Knowledge Management by Internet Communities: Is it an example for the corporate world?

Timo Staub, MSc  
_Berne University of Applied Sciences (BUAS), Hallerstrasse 10, 3010 Berne, Switzerland, timo.staub@bfh.ch_

Dorothea Christ, Dr. habil.; Katharina Tschannen, MA  
_Pädagogische Hochschule Thurgau (PHTG) and Bundesamt für Sozialversicherungen (BSV), Kreuzlingen and Bern, dorothea.christ@phtg.ch, katharina.tschannen@bsv.admin.ch_

Abstract: Many Internet users have joined virtual communities who meet around a certain Website. Some of these Websites have become a formidable resource for the management of formal and explicit knowledge in various domains, such as language terminology and micro-mechanical engineering. The present paper focuses on two of these community types:

1. The first community type is centered around a collaborative editing tool – these are Wiki sites\(^1\), the most famous one being Wikipedia, especially in its English and German version – the user paradigm is that of an encyclopedia.

2. The second community type is centered around discussion boards, and the user paradigm is either a focused discussion around a certain subject, or of a Question&Answer (Q&A) type.

The authors conclude that some ideas and concepts coming from open Internet communities should be re-used in the learning and knowledge management endeavors done in the corporate and government world - these ideas will be listed as a proposals for lessons to be learned in the last chapter of the current paper.

Keywords: Social Media, Web communities, Open Knowledge Domain, Quality Control, Motivation, Tagging

I. Introduction

1.1 Research topic

The World Wide Web has changed a lot during the last years: It started as a hypertext medium used for research and technical documentation, and developed into a means for the publication of multimedia content, collaboration among hobbyists, and professional work.

In recent years the World Wide Web became a baseline for the so-called „Web 2.0“, the basic idea being that the content of a certain Web site is produced by its users rather than a centralized publication authority: Nowadays ubiquitous terms like “Social Media”, “Web 2.0”, “Web communities” and “Crowd sourcing” describe this change. But all these terms are very imprecise because the tangible Web sites behind them (such as Ebay, Youtube, Flickr, Facebook, Wikipedia, Sourceforge and many, many others) are very different one from another, and work in inherently different ways.

---

\(^1\) Wiki sites are collaborative editing sites such as [www.wikipedia.de](http://www.wikipedia.de) (which belongs to the Wikimedia foundation), and other collaborative editing sites such the ones using the pmwiki software ([www.pmwiki.org](http://www.pmwiki.org)).
When looking at a certain Web community, one of the most interesting questions is the question about the motivation behind the community members’ collaboration - what keeps the community in question alive, why do the community members work for this community at all? Answers can range from pure fun and personal fulfillment (Flickr, Facebook) to professional networking (Xing), earning money and trading (Ebay), and to the very tangible need for a certain piece of information (Q&A forums). Based on the two community types mentioned above (Wikipedia authors and the discussion boards used by many horology and photographic hobbyists), this paper addresses the question if lessons can be learned for the corporate learning and knowledge management done by private business corporations and government agencies.

1.2 Research questions

In this paper the following research questions will be analyzed in a generic way:

1. How can we recognize an Internet community, and what constitutes it? Who are their members? How can we classify the various communities which can be found in the Internet?
2. How do the community members collaborate? What Internet tools are used, and how are they being used? How are the rights to access and alter data handled? And what about quality control?
3. What attracts the community members to the Web site in question, and why do they contribute at all?
4. What could be the lessons to be learned for the corporate or government world? Can we, for example, find any “Web 2.0 learning resource management models” which corporations could re-use within their own business, even if their activities are not centered around the Internet?

1.3 Research goals

The research goals are:

1. The first two research questions mentioned above set the framework for the present paper, thus giving a general definition of the terminology, subject, and scope.
2. The third research question is meant to deliver a very short analysis of possible motivation factors driving successful Web communities. This is an important aspect if we want to apply similar models to the corporate world.
3. The fourth research question, finally, deals with the goal of the current research project, which is finding some useful lessons which the corporate world could learn. Here again, the two specialized community types mentioned above are considered more closely, and we focus especially on the details of their respective technological tools (Wiki editors, forum software).

