

Refining the Semantics of Open Access¹

Stefan Gradmann, Katja Mruck & Maximilian Stempfhuber

Key words: open access, publication process, interactive aspects of e-journals, publication platform

Abstract: In an attempt to stress a specific aspect of the "open-access" paradigm, experiences in setting up a collaborative publication platform in the social sciences are reported. The guiding principles in this context, the functional and technical requirements of the content producing partner, the open-access journal *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, are therefore presented first. Thereafter, a more general perspective of the project *German Academic Publishers* as an initiative mediating and federating expertise in the field is introduced before proceeding to the core section, dealing with aspects of technical standardization and transparency as a vital basis for open-access strategies and which is provided by the *Social Science Information Centre*.

Table of Contents

- [1. Introduction](#)
- [2. FQS—The Editor's Perspective on Open Access](#)
- [3. GAP—The Publisher's Perspective on Open Access](#)
- [4. DBClear— Generating the Interactive Environment](#)
 - [4.1 Modeling an e-journal with DBClear](#)
 - [4.2 Interactive aspects of e-journals](#)
 - [4.3 Overview of the publishing process](#)
- [5. Conclusion](#)

[References](#)

[Authors](#)

[Citation](#)

1. Introduction

The management of open-access journals is expected to be an interesting opportunity for current research information systems as these journals are available online and free of cost, making it possible to provide extensive information to the interested research communities in a fast manner. [1]

This paper is a joint contribution involving three distinct—and yet complementary—perspectives on the publishing process of open-access journals: namely the view of an information provider experienced in international co-operation, the online journal [Forum Qualitative Sozialforschung / Forum: Qualitative Social Research \(FQS\)](#), the view of [German Academic Publishers \(GAP\)](#), a German based initiative federating institutions and expertise in the area of open academic

1 This contribution was made possible by grants III N – BIB 46 BEfu01-01/54595, BIB44 HHuv 01-01, and III N – 554 922(1)00 from the [Deutsche Forschungsgemeinschaft](#) [German Research Foundation] (DFG), the central self-governing research organization in Germany. It was first presented at the 7th International Conference on Current Research Information Systems (CRIS 2004), Antwerp, Belgium, 12/13-15 May 2004.

electronic publishing, and of a research information provider, the [Social Science Information Centre \(IZ\)](#). IZ with GAP's assistance is currently involved in the integration of FQS into the national [Social Sciences Virtual Library \(ViBSoz\)](#). The paper reflects the model developed in the course of the preparatory work. [2]

Since the launch of the first open-access journal of the [Public Library of Science](#) and the release of the [Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities](#) both accompanied by enormous attention from mass media, open access is an issue far beyond the inner circles of the academic communities.² When discussing the term, the open-access paradigm is conceptualized rather differently, depending on the respective discourse fields and predominant interests in these fields. [3]

In this contribution, we mainly refer to the open-access paradigm (definition and strategies) as defined by the [Budapest Open Access Initiative \(BOAI\)](#). Following the BOAI, the

"literature that should be freely accessible online is that which scholars give to the world without expectation of payment. Primarily, this category encompasses their peer-reviewed journal articles, but it also includes any un-reviewed preprints that they might wish to put online for comment or to alert colleagues to important research findings." (<http://www.soros.org/openaccess/read.shtml>) [4]

Accordingly, two complementary strategies are suggested: "(OAJ) Researchers publish their research in an open-access journal if a suitable one exists, otherwise (OAA) researchers publish their research in a suitable toll-access journal and also self-archive it in their own research institution's open-access research archive." (<http://www.eprints.org/signup/sign.php>) [5]

In the following, we will deal with one of these strategies, the setting up of open-access journals, and in this specific context we basically aim at refining the (technical) semantics of the term "open access." We try to stress the necessity of a transparent, standard-based and technically open approach as a requirement for "access" to scientific content, being "open" in any conveyable meaning of the attribute. We are convinced that such a technical complement to the current politically focused debate around "open" or "free" access³ is badly needed in order to clearly identify the fundamental technical components that are required to make open access work at all! [6]

"Open" in this sense translates as "without restrictions"—and this not only means without financial restrictions, but also without any technological barriers, such as business models built around the use of proprietary formats and components or

2 See MRUCK, GRADMANN and MEY (2004) for an overview of the history, current state, problems and perspectives of the Open Access movement.

