

Managing Information Quality in Virtual Communities of Practice

Lessons learned from a decade's experience
with exploding Internet communication*

Andreas Neus

andreas.neus@de.ibm.com

IBM Unternehmensberatung GmbH, Am Sandtorkai 73, 20457 Hamburg

Abstract

This paper examines why the rise of computer-mediated communication, driven by dramatically lowered cost, creates new structural problems from an information quality standpoint. We review how the new economics of information enable a new paradigm of collaboration. Spearheaded by the Open Source community, this evolutionary approach to collaborative content creation is a way to address information quality in virtual communities of practice. Based on experience gained in community projects, a few simple steps toward improving the quality of information in virtual communities are presented and illustrated.

1 Introduction

Virtual Communities of Practice are becoming more important as a means of sharing information within and between organizations.¹ While there is widespread agreement that controlling the quality of information exchanged is critical to the community's success (triggering either a vicious or a virtuous cycle) much of the published work has been focused around advances in technological or theoretical concepts for semi-anonymous² mega-sites such as amazon.com or ebay.com,³ with users po-

tentially in the millions. However, those may well be the wrong kind of approach to address the issue of information quality for the increasingly important, more closely knit Communities of Practice that companies are initiating to speed up cross-functional knowledge sharing. As a result, we have observed that for smaller-scale virtual communities, the issue of information quality is addressed in one of three typical ways:

1. *It is ignored altogether:* everyone posts and there is little structure or quality control, making it difficult to re-use knowledge.
2. *It is over-controlled, resulting in tunnel vision:* every communication addressed to the community has to be approved by a leader or moderator first.
3. *It is buried inside an unwieldy tool:* the lively discussion that is the essence of a Community of Practice is squelched by a tool that was designed for information storage and retrieval, not for discourse and collaboration.

This paper attempts to give a general overview of some key issues faced by practitioners addressing the quality of information in virtual communities and showcases some solutions tried over the years on the Internet to solve information quality problems. Finally, some experience gained in projects charged with planning and setting up virtual communities over the course of the past decade is distilled into a few simple recommendations for raising the information quality in virtual Communities of Practice.

*Citation details: NEUS, A. (2001). Managing Information Quality in Virtual Communities of Practice. In: PIERCE, E. & KATZ-HAAS, R. (Eds.) Proceedings of the 6th International Conference on Information Quality at MIT. Boston, MA: Sloan School of Management.

¹See THOMAS STEWART & VICTORIA BROWN (1996).

²if only due to the sheer number of contributors.

³See JAKOB NIELSEN (1999) for a good overview.

2 Networks vs. Hierarchies: The appeal of using Communities of Practice

Traditionally, we have used organizational systems based on hierarchy and authority to identify high-quality information. But with the world around us accelerating in its need for both time-critical and high-quality information, these traditional means are often no longer sufficient. For whereas traditional, hierarchical organization structures are very good at getting tasks done in a “divide and conquer” paradigm, the success of companies today increasingly depends not on *dividing the work*, but rather on *sharing the knowledge*. Yet for facilitating the free exchange of knowledge, networks are inherently better equipped than hierarchies. From an information sharing standpoint, a hierarchical, tree-like organization is a worst-case scenario because it is a collection of bottlenecks: There is only one “official” path between any two nodes in the graph and the likelihood of people sharing information can drop as a function of their distance in the corporate org-chart.

Faced with this dilemma, advanced companies have started overlaying their primary “command-and-control” structure not only with a subject-driven layer, forming a matrix, but also with “Communities of Practice”: semi-formal networks of practitioners who exchange information on a common subject or problem of interest. They are alternately called “Competencies”, “Communities”, “Knowledge Networks”, “Special Interest Groups”, etc. and run all over the organization’s chart (and sometimes even beyond a single organization) to facilitate the exchange of information and lessons learned among those who are dealing with a common set of problems or issues. That a *network* is extremely efficient for passing information is a phenomenon which has been studied and confirmed numerous times under the names “small world effect” or “six degrees of separation”⁴. But with a network where every member is also an instant publisher, a new challenge emerges.

