trento, Italy
+39 0461 882142
lgaio@cs.unitn.it

Matthijs den Besten
Ecole Polytechnique
Centre de Recherche en Gestion
ENSTA PREG CRG 32
boulevard Victor
75739 Paris Cedex 15, France
+33 1 45 52 64 27
matthijs.denbesten@polytechnique.edu

Alessandro Rossi
University of Trento
Department of Computer and Management Science
Via Inama 5
I-38100, Trento, Italy
+39 0461 882101
arossi@cs.unitn.it

Jean-Michel Dalle
Université Pierre et Marie Curie
4 place Jussieu
75252 Paris cedex 05
France
+33 1 44 18 07 15
jean-michel.dalle@upmc.fr

ABSTRACT

In the paper we investigate an organizational practice meant to increase the quality of commons-based peer production: the use of template messages in wiki collections to highlight editorial bugs and call for intervention. In the context of SimpleWiki, an online encyclopedia of the Wikipedia family, we focus on {complex}, a template which is used to flag articles disregarding the overall goals of simplicity and readability. We characterize how this template is placed on and removed from articles and we use survival analysis to study the emergence and successful treatment of these bugs in the collection.

JEL Codes

M15 (IT Management), L86 (Information and Internet Services)

Keywords

Commons based peer production, Wikipedia, wiki, survival analysis, quality, bug fixing, template messages, coordination.

1. INTRODUCTION

Issues pertaining to the reliability of open content collections are at the core of the agenda of both scholars and practitioners interested in commons-based peer production. As put forward by Larry Sanger, Wikipedia co-founder:

> "It's fun, first of all. But it can be fun for intellectually serious people only if we know that we're creating something of quality. And how do we know that? The basic outlines of the answer ought to be fairly obvious to anyone who has read Eric S. Raymond's famous essay on the open source movement, 'The Cathedral and the Bazaar'. Remember, if we can edit any page, then we can edit *each other's work*. Given enough eyeballs, all errors are shallow. We catch each other's mistakes and enjoy correcting them." [14]

Others have been more agnostic regarding the possibility of large mass peer screening to act as a substitute for source authoritativeness as a means for assuring quality [9]. Obviously, as far as trustworthiness is concerned, content peer production has also its share of skeptics in the scientific literature [7], in practitioners' view [10] and in the popular media [2]. Despite some exception [8], this lively debate has mostly being fueled by claims that have still to move towards the stage of sound empirical validation.

We build from previous empirical research in the field which has started to shed light on the role of institutions and organizational practices in channeling the largely unstructured efforts of voluntary contributors [5-6-12]. According to this line of research, peer production within wiki platforms makes extensive use of template messages – standard info-boxes placed on top of a given page – as coordination tool which ease the contribution to the production process of the various participants. In Wikipedia, for instance, there is an overwhelming number of templates, a.k.a. tags, which are used as a means to facilitate various goals and activities, such as to flag particular anomalies and dysfunctions of pages (e.g., violations of common policies or guidelines), and to call for specific actions for contributors (e.g., cleaning up, improvements in the organization of the text, and so on).

In what follows, we concentrate on templates signaling breaches of important policies or guidelines – consensual standards and advisory statements which every editor should bear in mind when editing an article in the collection – and, drawing a parallel with open source software methodologies, we treat placing/removing such templates from the text of an article a process similar to filing /closing a bug report in software development.

Previous research on bug fixing in open source development [1-3-4] has shed light on various important organizational issues, e.g., which are the main drivers of coders' attention, which elements account for quick fixing of bugs, and so on. Conversely, the same topic has received much less coverage in the realm of open content peer production. We believe that the analysis of bug fixing activities is crucial in improving our understanding how and to what extent it is possible to reconcile the apparent contrast between spontaneous collaborative authorship and quality assurance of a wiki collection.

