

Proliferation in Enterprise Wikis

Hans-Jörg Happel*, Marius Treitz**

*FZI Forschungszentrum Informatik
Information Process Engineering (IPE)
76131 Karlsruhe, Germany
happel@fzi.de

**International University
Campus 3
76646 Bruchsal, Germany
marius.treitz@i-u.de

Wikis have become a popular tool for the collaborative creation of hyperlinked documents. Besides the well-known Wikipedia, enterprises are increasingly adopting wikis for internal documentation and knowledge management purposes. A typical attribute of wikis is their openness concerning the modification and creation of content and the lack of pre-defined structures. While this enables an easy access with low barriers for contributors, it also introduces certain problems – such as redundant or outdated content – which we subsume under the term of "wiki proliferation". In this paper, we analyze wiki proliferation in terms of its causes, consequences and common countermeasures. Our analysis is based on six qualitative interviews with companies which are using wikis internally. The interviews show that proliferation is a serious problem for these organizations which is actively addressed by different countermeasures.

1. Introduction

The software genre of a "Wiki" describes a lean approach to web-based content management, allowing multiple users to collaborate on the creation of documents. Basically, a wiki consists of titled pages that are mutually connected by hyperlinks. Thus every concrete installation of a wiki can be regarded as a small "world wide web", because it mainly consists of hyperlinked documents. However, in contrast to the web, wiki pages can be directly created and modified in the browser – often without any authentication required. When a user edits a page, a revision history is saved so that previous versions can always be restored. Thus, social control replaces complicated rights management to prevent malicious changes [LC01].

The most popular instance of a wiki – the Wikipedia project – shows how this concept works for the collaborative creation of an online encyclopedia. A large number of users take care of cleaning, repairing and other maintenance tasks [VWD04]. The success of Wikipedia makes wikis increasingly popular as intranet and knowledge management solutions in organizations. Recent studies show that up to 30% of the surveyed organizations make use of wiki technology or plan to do so in the future

[Eco07, McK07]. This is stressed by the availability of enterprise-ready wiki solutions such as Lotus Quickr¹ or Intel SuiteTwo².

While enterprise wikis are mostly free from vandalism or malicious edits, they suffer from proliferation as they grow (c.f. section 2). This is mostly due to the lack of pre-defined structure in wiki systems. While this is not a problem for the Wikipedia, since its flat article structure is predetermined by the characteristics of an encyclopedia, content in enterprise wikis stems from an open domain, spanning ideas, new projects, new products and so on. Thus, managing the growth of enterprise wiki is an important topic, which is often addressed using the term "wiki gardening".

Surprisingly, both issues – the growth and proliferation of enterprise wikis and its countermeasures such as wiki gardening – are not well-defined in practice and research. Also research on enterprise wikis has concentrated on wiki adoption and usage [MWY06, Ba07].

Our goal is to analyze the phenomenon of proliferation in enterprise wikis. Therefore, our paper makes two contributions. First we provide an initial characterization of wiki proliferation and possible countermeasures. Second, we present results of qualitative interviews in companies which are running wikis for internal usage.

2. Wiki proliferation and its countermeasures

While many authors stress the importance of a special "wiki philosophy" and the necessity of measures such as wiki gardening, the implied underlying problems are only fuzzily described. Although the term "wiki proliferation" is mentioned sometimes, it is not precisely characterized, but defined indirectly by describing resulting problems or countermeasures.

However, most authors seem to agree that proliferation is related to the size and growth of a wiki. [Buf06] writes: "The number one problem reported by people ... is related to the wiki open structure that makes navigation, orientation and search sometimes difficult. This occurs when the wiki reaches several thousand pages". [CM05] also refer to growth as a main cause: "...the ease of contribution has resulted in a proliferation of wiki pages. In fact, one user even stated that 'retrieving becomes harder every day, because of the growing amount of content and the lack of structure'". In this paper, we define the term proliferation with respect to wikis as follows:

Proliferation describes all phenomenons that occur as a side effect of a wiki's growth and lead to problems that can be measured; such problems are either problems related to the access of content, the quality of content or the structure of the pages as a whole.

Examples of proliferation are old and outdated content, information that is redundant, as well as the poor linkage between content.