⁴See MILGRAM (1967).

3 Problem Outline: The vanishing cost of communication

Before the proliferation of the Internet technology, there was a simple filtering system that kept the quality and relevance of transmitted information mostly above a certain threshold. It was called *cost*. Because every copy of information and every movement of that copy was tied to physical matter, it always incurred real cost that someone had to shoulder.

Therefore, only such information which was deemed by *someone* to be worth spending money on copying and distributing, had a fighting chance of ever being seen by more than a trivial number of people. Yes, there have always been tabloids whose information quality, when measured against objective criteria, did not do terribly well, but in terms of a *fitness for use* definition, this content still had to have a certain “quality”, because people actually paid to read it. And while there have always been “nut cases” creating their own newspaper from their flat on a copying machine, your chances of ever coming across that content fell dramatically (probably to the tune of N^2) with your physical or social distance from the source.

So traditionally, a piece of information had to pass through several layers of weeding, selecting and editing before it would get a fighting chance to come to the attention of a wider audience. Publishers and advertisers were effectively making a bet, with their very real money, on the quality (again in terms of *fitness for use*) of the content being suitable for its intended recipients, because every copy of information incurred real costs. Therefore, it *used* to be good business to limit the proliferation of your information to coincide with your intended audience, i.e. the target segment or market. And the better you were at addressing only your target segment with the information (and in turn designing the information to be relevant to the recipients) the less money you wasted. This incentive brought about editors, proofreaders, market analysts, etc.

The Internet, however, dramatically lowered the cost of copying and distribution – to practically zero. The last time such a major drop in the cost of information proliferation occurred was probably

Gutenberg's invention of movable type⁵. As with Gutenberg's invention, this more recent drop in the economic cost of distributing information has created major shockwaves around the world. The success of the Internet as the undisputed global communications medium of the future and the rise (and often fall) of countless businesses built around it are an example of this kind of shockwave.

4 Spam: A New Word for Irrelevance

The vanishing cost of communication, coupled with commercial interest, has given rise to a bane of modern society, called "Spam"⁶. Whereas today "Spam" is mostly understood to mean "unsolicited commercial e-mail" or "unsolicited bulk e-mail" (the *electronic* junk mail that clogs your mailbox), it originally came up in the early nineties as a term for "massive crossposting": posting the same article to several thematically unrelated newsgroups on Usenet. When companies first stumbled upon the mostly academic global discussion network that drove the growth of the Internet in the pre-Web years, some were lured into abusing the infrastructure for posting commercial messages into thousands of newsgroups by the same economics that sustain e-mail spamming today: The cost incurred *by the sender* for sending information to be viewed by one recipient or 1,000,000 recipients, is practically identical, and in both instances negligible. This is because the bulk of that cost is carried by the providers of the infrastructure and the recipient, *not by the sender*. As Vint Cerf puts it:

Spamming is the scourge of electronic-mail and newsgroups on the Internet. It can seriously interfere with the operation of public services, to say nothing of the effect it may have on any individual's e-mail system. ... Spammers are, in effect,

⁵Not only did Gutenberg make it possible for Luther to copy his translation of the Bible, something that the copyshops of the day in central Europe - catholic monasteries - would not have approved of, it also created a whole new economy around typeset books and later newspapers.

⁶Named after the Monty Python sketch featuring the recurring canned meat product. Crossposting in the newsgroups amounted to reading the same message over and over and over again. See ERIC RAYMOND (1993) for more information.

taking resources away from users and service suppliers without compensation and without authorization.