In particular, in this paper we aim at understanding through descriptive and survival analysis which are the variables in an article production process which account for the emergence of a bug or have influence on how the bug is fixed. In particular, we model the dynamic of tagging as a survival process, linking the probability of entry/exit of a page into the "pathological state" to various explanatory variables. These variables can be traced back to three large families: measures of effort and intensity of work, measures of participation of users/division of labor and measures related to morphological features of pages (size, readability, similarity overtime, and so on). According to this framework, we perform survival analysis (*i*) on the duration of the initial non-pathological state, thus studying which variables shorten or increase the amount of time needed for an encyclopedic article to develop undesirable features (*ii*) on the duration of the pathological state, exploring how different variables affects the treatment or the persistence of such undesirable features.

This paper is organized as follows: we first offer some justification for the choice Simple Wikipedia as the empirical field of investigation. We then describe the methods used to retrieve data and how we performed the analysis. Then we present our main findings, distinguishing between descriptive and survival analysis. Finally we offer some concluding remarks and suggestions for further research on related topics.

2. DATA

In order to perform our study we mined data from Simple Wikipedia, an online encyclopedia – which belongs to the larger Wikipedia project – intended for readers whose first language in not English.

The fundamental reason of the choice of Simple Wikipedia over various other publicly available wiki–based collections lies in the strong commitment by the active participants in the activities of collaborative editing to a writing style which poses a strong emphasis on simplicity and readability [15-16]. Accordingly, the template {{complex}} (which in the early days of the collection was labeled {{unsimple}}) is used by editors in order to signal that a particular article is unsatisfactory as far as readability is concerned. This is certainly not the only instance of use of template messages which closely parallels how bug reports work in software development (e.g., a prominent Wikipedia template in this respect is represented {{NPOV}}, which flags the breach of the fundamental policy of writing articles according to a neutral point of view). Nevertheless, an additional rationale pushing further the argument for choosing {{complex}} from Simple Wikipedia over other alternatives is represented by the possibility of computing measures of simplicity/readability, derived from the computational linguistic tradition, which can be used as objective appraisal of the gravity of a bug and of the effectiveness of the work done to fix it.

Data mining and preparation of the database for the analysis. We used the July 2008 archive of Simple Wikipedia, available at: http://downloads.wikimedia.org, which for every revision made on an article page lists the following data: the user–id of the editor (IP address in case of anonymous edits), date and time of the edit, comments made by the editor and the full (wiki markup) text of the revision. We selected from the archive all the revisions corresponding to article pages which had been tagged at least once with the {{unsimple}}|{{complex}} templates. In order to avoid biases due to very short series for some datapoints in the survival analysis, we restricted the analysis to article pages which had been revised at least 15 times. For each article page we limited our extraction to all revisions belonging to the interval spanning from the first edit to the revision antecedent to the removal of the {{complex}} template[1]. After this selection, we ended up with 378 article pages for the analysis. De-wikification of the text and categorization of registered users (in terms of administrators, bots and "plain" registered users) have been performed according to previous literature [5-6]. While some studies on the English Wikipedia have shown that actual changes in a given article page are sometime the result of longer discussions occurring at the level of the corresponding talk page [11, 18], the use of talk pages as a means to anticipate and discuss actual changes is relatively infrequent in Simple Wikipedia. Thus, the dataset employed in our analysis is restricted to data from article pages only.

Vandal filtering. Previous work has highlighted the short life span for vandal edits in wiki collections [17]. While this generally reassure on the limited impact of these malicious activities on the quality of the whole archive, at the same time when studying the process of development of articles one has to carefully evaluate whether vandal edits might introduce distortions in the interpretation of the data. In our case filtering out vandal edits was essential in order be able to perform unbiased measures of work activity related to article pages (e.g., number of revisions, number of unique contributors, etc.). While we were aware of the existence of algorithms for the automatic detection of vandalisms [13], given the relatively limited number of revision involved, we preferred to manually check for vandalisms. We performed this activity both using comment analysis (in order to single out reverts which were explicitly accounted by editors as fixes to vandal edits) and MD5 hash (computed over the full text of a revision) comparisons across subsequent revisions of an article page. Overall, we filtered out from the dataset 11% of revisions which were categorized as vandalisms or revert edits fixing previous vandalizations.