Methodology: We did consult literature as, for example, listed in Hasler Roumois (2010), but we found only very little material that we found helpful to reach the above mentioned research goals. We decided, therefore, to proceed inductively and focus on “hands-on”-research in the Internet.

II. Constitution of Internet Communities

2.1 What constitutes an Internet community?

The answer to this question might seem easy, but it is not. In fact, we find that Web communities can be very heterogeneous, and it is quite difficult to find a common denominator. E.g.,

---

2 “Q&A forum” stands for “question and answer forum” – these are discussion boards using a question and answer paradigm – first one user asks a question, then other users offer their answers to the question. Q&A forums should not be confused with FAQs (frequently asked questions), which are mostly just questions and answers collected by one single author.
some communities gather around one specific Website (such as the Wikipedia users and authors), whereas others meet around several dispersed Websites.

We find evidence for the claim that a person who does not invest a certain amount of own work, and time, to a certain community is not really a part of this community: It would be, for example, rather strange to call all google users the “google community”, although google users actually do contribute to the improvement of the search engine by providing their search patterns. Furthermore, we face an observation problem: Website users who do not contribute explicitly are invisible, because they don’t leave visible traces on the Website.

The more we browse through the Internet, the better we see that many Internet communities lack genuine contributions and people - there are innumerable more or less unused Wikis and discussion boards in the Web. These Web sites clearly cannot serve as an example for the knowledge management within the corporate world. Therefore we also need some criteria to help us distinguish the “successful” online communities from the “unsuccessful” ones - after some deliberation we propose two general definitions:

- People making deliberate contributions on a website - thus sharing their knowledge with others - constitute an online community. The contributions can be manifold, but they have to be done knowingly (manual blog entrance, picture posted, improvement of a Wiki article, etc.), and they have to be done on a shared and explicit content base, such as a Website focusing on a certain subject matter.
- Communities that have regular content changes/add-ons and many user registrations are lively and are, therefore, considered successful. On the other hand, communities with very few content changes and almost no visible members are seen as static, ineffective and thus unsuccessful.

2.2 Who are the members of a given Internet community?

Sometimes members can be easily found via member profiles, sometimes not. Some members appear under their real names and photographs, others use nicknames, sometimes even several nicknames. In some communities, members seem to gather something like research reputation because they are quoted by other members. Some communities seem to have a rather quick member turnover, others not. And sometimes we can only “see” the members of an online community via their contributions, which can vary from superficial comments up to well researched articles.

We again suggest a generic definition:

- A member of a certain online community is anybody engaging within the community in a way that results in visible contributions over a certain period of time. As long as we can see that most contributions within a certain community come from human individuals, we shall regard the (real of fake) name behind these contributions as a community member.

2.3 How can we classify Internet communities?

There are some very well known and active Web communities producing a lot of content used throughout the world. Some of these communities have developed very rapidly, and they attract millions of users. Obvious examples are Wikipedia (an online encyclopedia), flickr (exchange and comments on photographs), del.icio.us (exchange of Web links through shared classification), youtube (Web videos), and ebay (selling and trading goods). When analyzing these examples, it quickly becomes obvious that Web Communities are manyfold and difficult to classify. The classification suggested in this paper focuses on the tools used by some of these communities and distinguishes between two types of tools:

- Wiki software, as used for online dictionaries, with Wikipedia as the obvious example.

---

3 Unused Web sites tend to disappear, so it is difficult to give stable reference URLs here, but such Web sites can be found via google via search terms such as “mechanical” “repair”, and “forum”.

4 This distinction was done on the basis of observed differences: A successful Web community such as the technical forums on www.uhrforum.de has usually at least 30 concurrent users, and the latest discussion contribution is not older than some minutes. Less successful Web sites such as most of the forums in www.watchlounge.com have few or no concurrent users, and the latest contributions are several days or weeks old, frequently remaining unanswered.
Discussion boards, as used for Q&A forums and diverse discussions. As example we have chosen the discussion boards used in the hobbyist domain of horology.