We consider the following aspects of our definition as important: (1) proliferation may not only affect single pages of content, but the entire wiki, (2) proliferation is a natural side effect of a growing number of wiki pages and authors and (3) proliferation results in measurable problems.

¹ <http://www-306.ibm.com/software/lotus/products/quickr/>

² <http://spikesource.com/suitetwo/>

Furthermore, we differentiate between three domains of proliferation in a wiki environment:

1. *Content*: Problems that affect the content of single wiki pages. Such problems are mostly related to the quality of the content, such as the occurrence of inaccurate or old content (see e.g. [STGS05, BS07]).
2. *Structure*: Problems related to the structure of the wiki pages as a whole. This includes problems such as the unnecessary scattering of closely related content across multiple pages.
3. *Access*: Problems that hinder the access to relevant information in the wiki. This may occur when pages are untagged, poorly titled or in some other way harder to access than other content.

In contrast to common metrics from the area of information quality (such as described in [STGS05]), we do not only address properties of the content, but also the overall structure of the wiki-system as well as access problems. For each of these three categories, we derived a number of problem patterns based on existing problem scenarios from our interviews (c.f. section 3) as well as literature. Examples for such patterns, are depicted in Figure 1.

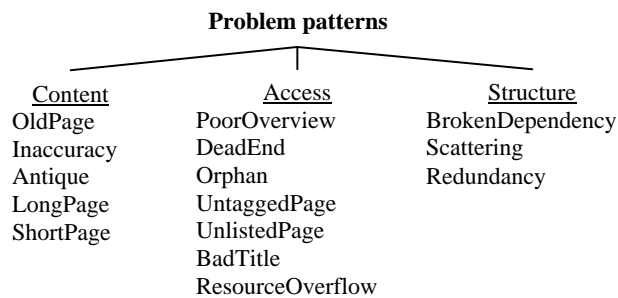


Figure 1: Wiki proliferation patterns

The growth of a wiki (in terms of the number of articles) can be regarded as a primary cause of these problems [Buf06]. Related to this, also the number of users resp. authors is influencing proliferation. "A wiki is a place where people gather to collaborate. Like a meeting room or a living room, it gets cluttered after a little while" [WG07]. A further cause can be vandalism – i.e. the conscious damage of wiki content [VWD04]. However, the latter is not wide-spread in enterprise wikis due to better control mechanisms [MWY06].

Similar to wiki proliferation itself, also its consequences have not been systematically analyzed so far. [Buf06, CM05] show that problems with search and

	Preemptive	Subsequent
Manual	<ul style="list-style-type: none"> • Employee training • Quality standards, rules of conduct • Tagging, categorization • Use of templates 	<ul style="list-style-type: none"> • Wiki Gardening
Automated	<ul style="list-style-type: none"> • Rights management • Workspaces 	<ul style="list-style-type: none"> • Bots, Plug-ins, Scripts • Watchlists, Patrolling of edits

Table 1: Countermeasures against proliferation

access of relevant information and the resulting increase in search time and effort are a significant problem. Further consequences of proliferation are the lack of acceptance of the wiki by new users [WP07] or the loss of already existing users [Buf06].

Although the concept of proliferation is not explicitly defined, a wide range of tools and practices exist that are concerned with the removal or prevention of proliferation in a wiki environment. Here we differentiate between preemptive, as well as subsequent measures: measures targeted to prevent the occurrence of proliferation and measures to remove existing proliferation, respectively. Furthermore, we differentiate between manual practices and automated measures (see Table 1 for some examples).

3. Wiki proliferation in practice

To investigate the phenomenon of proliferation and its countermeasures, we conducted six semi-structured expert interviews [Wi00] in organizations which are using wikis.