3. METHODOLOGY

We applied survival analysis to study two different albeit intertwined phenomena: (*i*) transition of article pages from the initial "simple" phase (from now on: regime 1) to the subsequent "unsimple" phase (from now on: regime 2) and (*ii*) exit from regime 2. The observation periods are, respectively, from the very first version of an article page to the revision in which the template {{complex}} appears, and from the latter to the revision in which the template is edited out. By definition of the sample, for the first event (exit from regime 1) all observations are uncensored, while for the second

[1] In the case of pages in which the {{complex}} tag has never been removed (a.k.a. censored pages) we took all the available revisions. Also, we did not consider instances of repeated flagging of one page, where one page, after returning in the "simple" regime, is flagged once again as complex.

event some observation are censored, meaning that in some cases the template has never been removed from the article page.

In order to perform the survival analysis we used the dataset extracted in the previous Section to calculate the following variables:

- *duration of regimes*:
  - `duration1`, `duration2`: duration (days) of regime 1, 2;
  - `reactime`: reaction time (days) measured as time between tagging and the first subsequent revision of the article;
- *intensity of efforts and division of labor*:
  - `revs1`, `revs2`: number of edits in regime 1, 2;
  - `regrevs1`, `regrevs2`: number of edits by registered users in regime 1, 2;
  - `admrevs1`, `admrevs2`: number of edits by administrators in regime 1, 2;
  - `anonrevs1`, `anonrevs2`: number of edits by anonymous users in regime 1, 2;
  - `botrevs1`, `botrevs2`: number of edits by bots in regime 1, 2;
  - `anon1`, `anon2`: number of unique anonymous editing in regime 1, 2;
  - `admin1`, `admin2`: number of unique administrators editing in regime 1, 2;
  - `bot1`, `bot2`: number of unique bot editing in regime 1, 2;
  - `reg1`, `reg2`: number of unique registered users editing in regime 1, 2;
- *textual measures*:
  - `char1`, `char2`: character count at the end of regime 1, 2;
  - `word1`, `word2`: word count at the end of regime 1, 2;
  - `read1`, `read2`: Flesch readability score at the end of regime 1, 2;
  - `cossim`, `jacsim`: Cosine and Jaccard similarity measures computed between the end of regime 1 and the end of regime 2.

Readability and similarity metrics were computed according to [6].

4. RESULTS

4.1 Evidence from regime 1

According to a Kaplan Meyer estimate, regime 1 seems to fit quite well to a Cox Proportionality Hazard class model. Figure 1 depicts the survival function, using Kaplan Meyer estimation, fitted using duration times for regime 1.

In order to assess the different effects of covariates in the termination of regime 1, we start considering the impact of division of labor, and in particular the incidence of efforts by different kind of users towards duration of regime 1. For this purpose we need preliminary to screen for the possible existence of multicollinearity issues between the various variables.

Figure 1. Survival function for regime 1 (Kaplan Meyer estimator).