Below you will find a screenshot for both community types:

![Figure 1: Example of a Wiki community - the German Wikipedia](image1)

![Figure 2: Example of a discussion forum about micro-mecanical engineering](image2)

III. How do these communities work?

3.1 General observations

Within both discussion boards and Wikipedia, it is obvious that certain people gather around a common goal, and this “community” is then supported by the appropriate information tools.

These tools are mostly quite common, and their user paradigm is easy to understand – we have moderated discussion boards, usually with additional functionalities such as file exchange, and the functionalities offered by the Media Wiki software used by Wikipedia.

Both the Media Wiki and the discussion board software analyzed during this research offer quality control functionalities such as content rating and simplified peer review.

---

5 This screenshot comes from www.wikipedia.de, as seen on September 4, 2010.
6 This screenshot comes from www.uhrforum.de, and its mechanical sub-forums – the screenshot was also taken on September 4, 2010.
Within Wikipedia and successful discussion boards, the tools are designed to support the community goals directly – these tools are easy to understand, and a potential user can see quite quickly how these tools should be used and what the outcome will look like.

There is always a clear user paradigm – it is a dictionary in Wikipedia, and focused discussions around precisely specified subjects in the discussion boards. Furthermore, there is an easy integration of the various individual contributions into one “community content-pool”, which can be accessed by community members through specialized search tools such as search windows, unified content views, and links. Sometimes even fully automatic re-linking is offered7.

Interestingly enough, the content quality of a community tends to correlate with the look and feel of its Website and the tools used: We see flashy discussion boards with sparkling icons – they tend to attract users who write “flames”8, whereas more research-oriented users seem to gather around Wikis. Some of the Wiki content we find can only be qualified as top quality – it is well referenced and could easily serve as a basis material for a University course about natural sciences such as chemistry9.

3.2 Quality Control

Many communities offer high quality content which can be used as a valuable resource for professional work – the obvious example is Wikipedia. But some discussion boards also offer very good knowledge, too. How is this level of quality being achieved?

The access is totally free, but communities offering quality content are organized in a rather hierarchical way – everybody is invited to participate, but the contributions of participants are constantly scrutinized by other community members. Very often there is even a formal membership hierarchy with superior access and editing rights at the top. For instance, some Wikipedia articles (such as the one about George W. Bush on the German Wikipedia)10 have been blocked for anonymous editing by their administrators, in order to prevent unwanted updates by politically biased users.

Most Internet communities we see have grown anonymously - either community members did not have to reveal their real name or the names they gave were never really checked. However, most of these communities tend to have identifiable individuals at the top of their hierarchies. Very often, these persons effect a certain quality control by eliminating bad content quality or language deviations: discussion board administrators delete, for example, unwanted comments, they close discussion board threads that are not in line with the community goals, and sometimes they even exclude unwanted members from the respective on-line community. Interestingly enough, the initiative for these quality interventions usually does not come from the top of the hierarchy, but from simple community members complaining about the contributions of others.

In Wikipedia experienced and engaged community members overwrite false contributions. We have tested this by deliberately putting some false information in a rather minor and specialized Wiki article about the German photographic industry. It took another Wikipedia author only 3 hours to rectify this false information, whereas correct contributions remained unchanged11.

If we take a step backward and look for comparisons, we see that the Wikipedia methods are very similar to the ones used in the scientific world:

- **First, open access (everybody can write a contribution)**
- **Second, open discussions and assessment of people on the basis of their texts**

---

7 See for example the related links which are automatically shown on the bottom part of the interface of www.watchuseek.com – these are probably the result of automatic text re-indexing based on manual keyword lists. Or look at the discussion boards on www.uhrforum.de and their “Ähnliche Themen” feature: If the subject is about a failure of an ETA 2824 movement, then the discussion forum will automatically show other discussion threads about the same movement.

8 A flame is a rather insulting comment done in a discussion board, usually discrediting the discussion contribution of another board member.

9 See for example the chemistry pages on www.wikipedia.de, such as the pages about the basic elements, e.g. http://de.wikipedia.org/wiki/Kalium.

10 See http://en.wikipedia.org/wiki/George_W._Bush, then click on the “View source” source tab above – you will find a remark that the page has been blocked for editing.