Name	Company size ³	Using wikis since	Wiki users ⁴	Wiki size (pages)
Alpha	Very large	~ 2 years	Medium (50-100)	Medium (100-500)
Beta	Medium	~ 1 year	Few (<50)	Small (<100)
Chi	Medium	> 2 years	Many (>500)	Very large (>1000)
Delta	Very large	> 2 years	Many (>500)	Very large (>1000)
Epsilon	Very large	> 2 years	Medium to many	Large (500-1000)
Thy	Small	> 2 years	Few (<50)	Large (500-1000)

Table 2: Descriptive data of the surveyed enterprise wikis

³ Company size: Small (< 100 employees), Medium (100-1000), Large (1000-5000), Very large (> 5000)

⁴ The numbers of users and wiki pages concerning Alpha, Beta and Epsilon are estimations

The interviews consisted of a short questionnaire which was to be completed prior to the interview, the interview itself, as well as a short questionnaire afterwards. They were conducted with wiki experts from six different organizations. Table 2 shows the characteristics of the six wikis. Four interview partners reported that their company maintains a central wiki instance while two companies reported to maintain several independent installations. Two organizations introduced wikis formally, while the grass-roots effort of employees was responsible for wiki uptake in the remaining four organizations.

Table 3 describes the status of the integration of wikis into the IT infrastructure and processes. Phase 1 denote wiki usage without explicit administration and maintenance. Phase two denotes centrally administered wikis, which are not integrated into the IT processes. Phase 3 denotes that organization have integrated wikis completely.

The third column characterizes the status of proliferation. Group 1 denotes that proliferation exists, but is not considered a problem. Group 2 means that proliferation occurs and causes problems, but is not addressed by active countermeasures. Group 3 finally denotes that problematic proliferation is explicitly and systematically addressed.

Name	Status of integration	Status of proliferation	Dominant problems ⁵
Alpha	Phase 3	Group 1	Content
Beta	Phase 1	Group 2	Content
Chi	Phase 3	Group 2	Content, Structure, Access
Delta	Phase 2	Group 3	Content, Structure, Access
Epsilon	Phase 1	Group 2	-
Thy	Phase 3	Group 2	Content

Table 3: Integration, proliferation and dominant problems

The results show that proliferation is a severe problem in enterprise wiki settings. Of the three problem domains (*Content*, *Structure*, *Access*), the interview partners reported that problems related to *content* – such as the aging of content or the inaccuracy of content – are the most frequent, said to occur *often*⁶ on average. Consequences related to the *structure* – such as the scattering of content – and problems related to the *access* of the content are occurring *sometimes* on average. The severity of the consequences varies: Some cases of proliferation merely impair the usage of the respective wikis, while exceptional cases also lead to severe damages.

All surveyed organizations notice the occurrence of proliferation and make use of countermeasures to target it. One organization employs a person whose sole task is to manage the wiki, coordinate efforts and support users; other organizations usually administer the wiki technically while the users are asked to maintain the content by themselves. The interviews show that especially the practice of *Wiki Gardening* is

⁵ Includes problem domains which were reported to occur „often“ or „very often“

⁶ Given the options: Never, Sometimes, Often, Very Often

applied; this means that contents affected by proliferation are independently refactored by participating users (manual, subsequent). This practice usually works well on a small scale, while the low number of active maintainers makes this problematic in larger wikis.

	Preemptive	Subsequent
Manual	Alpha, Delta, Thy	Alpha, Chi, Delta
Automated	Alpha, Chi, Delta	-

Table 4: Countermeasures against proliferation in the surveyed organizations

Four organizations also made use of preemptive and automated measures, such as implementing a user rights management that restricts the user's ability to edit the wiki to a plausible extent. Again four organizations made use of preemptive, manual methods, such as imposing quality standards which inform the users about general policies related to editing the wiki. Automated, subsequent methods, such as automated scripts, have not been incorporated by any of the organizations. The use of automated methods however is naturally limited by the support of the software for such methods.

The interviews show that, while proliferation existed in both small and large interviewed enterprises, the experts from larger enterprises dealt with larger wikis and mentioned the use of more sophisticated countermeasures, such as rights management, the use of spaces or improved user interfaces. The small enterprises relied more heavily on manual countermeasures, especially wiki gardening. This practice however seems to be of limited use in larger wikis due to the inability to manually control all contents with the help of a small set of active maintainers. Therefore, in such larger wikis (systems with more than 1000 pages/500 users), automated solutions would be helpful.