Certainly the most famous first instance of low-quality information massively posted to Usenet was the "Green Card Lottery" Spam perpetrated by two lawyers who decided to advertise their services to thousands of newsgroups simultaneously, creating a huge backlash by people who found their time and their resources abused. At the time, of course, network bandwidth and storage was much more expensive than today and a much higher proportion of people were on the Internet through dialup connections which had to be paid by the minute. But even today, the most precious resource is still very much threatened by low-quality information: *human attention*. With communication increasing (and a lot of economic interests around) the "signal-to-noise ratio" dropped rapidly on Usenet newsgroups and other forms of virtual communities such as web-based discussion boards, chat fora such as IRC or instant messaging services such as ICQ or AIM. For example, the number of Spam postings to Usenet each month skyrocketed from below 100,000 to 1.8 million in just two years between 1995 and 1997.⁷ This deterioration of information quality, particularly on Usenet over the years, has caused many people to abandon public newsgroups and seek refuge in other, less open virtual communities, because finding relevant and high-quality information had become a hunt for the proverbial needle in the haystack. This information overload is a symptom of what can no longer seriously be called the *information society*. Rather, it could be more appropriately called an *attention deficit society*.

What information consumes is rather obvious: It consumes the attention of its recipient. Hence a wealth of information creates a poverty of attention.

HERBERT SIMON
(*Nobel Laureate Economics*)

It is this scarce resource, the *human attention*, which we must learn to better manage and direct toward the high-quality, relevant information in the

⁷See the Cancelmoose Page at <http://www.cm.org/> for details.

exponentially growing haystack of low-quality information in collaborative environments.

5 Brooks' Law: Quality as the result of a single mind's integrity

These information quality problems, found in many virtual communities today, appear to confirm the old saying "Too many cooks spoil the broth". A modern and perhaps more "scientific" version of this notion has become known as *Brooks' Law*, named after FREDERICK P. BROOKS, author of the classic book "The Mythical Man-Month".⁸

BROOKS talks about the inherent complexities of coordination and states that as the number of involved programmers N rises, the work performed also scales as N , but the *complexity* and *vulnerability to mistakes* rises as N^2 , in accordance with the number of communication paths required to coordinate the contributors. To achieve quality, BROOKS therefore recommends a *minimum* of contributors: "*Conceptual integrity in turn dictates that the design must proceed from one mind, or a very small number of agreeing resonant minds.*"⁹ Conversely, Brooks' Law predicts that "*a project with thousands of contributors ought to be a flaky, unstable mess*", as ERIC RAYMOND put it.¹⁰

6 The new paradigm: Collaboration

Looking at some of the more prominent results of virtual collaboration, it becomes obvious that Brooks' Law cannot be the only force at work here. It used to be common knowledge that high-quality *software* could only be produced by a handful of highly skilled experts who are organized in the dedicated, hierarchic environment of big companies and headed by professional management.

However, during the 1990s, a radical new paradigm of collaboration, enabled by Internet technology,

emerged, which seems to violate (or maybe naively ignore) Brooks' Law. This new paradigm has become known as the *Open Source development model*¹¹ and has brought about the creation of the successful free Linux operating system, which has been called "the impossible public good". Public Goods Theory predicts that a product which can be used by anyone, irrespective of whether they contributed to its creation or not, would never get created in the first place as everyone would attempt to free-ride. Instead, Public Goods Theory postulates that something like Linux could only be created with public money or by a government. But the dramatic drop in transaction costs suddenly allowed for the collaborative creation of such high-quality *software*.

Common knowledge still holds today that high-quality *information* can only be produced by a handful of highly skilled experts who are organized in the dedicated, hierarchic environment of universities or research centers and led by professional management. But by using the same dynamics that made the "impossible" open source goods possible, people are already busily attacking this notion as well, suggesting that a loosely knit network of skilled amateurs can produce comparable or better quality information in a collaborative paradigm than traditional solitary authors, institutions or publishers are able to create.

As an illustration, consider the discussion around *PublicLibraryOfScience.org*, an initiative that is calling on journal publishers to hand control over published articles back to the scientific community after 6 months. Publishers have typically argued that information quality and integrity of published research can only be assured if they remain in exclusive control of the information. Practice, however, tells a different story: When some previously published articles were moved to public Internet repositories, several errors that had gone undetected during the original publishing were found and corrected, thereby *increasing* the quality of the information. As David Lipman, director of the National Center for Biotechnology Information, states: "*The more eyes to look at it and fingers trying to work with it, the more things you can find.*"¹²

⁸BROOKS (1975).

⁹Brooks (1975), chapter 4.

¹⁰See Revenge of the Hackers, in Raymond (2001).