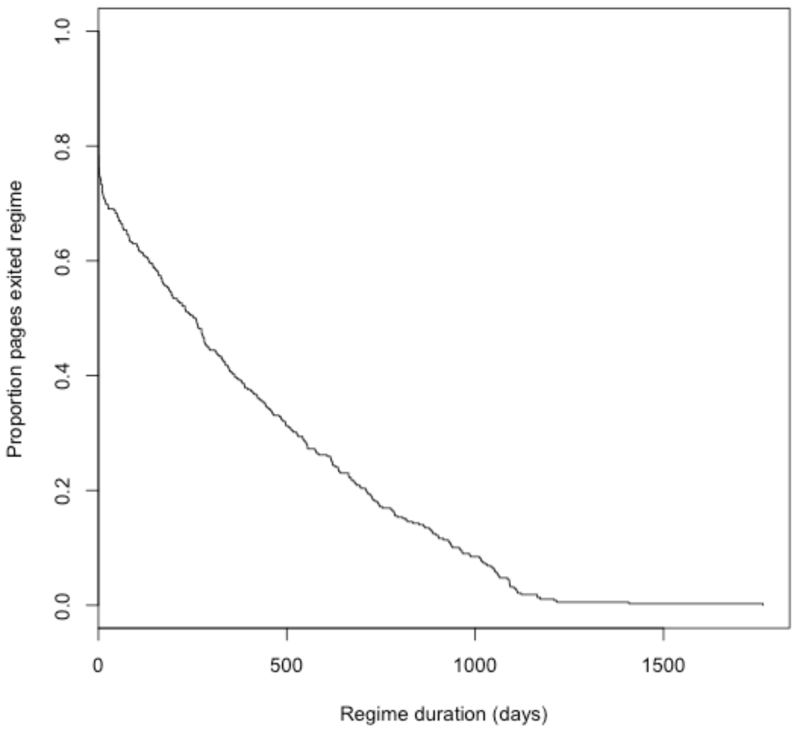

Table 1 summarize correlations existing between the duration of regime 1, and two families of covariates: variables related to efforts (edits) exerted by different categories of participants and variables measuring the number of participants (for such categories). The table shows the existence of strong correlation between participants and edits for each category considered: this suggests not to use both families in survival estimation.

Table 1. Correlation matrix for regime 1 covariates

	duration1	regrevs1	admrevs1	anonrevs1	botrevs1	reg1	adm1	anon1	bot1
duration1	1	.401	.398	.511	.615	.710	.575	.618	.640
regrevs1	.401	1	.296	.459	.304	.651	.364	.465	.324
admrevs1	.398	.296	1	.525	.387	.467	.807	.590	.367
anonrevs1	.511	.459	.525	1	.447	.652	.605	.912	.406
botrevs1	.615	.304	.387	.447	1	.525	.573	.506	.895
reg1	.710	.651	.467	.652	.525	1	.588	.755	.531
adm1	.575	.364	.807	.605	.573	.588	1	.673	.537
anon1	.618	.465	.590	.912	.506	.755	.673	1	.441
bot1	.640	.324	.367	.406	.895	.531	.537	.441	1

Consequently, a CoxPH model has been fitted in order to explain the impact of the three families of covariates depicted in Section 1. Overall, only the variables pertaining to intensity and division of labor seem to have a significant effect in explaining the length of regime 1, while variables regarding other features of the pages, such as size, readability, similarity have no explanatory power. For sake of compactness we present the final models only, which are summarized in Table 2.

In Model 1 the duration of regime 1 is negatively affected by the number of revisions by all categories of contributors. Similarly, there is a negative impact on duration when considering the number of different contributors per category (Model 2). The latter model seems to have a higher descriptive power as far as Rsquare and model tests are concerned.

Table 2. Survival Analysis on Regime 2 Inception

Variable	Model 1	Model 2
`regrevs1`	-0.039*** (0.0129)	_
`admrevs1`	-0.037* (0.0203)	_
`anonrevs1`	-0.027** (0.0089)	_
`botrevs1`	-0.105*** (0.0150)	_
`reg1`	_	-0.147*** (0.0267)
`adm1`	_	-0.135*** (0.0510)
`anon1`	_	-0.025* (0.0157)
`bot1`	_	-0.254*** (0.0375)
Rsquare	0.370	0.492
L ratio	175	256
Wald	122	177
logrank	120	185

p-values significance: *<0.1, **<0.05, ***<0.01

Overall the two models seems to suggest that both the level of effort on a page (in terms of revisions) and the number of participants in the editing process seem to anticipate the emergence of readability concerns. At this point of the analysis it is still difficult to judge whether this shortening is more due to a variant of the Linus' law (more eyeballs resulting in the anticipatory detection of a defect) or rather due to diminishing returns related with increases in the number of contributors. While the second model seems to be more ambiguous in this respect, the first one seems more clearly to suggest a connection between increases in work intensities and the emergence of a bug as the result of coordination conflicts. Nevertheless this issue seems to be worth of further scrutiny.

