The IASSIST Wiki – a Copyleft Solution for the 4th Decade

"All together now...." The Beatles

by Oliver Watteler [1]

Introduction

Entering its fourth decade, IASSIST has set out to plan the future. According to the Strategic Plan for the period until 2009, the three main strategic directions are Education, Outreach and Advocacy.[2] What do these three points have in common? They are dealing with the distribution of information. IASSIST currently publish a magazine, maintain a website and run an e-mail list, which includes logs that are searchable, thereby realizing one of its prime objectives as described in Article III of the Constitution that it wants to "foster international exchange and dissemination of information regarding substantive and technical developments related to social science machine-readable data."

There is no question concerning the importance of information. Thus, Donald Waters wrote in 2002 when talking about the relationship between researchers and archives:

"There are many dimensions to the good to be achieved, but two of them merit special mentioning. On the one hand, there is the joining together by scholars and the agents of education—universities, libraries, scholarly societies, and publishers—in serving the common interest of future scholarship by keeping good, or preserving, the digital resources now being created. On the other hand, there is the research and learning thereby made possible, which are the indelible marks of a good scholar. In other words, good archives make good scholars."[3]

If one continues on this line of thought it is logical that good archivists and good data librarans are the basis for good archives and that they are well trained specialists that keep up their knowledge e.g. through participating in the activities of IASSIST.

One possible way of sharing information is of course the Internet. There are numerous information gateways on the Web, for example:

- the Virtual Training Suite[4], which was presented by Heather Dawson at the IASSIST Conference in 2001,

- the Social Science Information Gateway[5], both being part of the Resource Discovery Network in the United Kingdom, or

- the Vascoda[6] portal in Germany.

and IASSIST itself which revised its site only recently.

Most of the gateways use state of the art technology like "harvesters" sifting through the enormous wealth of information on specific topics that can be found on the web. They are also encountering sites maintained by individuals and institutions that hold bits of information and hyperlinks of particular interest, which means that they are relying on existing content. But what if you want to foster information exchange on an international level? Link again to the same sites as so many other pages have before?

The Wikipedia offers an interesting concept of a website maintained through the joint effort of a dedicated community, much like the members of IASSIST.

Wikipedia – what is it?

Most members of IASSIST will probably know the Wikipedia.[7] In its own words: "Wikipedia is a Web-based, multi-lingual, 'copyleft' encyclopedia designed to be read and changed by anyone. It is collaboratively edited and maintained by thousands of users via the wiki software […] and it is hosted and supported by the non-profit Wikimedia Foundation."[8]

The Wikipedia allows anybody who is registered to "boldly" edit whatever content she or he is interested in. A wiki generally means a web site comprised of the perpetual collective work of many authors. Similar to a web log in structure and logic, a wiki allows anyone, using a web browser, to edit, delete or modify content that has been placed on the web site including the work of other authors.

The Wikipedia offers style guides, codes of conduct and several guidelines on how to proceed when working on an entry. The overall work is being loosely coordinated via the "Community Portal" including information on possible contributions (e.g. a task list) and basically everything you need in order to participate.[9] Furthermore, the Wikipedia Foundation encourages everybody to translate entries into

their respective language, which leads e.g. to 180,000 German or 90,000 Japanese equivalents of some of the 450,000 English articles.

What does a Wikipedia article look like?
An article usually consists of:

a) the entry itself, possibly extended by sections such as "See also" "External links" and "Further reading",

b) four index-tabs above the entry, allowing the user to get to the discussion list, the editing view and the history log

c) the toolbox offering links to a list of other entries that link to the entry in question ("What links here"), to a list of recent changes made in the Wikipedia as a whole ("Related changes"), and to a list of special pages e.g. pages that offer help to the user.

Articles that concern ambiguous topics are listed on so called "disambiguation pages". If you are e.g. looking for "data" you are being told that a "datum is a statement accepted at face value (a "given")" and some things about the etymology and the usage of the term. The disambiguation page then also tells you, among other options, that Data is "a fictional android character in the Star Trek universe".

What are the pros and cons of the Wikipedia?
Of course, an open source, freely editable pool of information also raises doubts and criticism. Some of the cons should be mentioned here in order to reflect the experiences that have been made with the Wikipedia.

If you go through the pages criticising the Wikipedia - but which are part of this platform and thereby make the development progress more transparent - you will find several severe points of negative assessment. The entry "Wikipedia: Why Wikipedia is not so great"[10] talks for example about the lack of accuracy, completeness and NPOVness (NPOV = Neutral Point of View) of several articles. The author argues that the credibility of sources "can be dubious because of the anonymous nature of the wiki", that too many people merely add so called stub (short articles with little content marked to be expanded) and do not care about expanding existing articles, and that some participants refrain from maintaining a neutral point of view. The complete freedom of editing articles leads to bad style, nonsense content and arbitrary changes of articles in order to support individual points of view. This freedom can even end in what is called "edit wars". This term refers to conflicts between competing content providers, a rather odd thing for an information system that was actually meant to serve a beneficial goal. "Articles are sometimes copied virtually verbatim from other sources infringing on (international) copyright, particularly when no credit

is given." The Wikipedia thus obviously has a serious copyright problem.

Those are just some of the arguments speaking out against "vandals" and "geeks" that seem to have a strong impact on the quality of the Wikipedia. They uncover the immediate problem of completely open systems: they are democratic, but seemingly out of control.