11 The experiment was carried through on the German Wikipedia article about the Rolleiflex camera system (http://de.wikipedia.org/wiki/Rolleiflex).
• Third, people can gain a certain reputation within the community, and climb a hierarchy
• Fourth and last, quality control is effected by the people at the top of the hierarchy

3.3 Conclusions

We see a wealth of inherently different collaboration and contribution models. Some online communities are wide-spread and have no precise shape: Some communities gather around one very specific Web site, others find their focus on at least 2-3 specialized Web sites. We find different tools, different looks, different people, different languages and varying content quality.

Successful communities work around Websites which are easy to understand for their respective target audience, at least at the beginning when using the basic functionalities. Most of these Websites offer some kind of value to their new members, such as highly targeted information, precisely tailored to the user’s needs. Usually, this information comes out of an integrated information pool containing all of the different users’ contributions.12

Quality control is an important issue for all Web communities, if they want to stay alive and attractive. It is interesting to see how the different communities handle the issue – as sketched above some of these models eliminate potential content errors rather quickly and in a very effective manner!

Quality content can be modeled in two cycles: The first cycle focuses on the motivation to contribute, the second cycle outlines quality control on the contributions already done, both cycles are interlinked:

Cycle model with quality control and rise in motivation13

---

12 An integrated data pool usually consists of a database containing data entries such as Wikipedia articles or discussion board contributions, which can then be searched via an automatic text index such as the text search field on the Wikipedia entry page.

13 The current cycle model has been created by the authors of the current paper and is mentioned only for illustration purposes.
IV. What attracts the community members to the Web site in question?

There is no clear answer to this question - we suggest the following hypotheses:

- In many communities, people just publish for the fun of seeing their contributions online, and have other people comment on them. Or they want to help other community members with information, and maybe get the same kind of help in return. We also see a few exceptional cases where people just seem to struggle against each other, or try to impose their point of view on other community members, very much like in the real world\textsuperscript{14}.

- In comparatively few communities, money seems to be the main driver, sometimes supported by very sophisticated business models. A classic example would be the exchange corners found on many horology websites, where people contribute their classifications, pictures and content descriptions in order to get their watches sold\textsuperscript{15}. Other examples would be advertisements, sponsored links, or the open contributions of industry employees done within online forums, who are writing in order to defend, for example, the quality of the watch movements they are producing.

- There are many formal but non-monetary reward systems: discussion forum software can, for example, assign a little icon (stars) to the user name of every participant. These icons are usually based on the amount of articles that the participant has written within the forum and/or on the acceptance that his or her contributions have received from the other participants.

- In many communities, there is something like an “online prestige” – people are being constantly judged on the basis of their contributions. Active and valuable contributions are publicly praised by other community members.

- Some community members seem to spend a big amount of their personal time on their contributions; those members are usually very active and occupy a higher step within the community hierarchy.

Considering our observation sample, we summarize our observations and suggest the following conclusion:

\( \Rightarrow \) Online communities work if they offer some kind of “reward” to their users. This reward can be manifold: In most cases, it is only based on personal motivation such as the fun to see one’s own content online, the motivation to support a noble idea, or the desire to help somebody. The hope to get some added value for oneself is also an important driver. In some communities, money seems to be the principal motivator. Finally there is prestige: There are open praises by other community members, which are very often supported by automatic reward system.\textsuperscript{16}

V. Which lessons can be learned by the corporate or government world?

5.1 Pragmatic lessons can be learned

Based on our findings, it would be difficult to suggest some ready-made “Web 2.0 learning resource and knowledge management models” that private companies or government corporations could re-use directly within their own businesses.

However, we think that both private companies and government agencies could learn something from successful Web communities: It is obvious that we cannot copy blindly all their development ideas, but some of their tools and methodology should be taken as an inspiration for the

\textsuperscript{14} This behavior can sometimes be seen in the discussion boards analyzed for the present paper, or in the discussion pages on Wikipedia, but it is actually rather rare – it happens much more often in discussion boards about political issues.

\textsuperscript{15} See for example the sales and exchange corners on www.watchuseek.com or www.uhrforum.de.