4.2 Evidence from regime 2

Similarly to the previous regime, for regime 2 durations a Kaplan Meyer estimate has been computed (Figure 2) and the model seem again to fit quite well a Cox Proportionality Hazard class model.

Figure 2. Survival function for regime 2 (Kaplan Meyer estimator)

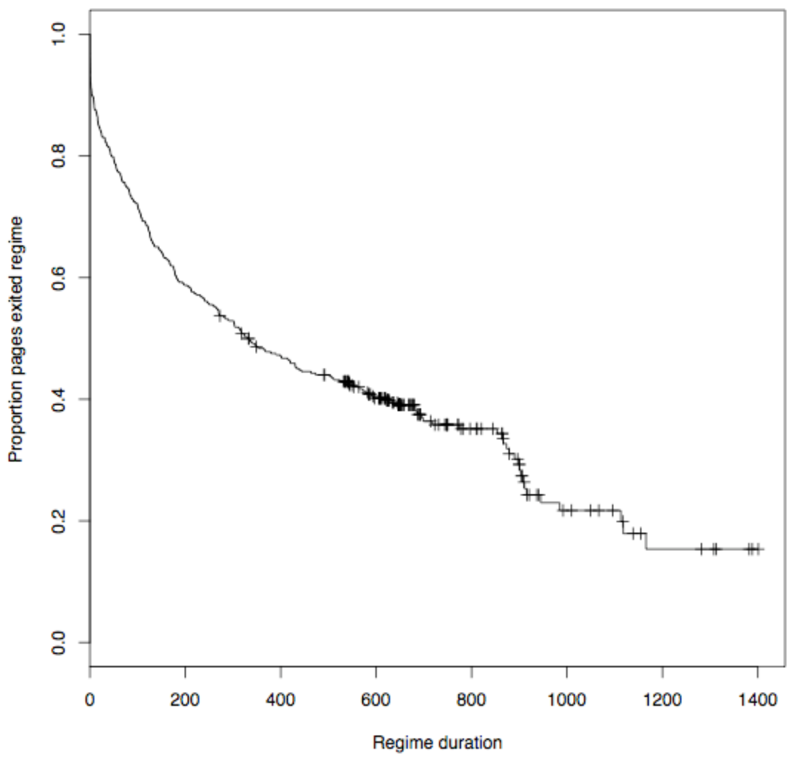

Table 3 confirms the existence of a strong correlation between the number of revisions made by different classes of participants and the number of participants (for the same classes), again suggesting to avoid the use of both families in the same model estimation in order to avoid for multicollinearity problems.

Similarly to what has been done for regime 1, we test for the same hypotheses related to efforts and division of labor; we look whether the total number of revisions and the number of different contributors in the various classes do play a significant role in exiting from regime 2.

Results are reported in Table 4. Here the variables related to the intensity of efforts (number of revisions) are not significant with the exception of revision made by bots (Model 1). On the contrary, all classes of users are significant when considering the number of different contributors per category (Model 2). In particular the shortening of the pathological regime seems to be affected by the presence of administrators, registered and bot users, while the presence of anonymous users seems to delay the fixing process.

Similarly to the previous Subsection, other covariates (in particular the textual-related covariate) have no incidence on the survival process. In particular, the reaction time to flagging has a negligible impact on regime 2 duration (for simplicity the model is not reported). Model 3 allows to introduce in the survival the duration of regime 1 (that can be also thought as the overall life of the page at starting of regime 2) as a covariate. This variable is significant and affects positively the duration of regime 2. A possible interpretation is that the older the page at time of flagging, the more difficult is to solve successfully readability issues.