¹¹See *Open Source Initiative* at www.OpenSource.org

¹²See the Scientific American article by KAROW (2001).

7 Linus' Law: Quality as the result of massive collaboration

How was the new paradigm, involving massive collaboration, able to overcome the limits postulated both by Brooks' Law and Public Goods Theory? The explanation, again, lies in changing economics of information and has been called "Linus' Law", honoring Linus Torvalds, the former computer science student who spearheaded the Open Source development model. Linus' Law is usually stated in its informal version, which resembles Lipmans statement above: *Given enough eyeballs, all bugs are shallow.*¹³ The key to the success of the collaborative development model is based on the lowered transaction cost for information, allowing the separation of the *identification* and the *solution* components of quality problems and spreading both tasks over a much, much larger population than could sensibly be done in traditional hierarchic approaches. Says Linus: *"Somebody finds the problem and somebody else understands it. And I'll go on record as saying that finding it is the bigger challenge."*

8 The Wiki Concept: Quality is what survives evolutionary pressure

The Wiki¹⁴ concept is an example of taking aforementioned collaboration paradigm to its extreme by practically eliminating any transaction cost in changing or correcting information. On a Wiki website, *anyone* can view and edit *any* page, without any prior clearing process by an editor or moderator. There is nothing to stop a malevolent user from deleting passages, or even whole pages, of existing information, or just adding complete nonsense. At first encounter, especially in the context of information quality, this concept appears to be a recipe

¹³The official version is as follows: *Given a large enough beta-tester and co-developer base, almost every problem will be characterized quickly and the fix obvious to someone.* This emphasizes the separation between characterization and solution of quality problems as a means to achieve greater efficiency. See RAYMOND (2001).

¹⁴For more information, see <http://www.c2.com/cgi/wiki?WelcomeVisitors>

for dramatic failure, an information quality disaster just waiting to happen.

Interestingly, we are still waiting for the disaster to happen – and it is nowhere in sight. The key is that although *any* user can change any page, the changes are stored in a log and *any other* user can review that log and instantly undo any change that he or she does not approve of. Using this deceptively simple safety net, the Wiki concept can be a very powerful accelerator for collaboratively creating and improving information.

As one example, consider *Nupedia.com*, a project dedicated to creating a freely available¹⁵ encyclopedia online. In accordance with their goal of high information quality, Nupedia.com adopted the traditional review process of publishers, where a (volunteer) author would first write an article and then submit it to Nupedia for review – a cumbersome process that resulted in only very few articles being contributed. When Nupedia.com still had only 20 articles to show for 18 months of operation, the founders realized that they had a problem and looked around for a solution. They found the Wiki concept and decided to start a complementary site, *Wikipedia.com*, as a hot-bed for collaboratively creating and improving articles. The best of these articles would then undergo the rigorous review process to become part of Nupedia.com. Something clicked and in merely 6 months from January 2001 until July 2001, Wikipedia.com has generated over 6000 articles, including many of very high quality, using this extreme interpretation of the collaborative paradigm. Instead of falling victim to vandalism, as might be expected, the site's wide-open concept quickly turned it into a thriving generator of information.¹⁶

What drove this astonishing result? In my view, the key to the demonstrable success of the Wiki concept is based on two pillars. The first is the elimination of practically all transaction costs for collaboration. Instead of informing an editor of a change you'd like to see and talking him into accepting it (possibly taking many exchanges back and forth), a Wiki system lets you make the change yourself, on the spot, with minimum effort. The second pillar is the creation of an *artificial information economy*

¹⁵Nupedia.com content is licensed under the GNU Free Documentation License.

¹⁶For more information, see www.kuro5hin.org/story/2001/7/25/103136/121

as a context for collaboration, which discourages low-quality or offending input, *because it is much "cheaper"*¹⁷ *for person B to undo the low-quality change that person A caused, than it is for person A to cause it.* This process weeds out low-quality information in an evolutionary paradigm. As Richard Dawkins puts it: *"Life is the result of the nonrandom survival of randomly varying replicators"*. The evolutionary paradigm has been demonstrated to be so potent that it can create order out of apparent chaos even based on *random* mutations, given a *nonrandom* selection. The evolution in terms of the collaborative Wiki concept has the additional benefit of the changes being anything but random.