\textsuperscript{16} Automatic reward systems are done in many different ways, the most obvious one consisting of small stars or tiny résumés and praises which are being published close to a certain user name, as you will find it in many discussion boards such as www.watchuseek.com (tiny résumés) and of course www.ebay.com (tiny stars). Furthermore, some systems allow their users to easily overview all the contributions which an author has written, which will result in further comments and praises.
learning content repositories and/or knowledge management models implemented within bigger organizations. Nowadays, the Web is driving the usability and innovation standards – the classic example for this is the use of a text search tool such as Google, which has become commonplace for everybody.

Ideas coming from open Web communities can be-reused within companies, and we would recommend an approach which is both pragmatic and open for inspiration. The “lessons to be learned” outlined in the paragraphs below illustrate what we mean:

**Think about motivation**

Successful Web communities include motivational interactions among humans, and they use good software paradigms – these two dimensions drive users to commit themselves to the Web community in question: the software tools are designed to enhance the users’ motivation, and the interactions among humans have an even more motivating effect.

**Have an integrated content pool**

The paradigm of “loosely coupled” Web 2.0 applications is very often mentioned in conferences and meetings\(^{17}\). However: most of the successful Community Web sites we see take a rather integrated approach - different types of information are gathered into one single content pool, and can then be searched quite effectively, and in a simple way.

**Have a low entry threshold**

Knowledge management tools must be self-explicable and easy to use. The basic recipe seems to be: take an easy user paradigm (online dictionary, Q&A forum), combine it with an intelligent and powerful text search tool, and you are almost there.

**Focus on quality control rather than on authoring restrictions**

Most companies still have a barrier between “the content publishers” (e.g., the instructional designers working on a corporate wording glossary) and “the readers” (e.g., the people reading the glossary). Therefore, quality control is mostly been implemented by an a priori limitation of the members who are allowed to contribute content. This can be justified in some cases (such as instructional design of media intensive learning content), but it doesn’t take into account some of the simple quality control solutions which can be found in the Internet.

**Develop a quality control model**

Rather than limiting authoring from the beginning, companies should focus on community support and automated quality tracking. E.g., quality issues could be observed by everybody and enforced via informal hierarchies based on trust. And the system should offer functions such as user tracking and content role back\(^{18}\). In short: a quality concept is needed and should focus on three aspects:

1) **granularity of intervention**: on which level do we exercise quality control – e.g., do we care about the language quality of every single sentence in our system, or is it enough if we reject and accept articles as a whole?

2) **intervention triggering and reaction time**: is it OK if a potentially incorrect contribution is allowed to remain online for some time before it is detected and corrected? Or do we need a quality control which avoids all potential errors beforehand?

\(^{17}\) See Timo O’Reilly, What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, 2004-2007. Tim O’Reilly and his findings about loose coupling and Web 2.0 are very often quoted at conferences and meetings.

\(^{18}\) “Content role back” is a functionality offered by database engines and Wiki tools – if errors have been entered on a Wiki page, then the Wiki page can be “rolled back” to its previous errorless state and content.
3) **production cost**: how much do we invest in quality? There is a trade-off between quality and cost: online systems that avoid all errors demand more resources because all contributions need to be checked for their quality before they are allowed to appear on the screen.

Giving up some centralized manual control does not necessarily mean a loss in quality – it is not by accident that the concept of peer review is the driving force behind quality Web solutions such as Wikipedia, and scientific journals, too.

**Forget about manual classification**

Knowledge management tools must be self-explicable and easy to use. The basic recipe seems to be: take an easy user paradigm (online dictionary, Q&A forum), combine it with an intelligent and powerful text search tool, and you are almost there.

Private companies and government agencies still have a tendency to think in terms of classification, which is, for example, enforced by rigid filing plans and or by category tags such as the ones used in Microsoft Sharepoint\(^\text{19}\). The main goal of classification is to have people find the content they are looking for: the content classifier attaches a metadata tag to a content object (e.g. to a document), and the content searcher will then use these tags to access the content object he/she is looking for. Therefore, classification tags are basically a communication means between the classifier (who assigned the tag) and the searcher (who uses the tag with the expectation to find the content he/she is looking for). The whole system makes only sense if the searcher interprets the tags in more or less the same way as the classifier did, which means that he or she would classify the content in more or less the same way as the classifier did.