Table 3. Correlation matrix for regime 2 covariates

	duration2	duration1	regrevs2	admrevs2	anonrevs2	botrevs2	reg2	adm2	anon2	bot2	react
duration2	1	.000	.307	.376	.327	.599	.438	.543	.355	.706	.199
duration1	.000	1	.182	.190	.197	.178	.220	.226	.204	.159	-.066
regrevs2	.307	.182	1	.816	.792	.467	.887	.710	.803	.358	-.031
admrevs2	.376	.190	.816	1	.929	.618	.904	.897	.931	.490	-.038
anonrevs2	.327	.197	.792	.929	1	.571	.888	.842	.988	.436	-.038
botrevs2	.599	.178	.467	.618	.571	1	.635	.645	.583	.922	-.061
reg2	.438	.220	.887	.904	.888	.635	1	.849	.908	.525	-.028
adm2	.543	.226	.710	.897	.842	.645	.849	1	.866	.578	.006
anon2	.355	.204	.803	.931	.988	.583	.908	.866	1	.453	-.028
bot2	.706	.159	.358	.490	.436	.922	.525	.578	.453	1	-.034
react	.199	-.067	-.031	-.038	-.038	-.061	-.028	.006	-.028	-.034	1

A final remark is worth on the variable measuring the efforts made by users which originally tagged the page. This variable is not significant, thus hinting to a quite different story with respect to open source development as far as to bug fixing is concerned, and reinforcing a view of open content creation communities as made more by “passers-by” users, rather than by contributors which commit themselves to a particular artifact on a long term perspective.

5. CONCLUSIONS

In the paper we study how bugs emerge and are treated in collaborative authorship collections, and we report our preliminary findings on a survival analysis performed on the durations of entry and exit times of pages from a pathological treatment.

In what follows, we summarize the main findings of the analysis and we suggest how to challenge further the evidence collected. As far as regime 1 is concerned, we showed that entry in the pathological regime is affected both by the number of users and their efforts, and the former model seems to be relatively more robust. Conversely, no structural feature of pages like size, readability, similarity, and so on are helpful in explaining the {{complex}} tagging. Overall, survival findings might highlight the existence of competing explanations regarding the shortening of regime 1 duration (complexity/coordination issues vs. "eyeballs" hypothesis), which call for further scrutiny.

Table 4. Survival Analysis on Regime 2 Termination

Variable	Model 1	Model 2	Model 3
regrevs2	-0.014 (0.0098)	_	_
admrevs2	0.015 (0.0196)	_	_
anonrevs2	0.007 (0.0064)	_	_
botrevs2	-0.091*** (0.0103)	_	_
reg2	_	-0.060* (0.0373)	-0.084** (0.0379)
adm2	_	-0.233*** (0.0516)	-0.262*** (0.0519)
anon2	_	0.074*** (0.0131)	0.081*** (0.0135)
bot2	_	-0.224*** (0.0222)	-0.228*** (0.0220)
duration1	_	_	.0004** (0.0001)
Rsquare	0.320	0.499	0.537
L ratio	146	262	291
Wald	83.4	152	174
logrank	89	189	213

Regarding regime 2, exit from the pathological state seems to depend on factors related on the number of participants only. In particular, while anonymous users have detrimental effects, all three categories of registered users seems to help in sorting the readability issue. during the regime shortens its duration.

Finally, we mentioned that both entry and exit cannot be traced back neither to reaction time measures, nor other structural features of pages, such as readability, similarity, and so on. In this respect we think that, other statistical models, i.e. event analysis, might represent a more suitable way to study in a more dynamic way their effect on pages being tagged.

The research here reported can be regarded as the first step towards a more comprehensive understanding on the effectiveness of various organizational practices on bug spotting and bug fixing in collaborative authored open content collections. In this line, we find promising to apply the survival analysis framework to the study of other template messages, which signal the breach of important policies (such as above mentioned {{NPOV}}) and to larger datasets, such as the ones which can be derived by mining the English Wikipedia.

ACKNOWLEDGMENTS

Loris Gaio and Alessandro Rossi gratefully acknowledge partial support of the work from institutional funds from MIUR (projects FIRB03 and PRIN05).

REFERENCES

[1] Crowston, K. and Scozzi, B. 2004. Coordination practices for bug fixing within FLOSS development teams. In: Proc. 1st CSAC. Porto, Portugal.