As an analogy, imagine a new method by which any passer-by could undo a night's work by a graffiti "artist" simply by snapping his fingers, if he thought that the house looked better the way it was before, without the new "decoration". What would the effect be?

1. There would be very little incentive for people to create low-quality graffiti "content", because they have to labor for hours, only to have their effort casually nullified by the next person to walk by.
2. The content that *survives* review by many people over a long period of time is likely to be of high quality, in the sense that there is widespread agreement that the wall looks better *with* the new graffiti than it did without it.

9 "It's the economy, stupid!"

A key problem with hierarchic approaches to information quality is that they don't scale well – and you have the issue of who chooses the editors, peers or raters. Instead, the solution may lie in creating an information economy that uses an evolutionary paradigm to *grow* and *evolve* high-quality information collaboratively, rather than to have a single author *construct* it. In this information economy, there should be a high incentive for contributing and maintaining high-quality information and a disincentive for contributing poor-quality information. The challenge then becomes one of creating

¹⁷"Cheaper" economically, i.e. in terms of low cost in attention, time or reputation.

such an information economy that produces high-quality information. Some important cornerstones for this to work are the following five factors:

1. Accountability for contributions as a basis for reputation
2. A thematic focus and "culture" for high quality contributions
3. A sense of trust and identity through personal profile pages
4. A common memory or knowledge repository which is developed in collaboration
5. Membership criteria to keep the level of discourse high and on topic

10 Conclusion

The Internet has provided us with the means to effectively collaborate across time and space with vanishing transaction costs. As the success of open source software such as Linux has proven, this new paradigm has the potential to break through well-established limits. Virtual Communities of Practice have the chance of using these same economics to redefine how high-quality information is created and shared in an organization. But one of the greatest obstacles to adopting this new paradigm are traditional notions in our own minds of how information quality is achieved, limiting our thinking about the non-linear potential of collaboration.

I believe that companies, which are today faced with increasingly well-informed customers, need to actively embrace and support virtual Communities of Practice as a way to bypass information bottlenecks, to speed up internal knowledge creation and sharing, and because they need to keep up with their increasingly well-informed customers. Or, as The Cluetrain Manifesto postulates: *Because markets, unencumbered by corporate bureaucracy and the need to ask permission at every turn, are learning faster than organizations.*¹⁸

¹⁸LEVINE et al. (2000).

11 Outlook: Networks of knowledge

The future of addressing information quality in virtual communities may well lie in supporting collaboration by mapping and analyzing the *underlying social networks*, revealing the now mostly invisible links between people and communities. To some extent, this is already being done, i.e. by the Google search engine.¹⁹ Some advanced research on analyzing social networks to assess information quality is also carried out in the CLEVER Project at IBM's Almaden Labs.²⁰

The essence of the mapping approaches is to analyze the micro-decisions made by people pointing to resources and to aggregate this information over a large number of people to derive information quality measures based on implicit human decisions. Besides opening up an exciting new way to tap human expertise for determining quality, these approaches also bring up a new class of challenges, especially in the area of privacy, that have only recently received attention.

The essence of the mapping approaches is to analyze the micro-decisions made by people pointing to resources and to aggregate this information over a large number of people to derive information quality measures based on implicit human decisions. Besides opening up an exciting new way to tap human expertise for determining quality, these approaches also bring up a new class of challenges, especially in the area of privacy, that have only recently received attention.²¹

Appendix: A brief community cookbook

Feedback loops are important for efficiently producing high quality information. Therefore, the system that allows for better feedback is the one with the potential to provide the better quality. The Internet technology provides us with such a system for much faster – and much broader – feed-

back loops than were previously possible. But the new technology is only the *enabler* for the virtual Communities of Practice, facilitating collaboration and feedback. The *driver* for the new paradigm is the organization, trust, commitment and interaction between the community members. This “soft” or *human side* often turns out to be the trickier part. Therefore, here are a few suggestions to help managing these softer issues.