One of the harshest critics of the Wikipedia concept was a former editor of the Internet version of the Encyclopaedia Britannica, Robert McHenry. In an online article he complains very much about the same details as mentioned above, but also admits that planning an encyclopaedia from the top is an almost unattainable aim:"I know, to begin with, that it can't be done in any thoroughgoing way. The job is just too big. Professional reviewers content themselves with some statistics -- so many articles, so many of those newly added, so many index entries, so many pictures, and so forth -- and a quick look at a short list of representative topics."[11]

On the pro side the openness of the system is what makes it very appealing to users. A forum founded on a data-base system that allows easy access which obviously results in a decent pool of information. More than 168,000 registered users show that there are obviously quite a number of people having an active interest in the wiki-kind of information. And the collaborative work has produced some remarkable results.

IASSIST and information sharing
I see the objective of reinforcing the international information exchange, as mentioned above, as being at the core of an organization that obviously has a very positive public image. The notion that there is an entire community of people out there that may have encountered the same problems that I have can be a relief. No bureaucratic framework is slowing down the work and the people active within IASSIST are highly motivated. Like many other voluntary organizations IASSIST wants to get more people to become involved and would like to see increased activities concerning training and education.

An IASSIST wiki could be an opportunity for people to contribute when they could and in the way they like. As the e-mail list shows, people are seeking contact and the large world of data archives and users suddenly becomes very small. This is a good starting point for further activities. A wiki encyclopaedia could serve as a framework for information on the topics that most members are interested in:

- data preservation,

- data editing,

- data access and dissemination,

- data confidentiality,

- data sources, and

- statistical and methodological issues.

This follows very much the framework of IASSIST's action groups and the organization of its website. The presentations from the IASSIST conferences could be utilized as quarries of information for the wiki knowledge base. Or better still it could be a repository for the information presented at the meetings i.e. prepare your slides and turn them into wiki content at the same time. Information bits that show up in e-mail conversations could be directly transferred into the data-base and scattered references could be bundled up in one place.

There should certainly be some rules in order to avoid the problems with the Wikipedia as described above:

- First of all, editing should be restricted to registered members of IASSIST. Since all of them are professional using the organization for day-to-day problems on the job, there should be no complications with "geeks" that try to wreck the system.

- Secondly, all entries should be in proper English.

- Thirdly, the recommendations of the Wikipedia on how to deal with your information should be taken over in order to make life easy for potential authors. The system should only be responsible for the rendering and formatting.

- And, last but not least, an open-minded notion toward changes and corrections of entries should be advocated, so that the harmony within IASSIST is not be disturbed by any "edit wars". But I honestly believe that this is no problem within IASSIST.

Conclusion
Both the Wikipedia and IASSIST live on the voluntary input from a dedicated community committed to sharing knowledge. This is a point they have in common, though the communities differ tremendously in size. The collaborative effort that ends up in well filled pools of information is probably the most appealing thing about the Wikipedia. It is obviously this aspect of the Wikipedia that draws a lot of interest toward it, which is an effect that could be utilized by IASSIST for its plans to reach out to the interested public.

As in many organizations few people do a lot of the work, but there is always potential for small contributions that could add up to a solid base. A wiki kind of information system could be a bank for this input. With regards to the scope, an IASSIST wiki should be between an Encyclopaedia Britannica style knowledge base and an open forum for volunteers to contribute 24/7. To avoid the problems of the Wikipedia a certain regulative framework should be established that facilitates the dissemination of

helpful and well considered information.

Of course, it will take funding and maintenance to run the system, but the idea of wiki could carry the IASSIST spirit one step further into the virtual world.

References
Members of the IASSIST Strategic Plan Action Group (ed.), IASSIST Strategic Plan, 2004-2009, 2004: <http://www.iassistdata.org/membership/plan_june2004.pdf>

The State of Digital Preservation: An International Perspective. Conference Proceedings, Washington DC 2002 (= CLIR Report No.83): <http://www.clir.org/pubs/reports/pub107/pub107.pdf>

Social Science Information Gateway: <http://www.sosig.ac.uk/>.

Vascoda: <http://www.vascoda.de/>.

Virtual Training Suite: <http://www.vts.rdn.ac.uk/>.

Wikipedia: <http://en.wikipedia.org/wiki/Wikipedia>.

Endnotes
[1] Oliver Watteler, GESIS – Zentralarchiv für Empirische Sozialforschung an der Universität zu Köln, Cologne (FRG). E-mail: watteler@za.uni-koeln.de

[2] Members of the IASSIST Strategic Plan Action Group (ed.), IASSIST Strategic Plan, 2004-2009, 2004 <http://www.iassistdata.org/membership/plan_june2004.pdf>

[3] The State of Digital Preservation: An International Perspective. Conference Proceedings, Washington DC 2002 CLIR Report No.83), p.91 <http://www.clir.org/pubs/reports/pub107/pub107.pdf>

[4] See <http://www.vts.rdn.ac.uk/>.

[5] See <http://www.sosig.ac.uk/>.

[6] See <http://www.vascoda.de/>.

[7] For an overview of the Wikipedia, see <http://en.wikipedia.org/wiki/Wikipedia>.

[8] Copyleft is a concept contrasting the idea behind copyrights. For details see article "copyleft" at Wikidepdia.

[9] On the technical side everything is open source and the development can be followed on the SourceForge platform. Wikipedia uses PHP and the MySQL DBMS.

[10] See <http://en.wikipedia.org/wiki/Wikipedia:Why_Wikipedia_is_not_so_great>.

[11] See <http://www.techcentralstation.com/111504A.html>