[2] Cuozzo, S. 2008. Urban Myths: What Wikipedia gets wrong about NYC. New York Post. http://www.nypost.com/seven/08242008/postopinion/opedcolumnists/urban_myths_125773.htm?page=0. Accessed 21 January 2009.

[3] Dalle, J.-M. and den Besten, M. 2007. Different bug fixing regimes? A preliminary case for superbugs. In: Proc of the 3rd Int Conf on Open Source Systems. Limerick, Ireland.

[4] Dalle, J.-M., den Besten and M., Masmoudi, H. 2008. Channelling Firefox developers: Mom and dad aren't happy yet. In: Open Source Development, Communities and Quality, B. Russo, E. Damiani, S. Hissam, B. Lundell, G. Succi, Eds., IFIP International Federation for Information Processing;275. Boston, Springer:265–271.

[5] den Besten and M., Dalle, J.-M. 2008. Keep it Simple: A Companion for Simple Wikipedia? Industrial Innovation 15(2):169–178.

[6] den Besten, M., Rossi, A., Gaio, L., Loubser, M. and Dalle, J.-M. 2008. Mining for practices in community collections: finds from Simple Wikipedia. In: Open Source Development, Communities and Quality, B. Russo, E. Damiani, S. Hissam, B. Lundell, G. Succi, Eds., IFIP International Federation for Information Processing;275. Boston, Springer:105–120.

[7] Denning, P., Horning, J., Parnas, D. and Weinstein, L. 2005. Wikipedia risks. CACM 12(48):152.

[8] Giles, J. 2005. Internet encyclopaedias go head to head. Nat. 438(7070):900–901.

[9] Ito, J. 2004. Wikipedia attacked by ignorant reporter. Joi Ito's Web. http://joi.ito.com/weblog/2004/08/29/wikipedia-attac.html. Accessed 3 February 2009.

[10] Keen, A. 2007. The Cult of the Amateur: How Today's Internet is Killing Our Culture. Doubleday Business, New York.

[11] Kittur, A. and Kraut, R.E. 2008. Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination. CSCW 2008: Proceedings of the ACM Conference on Computer-Supported Cooperative Work. New York: ACM Press.

[12] Loubser, M. and den Besten, M. 2008 Wikipedia Admins and Templates: The Organizational Capabilities of a Peer Production Effort. SSRN. http://ssrn.com/abstract=1116171. Accessed 1 December 2008.

[13] Potthast, M., Stein, B. and Gerling, R. 2008. Automatic Vandalism Detection in Wikipedia. In: Advances in Information Retrieval, C. Macdonald, I. Ounis, V. Plachouras, I. Ruthven, R.W. White, Eds., Boston, Springer:663–668.

[14] Sanger, L. 2001. Wikipedia is wide open. Why is it growing so fast? Why isn't it full of nonsense?. Kuro5hin. http://www.kuro5hin.org/story/2001/9/24/43858/2479. Accessed 10 November 2008.

[15] Wikipedia:Simple English Wikipedia. 2008. In: Simple Wikipedia, the free encyclopedia. http://simple.wikipedia.org/wiki/Wikipedia:Simple_English_Wikipedia. Accessed 10 November 2008.

[16] Wikipedia:How to write Simple English articles. 2008. In: Simple Wikipedia, the free encyclopedia. http://simple.wikipedia.org/wiki/Wikipedia:How_to_write_Simple_English_articles. Accessed 10 November 2008.

[17] Viégas, F., Wattenberg, M. and Dave, K. 2004. Studying cooperation and conflict between authors with history flow visualizations. In: Proc of the SIGCHI Conf on Human factors in computing systems 2004, New York, ACM Press: 575–582.

[18] Viégas, F., Wattenberg, M., Kriss, F. and van Ham, F. 2007. Talk before you Type: Coordination in Wikipedia. In: 40th Annual Hawaii International Conference on System Sciences (HICSS'07).