Here are five simple steps you can take to support a strong collaborative culture and improve the quality of information in your virtual Communities of Practice.

1. Accountability: The prerequisite to reputation

When someone makes a change to the knowledge, this change must be tracked so that there is accountability, i.e. the actor can be adequately credited with the cost or benefit of the change, and in order to allow selective reversing of changes. This way, other community members can make intelligent choices regarding how they spend their attention with respect to this user.

Impact on information quality: As an example, I may decide whether or not to review a contribution or a change to the pool of knowledge depending on whether the actor has produced high-quality information in the past. While “blacklisting” is difficult in a possibly pseudonymous virtual environment, accountability is still very important for “whitelisting” – because a positive reputation is an asset that the owner has an incentive to protect.

Recommendation: Addressing accountability can range from something as simple as making sure everyone has to log in with a username and password before contributing and keeping a log-file to highly complex rating and reputation systems.

2. Focus and culture: A community charter

A charter including clear rules on what behavior is expected and what may be done with the content created. One key to achieving high quality is to realize that a good community of practice is a self-

¹⁹www.google.com

²⁰See KLEINBERG (1997) and KUMAR et al. (1999) for more details.

²¹See ADAMIC & ADAR (2001).

regulating entity that will improve the information quality by peer-pressure. Traditions and customs governing what kind of information is accepted will develop and these will be enforced by the members as part of the community's culture.

Impact on information quality: A charter sets the tone for the discourse in the community of practice. It should be created jointly with the community members to ensure adoption. A strong culture on what kind of quality is expected from the information will go a long way to ensuring, via peer pressure, that the quality of contribution remains high. Explicit rules on re-use of posted information outside the community are necessary for a feeling of trust and comfort to develop, where people are willing to ask "stupid" questions or go out on a limb. I.e. "no external re-use without asking permission from the author first" could be such a rule. Because a community's membership changes, its activities ebb and flow, and its leaders change, a charter is a good way to provide a scaffolding that does not depend on individual members.

Recommendation: Write at least a draft charter for the community that sets a standard for behavior, expected quality, and in which circumstances information created in the community may be used outside it. For inspiration, have a look at some examples of such manifested traditions, like the famous "Netiquette" texts that are posted to *news.announce.newusers* group.²² They explain customs and the reasons behind them to new members. Other examples of standards for quality-checking information are the humorous "Crackpot Index"²³ circulated in the *sci.physics* newsgroups and the "Gullibility Virus" warning.²⁴

3. Trust and Identity: Personal Profile Pages.

Trust is a problem in virtual environments. The trust that forms very easily in face-to-face meetings is much harder to achieve when all you know about the other members is their e-mail-address. A key step to creating the trust and sense of identity nec-

²²See *news.announce.newusers* FAQs in the references below.

²³See JOHN BAEZ (1998) for details.

²⁴This text by ROBERT HARRIS (2000) warns readers not to become multipliers for false information – in a very original way.

essary for a thriving community of practice can be taken by providing a personal profile page for each member and encouraging its use. This page should include a picture, some self-description, and room for (links to) other resources relevant to the member's professional and maybe personal life, which they wish to share.

Impact on information quality: These profiles facilitates the exchange of ideas and the creation of trust between members, as they allow people to get a better concept of the other person's expertise and interests. Resources linked to from a profile are directly available to the other members, without having to ask and wait for an answer. It ensures that members get a good idea what the skills in their community of practice are and who to turn to with which kind of question.

Recommendation: Make sure your community tool supports such personal profile pages and encourage members to use them. When a few key members present themselves in this fashion, the others usually follow suit.