But there is no such thing as a stable “perfect” classification system – classification is only about the facilitation of search, and it stands and falls with the searcher’s interpretation (and understanding) of the classification tree. We come to the conclusion that most of the successful Internet classification systems use a rather simplistic approach: they don’t look for perfect classification, they don’t go too far into the details, they define a classification system only to the degree that they have content to search for, they live with uncertainties and different interpretations, and they recognize the fact that classification trees usually look different from one language to another. In short, they stick to the “stuff” that everybody understands. Furthermore, automatic text indexes are far easier to use than the results of manual classification: Depending on the context, manual in-house classification should be reduced or dropped altogether.

**VI. Conclusions**

The various learning repository and knowledge management projects which are currently being done in government agencies and private companies tend to be more complex than one would imagine - they can fail or succeed for various reasons such as corporate culture, power balances and (missing) motivation to contribute or learn. Recipes coming from open Web communities should be interpreted with care in a given organizational context, but some of these ideas and concepts can and should be taken as an inspiration for the learning repository and knowledge management projects done in the corporate or government world!

---

\(^\text{19}\) See product description of Microsoft Sharepoint, as published by Microsoft on www.microsoft.com and in the various technical handbooks about Sharepoint. Sharepoint uses meta tags, which can then be re-used within the menus and database folders of the Sharepoint system. URL last checked on Sept. 15, 2010.
References

Book and paper references

Web communities in horology
[3] www.timezone.com, as observed since 2008, then between July and September 2010, has a main site and 2-3 sub-sites, such as the watch school and the tool kit site. Timezone is one of the biggest portals for the horology hobbyists, and it carries all the classical elements of such a collector portal. There is a special forum about watchmaking and repair issues.
[4] www.uhrforum.de, as observed since 2008, then between January and July 2010, has a main site and collects many discussion boards with q&a forums, some of them about watch repair. This Website is very widely used in Germanic countries such as Germany and Switzerland.
[5] www.watchuseek.com, and its forums, which can be seen on www.watchuseek.com/site/forums.htm, as observed since 2007, then again between July and Sep-tember 2010.
[6] www.watchlounge.com, as observed since 2008, then between June and August 2010: Watch lounge is similar to timezone and uhrforum above, but less successful – it carries a discussion forum and some additional functions.
[7] www.eta.ch, as observed between August 2009 and August 2010. ETA offers a very interesting customer service for its watch movements – business partners can order spare parts and movements, and for some movements such as the ETA 7750 there are even multimedia disassembly/assembly instruction videos (http://www.eta.ch/swisslab/7750/7750.html, as observed in September 2010).
[9] www.uhren-wiki.net, as observed since 2009, then between May 2010 and September 2010. This page has grown a lot over six months, but it is difficult to tell if this growth is due to a lively community, or to some paid centralized staff work.
[10] http://usawatchservice.com, as observed between July 2010 and August 2010: One of the bigger watch repair services in the Internet, with a price list for factory services.

Web communities in photography
[11] http://www.photo.net, especially its forums, as observed since 2003, then again for this research in summer 2010 – even minor photographic accessories such as tripods and ball heads are extensively discussed.
[12] www.nikonians.org, and its forums, as observed since 2005, then again for this research in July/August 2010. Nikonians is a Web page gathering Nikon users, and photographic accessories are also discussed.
[13] www.flickr.com, as observed since 2003, then again in September 2010. People post their photos, and have other people comment on them. A lot of high quality classical photography can be found on flickr.
[15] www.kenrockwell.com, as observed since 2007 is not a Web community, but a technical blog in the photography domain – very often quoted in communities.

Online dictionaries
[16] www.wikipedia.org, www.wikipedia.de and others such as www.fr.wikipedia.org: These are the famous Wikipedia dictionaries belonging to the Wikimedia foundation (www.wikimedia.org). These Websites have been observed very closely between June and September 2010, and before since 2003. It is quite astonishing to observe the rapid growth of the German Wikipedia as to both quality and quantity of content – in some areas of knowledge, the German version is more developed than the English one. There are many Wikimedia sites on specialized subject matter areas and the offer of educational content (http://de.wikiversity.org/) is developing.