The separation of articles into subspaces, for example, provides better organization, while rights management helps to reduce unwanted edits. However, these automated methods are still limited in many ways by the lack of sophistication of the wiki engines. The interviews show that the (enterprise) wiki engines do not yet fully satisfy the needs of the users in this respect. Users have a strong desire for more professional tools and better solutions to support their use of wikis and to successfully prevent and remove proliferation.

4. Summary

In this paper, we provided a first systematic overview of the phenomenon of wiki proliferation and the associated countermeasures. Problems were differentiated into the domains of content, structure and access. Countermeasures were distinguished along the dimensions of preemptive/subsequent and manual/automated.

To further investigate and refine these characterizations, we conducted a series of six explorative expert interviews regarding the current problems and practices in enterprises using wikis. Our results show, that proliferation is observed in all six wikis, especially concerning content. The impact of consequences varies: while most

proliferation issues merely impair the usage of the wiki, also some severe problems were reported – e.g. when outdated information from wikis was used.

To address proliferation, several countermeasures were in place. Our results show, that especially “wiki gardening” in terms of subsequent manual cleanup is applied. Also, some organizations use preemptive methods such as access control or quality standards. Subsequent automated methods were not used by any organization.

Due to the rising interest in enterprise wikis [Eco07, McK07] and the reports of experienced organizations in our survey, we expect that the importance to manage wiki proliferation will grow in the future. Therefore, we aim to unify the heterogeneous set of countermeasures in order to suggest concrete guidelines and tools for managing the growth of wikis. Our paper provides first insights towards a deeper understanding of wiki proliferation and how to deal with it. Since most of our arguments are of qualitative nature, a future step will be a quantitative survey among a larger number of users.

5. References

- [Ba07] Bartel, T.: Wikis in Enterprises. Available at: <http://wikipedistik.de/survey/results.html> (Date of Access: 18.08.2007)
- [BS07] Braun, S.; Schmidt, A.: Wikis as a technology fostering knowledge maturing: What we can learn from Wikipedia. Proceedings of the 7th International Conference on Knowledge Management (I-KNOW 2007), Graz, Austria. Springer 2007.
- [Buf06] Buffa, M.: Intranet Wikis. In: Proc of IntraWeb workshop, WWW2006, Edinburgh, Scotland, 2006.
- [CM05] Chau, T. and Maurer, F. 2005. A case study of wiki-based experience repository at a medium-sized software company. In Proceedings of the 3rd international Conference on Knowledge Capture (K-CAP '05). ACM Press, New York, NY, 185-186.
- [Eco07] Economist Intelligence Unit: Serious business: Web 2.0 goes corporate. Available at: http://socialmediagroup.ca/wp-content/uploads/2007/06/smg_eiu_web20.pdf (Date of Access: 18.08.2007)
- [LC01] Leuf, B., Cunningham, W.: The wiki way: Quick collaboration on the web. Addison-Wesley, 2001.
- [McK07] How businesses are using Web 2.0: McKinsey Quarterly, March 2007. http://www.mckinseyquarterly.com/article_page.aspx?L2=16&L3=16
- [MWY06] Majchrzak, A., Wagner, C., and Yates, D. 2006. Corporate wiki users: results of a survey. In: Proceedings of the 2006 International Symposium on Wikis, Odense, Denmark, ACM Press, New York, NY, 2006, S. 99-104.
- [STGS05] Stvilia, B., Twidale, M.B. Gasser, L. & Smith, L.C. (2005). Information Quality Discussions in Wikipedia. Technical Report ISRN UIUCLIS--2005/2+CSCW.
- [VWD04] Viegas, F.B.; Wattenberg, M.; Dave, K.: Studying cooperation and conflict between authors with history flow visualizations. Proc. CHI'04, Vienna, 2004. p. 575-582.
- [WG07] Wiki gardening tips. Available at: http://www.eu.socialtext.net/exchange/index.cgi?wiki_gardening_tips. (Date of Access: 18.08.2007)
- [Wi00] Witzel, A.: The problem-centered interview [27 paragraphs]. Forum: Qualitative Social Research 1(1). Available at: <http://www.qualitative-research.net/fqs-texte/1-00/1-00witzel-e.htm> (Date of Access: 18.08.2007).
- [WP07] Wikipatterns. Available at: <http://www.wikipatterns.com/> (Date of Access: 18.08.2007)