4. Collective Memory: FAQs as efficient knowledge repositories

FAQs are a very powerful way to distill lessons-learned in virtual communities. They were originally invented out of sheer necessity: With the rising popularity of the Usenet in the late 80s came a problem: Newsgroups, which were home to communities discussing their chosen subjects, usually at a very high level of expertise, were faced with an influx of new users almost on a daily basis. This caused the ongoing discussion to be brought back down to basic, beginner's questions frequently, as a stream of new users, unaware of the discussion's history, asked the same questions over and over again. The regulars realized that they could only solve this by compiling the answers to those *frequently asked questions* in a file that could then be referenced in reply to those questions. Out of this necessity, accidentally, a very powerful didactic tool was born: In contrast to practically all other forms of documented knowledge, the FAQ is structured *not* from the perspective of the "knower", but is collaboratively created over time and structured from the perspective of the "knowledge seeker". It is therefore a much more efficient way of educating

people and bringing them to a common level of understanding than was available before. In a way, it is accelerated education. An FAQ is usually maintained by one or more people who have an interest and some expertise in the subject. Those people become a natural focal point for both questions and new answers regarding the subject at hand, starting a virtuous cycle.

Impact on information quality: By providing a focal point for the community's knowledge on given subjects, people can stop reinventing the wheel and instead focus on creating the best wheel for everyone. FAQs of active communities are typically of a much higher quality than i.e. Textbooks, simply because there are so many more eyes for scrutiny and the combined know-how of the community helps to polish the text over time, instead of having just a few authors write a text that only gets revised every other year at the most.

Recommendation: Create an infrastructure for maintaining FAQs and encourage members to start FAQs on their pet subjects. This way you will quickly seed the creation of efficient knowledge repositories that can quickly grow from half a page to several dozen pages in size and allow you to easily capture lessons-learned in the community of practice.²⁵

5. Membership Criteria

A community lives off its peers. If you get the right people together, you start a virtuous circle that draws in more of the right people simply by word-of-mouth. If you let everyone in indiscriminately, you soon have an unfocussed group of members, dropping information quality (in the fitness for use sense) and the experts, whose discussion you wanted to tap into, will drop out of sight again. This is why good clubs have bouncers who perform a very important *function* of quality control (which is not to say they always do a good *job*).

Impact on information quality: The impact of controlling membership is very straightforward: Having the right members in the community goes a long way to ensuring a good signal/noise ratio. This is doubly important as few experts wish to waste their

time in a community where they feel they do not have a lot to learn themselves, but rather always serve as unpaid teachers to the rest.

Recommendation: Think about some kind of barrier to entry. It does not need to be high, just something that keeps people with only a passing interest out. This can be as simple as asking people to send a CV or just give a few statements about why they believe they would make a valuable contribution to the Community of Practice. This encourages a self-selection that will increase the quality and focus of your community's members. For an example, observe the membership application in Howard Rheingold's successful "BrainStorms" community.²⁶ Managing the quality of members is possibly the most important single aspect and deserves thought. While doing this, consider the whole lifecycle: Individual members will become more and less active over time. Have a policy for weeding out people who have abandoned the community to avoid the sense of anonymity that comes with a community having too many members with whom no-one has communicated in a long time.

These steps should bring your Community of Practice closer to becoming a thriving, collaborative source of high-quality information – and they are largely independent of the underlying technical infrastructure employed.

Keep in mind that a community is a social creature that cannot be "created" in a traditional sense. Rather, you need to provide the right context for a community to prosper. Therefore, you need to be very careful how much control you seek to exert. Try to let the community organize itself to the greatest degree possible, rather than trying to micromanage it. Community leaders will emerge naturally: Those members who are perceived as the right mixture of being very knowledgeable, accessible and *active* in the continuing dialogue. Also, even in work-related communities of practice, "off topic" discussion should not be squelched, but accepted as a the necessary "social lubricant" that any efficient knowledge network needs. After all, *knowledge is human*.

²⁵To get a feeling for the vast amount of high-quality information thus captured on Usenet over the years, visit www.faqs.org

²⁶See Howard Rheingold at www.rheingold.com

References

- [1] Adamic, L. A & Adar, E. (2001). *Friends and Neighbors on the Web*. Web-site: <http://www.parc.xerox.com/istl/groups/iea/papers/web10/>
- [2] Baez, J. (1998). *The Crackpot Index*. Web-site: <http://math.ucr.edu/home/baez/crackpot.html>
- [3] Brooks, F. (1975). *The Mythical Man-Month*. Reading, MA: Addison-Wesley.
- [4] Cancelmoose (1997). *The Cancelmoose Homepage*. Web-site: <http://www.cm.org/>
- [5] Euro Cauce (2001). *Euro Cauce*. Web-site: <http://www.euro.cauce.org/en/index.html>
- [6] Harris, R. (2000). *The Gullibility Virus*. Web-site: <http://www.virtualsalt.com/warning.htm>
- [7] Karow, J. (2001). *Publish Free or Perish*. In: Scientific American, April 2001. Web-Site: <http://www.scientificamerican.com/explorations/2001/042301publish/>
- [8] Kleinberg, J.M. (1997). *Authoritative Sources in a Hyperlinked Environment*. IBM Research Report RJ 10076
- [9] Kumar, S. R., Raghavan, P., Rajagopalan, S., & Tomkins, A. (1999). *Trawling the Web for emerging cyber-communities*.
- [10] Levine, R., Locke, C., Searls, D. & Weinberger, D. (2000). *The Cluetrain Manifesto: The end of business as usual*. Cambridge, MA: Pegasus Books
- [11] Milgram, S. (1967). The Small World Problem. *Psychology Today*, May 1967.
- [12] news.announce.newusers FAQs. Web-Site: <http://www.ibiblio.org/usenet-i/groups-html/news.announce.newusers.html>
- [13] Nielsen, J. (1999). *Reputation Managers are Happening*. Web-site: <http://www.useit.com/alertbox/990905.html>
- [14] Open Source Initiative (2001). *The Open Source Initiative (OSI)*. Web-site: <http://www.opensource.org>
- [15] Raymond, Eric S. (1993). *The New Hacker's Dictionary*, 2nd ed. Cambridge, MA: MIT Press
- [16] Raymond, Eric S. (2001). *The Cathedral and the Bazaar*. Cambridge, MA: O'Reilly & Associates (Web-Site: <http://www.tuxedo.org/esr/writings/cathedral-paper.html>)
- [17] Raymond, Eric S. (2001). *The Jargon File*. Web-site: <http://www.jargon.org> (The online-version of the New Hacker's Dictionary).
- [18] Resnick, P. et al. (1994). GroupLens: An Open Architecture for Collaborative Filtering of Netnews. *Proceedings of the ACM 1994 Conference on Computer Supported Cooperative Work*, Chapel Hill, NC: Pages 175-186
- [19] Rheingold, H. (2001). *The Virtual Community*. Web-site: <http://www.rheingold.com/vc/book/>
- [20] Rheingold, H. (2000). Join the BrainStorms Community? Web-site: <http://www.rheingold.com/community.html>
- [21] Sanger, L. (2001). Britannica or Nupedia? The Future of Free Encyclopedias. Web-Site: <http://www.kuro5hin.org/story/2001/7/25/103136/121>
- [22] Stewart T. & Brown, V. (1996). The Invisible Keys To Success – Communities of practice are where learning and growth happen. *Fortune*, August 5, 1996
- [23] WikiWiki (2001). Web-Site: <http://www.c2.com/cgi/wiki?WelcomeVisitors>

Suggested further reading

- [24] Brown, J. S. & Duguid, P. (2000). *The Social Life of Information*. Cambridge: Harvard Business School Press.
- [25] Kim, Amy Jo (2000). *Community Building on the Web: Secret Strategies for Successful Online Communities*. Berkeley, CA: Peachpit Press
- [26] McKeon, D. (1997). *Moderated Newsgroups FAQ*. Web-site: <http://www.faqs.org/faqs/usenet/moderated-ng-faq/>

- [27] Oram, A. [Ed.] (2001). *Peer-to-Peer: Harnessing the Power of Disruptive Technologies*. Sebastopol, CA: O'Reilly & Associates.
- [28] www.faqs.org: A vast resource of Usenet-community generated content, including Netiquette, Emily Postnews, Newusers Introductions, FAQs, RFCs, etc.
- [29] www.isoc.org: The Internet Society. An authoritative source for information on *The Internet*, significantly higher quality than most other sources.