3 Cf. the repeated discussions between Stevan HARNAD and other protagonists in the [American Scientist Open Access Forum](#); for an example of "free access" see the "Washington DC Principles for Free Access to Science" from March 16, 2004, available at <http://www.dcpinciples.org/>.

the introduction of document identification and link resolving services that risk being controlled by a few commercial exploiters. Likewise "access" in this perspective needs a clearer technical conceptualization: Which technical operations exactly are implied when using this term? Will the simple distinction between "read-only" and "write" access rights be sufficient or do we need a richer taxonomy of potential operations (copying, annotating, versioning, quoting to name just a few examples)? And which rights do we give to which actor in which context? [7]

The questions above are just examples of one principle: open access is a valid and viable concept only as long as we are able to technically enable openness and as long as we are able to actually control and effectively model the technical means of access to information. Furthermore, open-access publishing strategies need to offer functional equivalents for all operations that were hitherto possible within the printed universe and which scientists are accustomed to. A good example is the need to offer an electronic counterpart of annotation facilities that can be used in a WWW based working environment transparently without infringing document integrity at the same time. We feel that we have some specific technical perspective to add in this context that is required to make scientists actually adopt such working environments as a valid substitute for their traditional techniques, which up to now have mostly been "electrified," using digital technology to basically re-implement analogous working models in a networked context. [8]

However, the prevailing part in this perspective—even though not quantitatively—should be the functional and technical requirements of the content producing partner, *FQS*, which therefore are presented first. Afterwards, the more general perspective of *GAP* is introduced before proceeding to sections dealing with aspects of technical standardization and transparency as a vital basis for open-access strategies. [9]

2. *FQS*—The Editor's Perspective on Open Access

*FQS*⁴ is a peer-reviewed, international and multilingual open-access journal, which has established a leading position in the field of qualitative social research. *FQS* publishes contributions in German, English and Spanish and using the competence and experience of more than 70 peer reviewers from 15 countries to ensure scientific quality, it has distributed 523 articles in 15 issues from January 2000 to May 2004. [10]

In the course of our work we realized that open-access is a crucial factor for improving the following five aspects of scientific publication considered important by our community: *transparency, impact, speed, quality and efficiency*. [11]

Whereas the aspects of transparency, speed and efficiency may be more or less evident, impact and quality require some specific comments. We have recently

4 *FQS* is based at the Freie Universität Berlin and funded by the DFG.

experienced a dramatic increase in *impact*⁵ that can be measured for example by the acknowledgment *FQS* receives by the international community⁶ or by having a closer look at popular search engines. When searching for terms such as "qualitative research," *FQS* is currently listed in one of the first places in popular search engines such as Google, Yahoo or Altavista. This is one clear indicator of *FQS* being a focus of international research using qualitative methods. This occurred as a result of several open-access strategies and namely as a consequence of systematically placing our publications in the [Open Directory](#), in various international websites (small ones as well as for example [SOSIG](#)), thus ensuring maximum web visibility⁷. [12]

As to *quality*, two dimensions evidently need to be taken care of: formal and content aspects. We have been successful in ensuring high standards in both areas⁸ at the price of an enormous amount of human labor both in our editorial work and in the time-intensive review process.⁹ While a certain amount of human and namely intellectual investment will always be required to maintain decent scientific standards, we are convinced that a lot of productivity could be set free once major parts of this quality assurance track are automated. This could be done using e.g. WWW-based workflow components, automated checks on formal consistency of documents and technological support for annotations during the publishing process and after a contribution has been published. It is in this area that we expect synergy from our co-operation with *IZ* and *GAP*. [13]

As a very important addition to our specific conditions of content production and publishing we would like to emphasize the need for low-barrier and standard technology in several senses of the term. Low-barrier access is required in terms of financial means for all instruments of content creation and submission, as many of our contributors live in countries outside the so called "developed world" and in economic conditions that do not support access to electronic production means that may be considered standard in Western Europe and Northern America. This applies to software as well as to data transmission (bandwidth): resource restrictions must be accounted for in both areas when making technological choices for our scientific community. [14]

-
- 5 HARNAD (2003) discusses some of the consequences of open access for maximizing research impact.
 - 6 See for example the consultative studies of Nigel FIELDING and Steve WOOLGAR for the Economic and Social Research Council; the latter used for the inclusion of German researchers as experts in his study more or less members of the *FQS* Editorial Board (FIELDING 2003, WOOLGAR 2003).
 - 7 To increase the visibility of texts published in *FQS* and also in traditional data bases, *FQS* signed a licensing agreement with EBSCO publishing in May 2004.
 - 8 *FQS* uses a double-blind peer review process. The need for peer review as a major tool for providing quality also in open-access journals is stressed continuously by Stevan HARNAD (see the contributions, available at <http://www.ecs.soton.ac.uk/~harnad/intpub.html> and MRUCK & MEY 2002 for a short summary on "Peer Review: Between Printed Past and Digital Future"). The importance of peer review within the open-access movement is obvious if one takes into account that BOAI explicitly refers to peer-reviewed contributions as core data, the open-access initiatives have to deal with; also the [Directory of Open Access Journals](#) only covers "free, full text, quality controlled scientific and scholarly journals."
 - 9 The efforts, necessary to establish such a resource is partly discussed in a report on the beginning and the first years of *FQS*, see MRUCK and MEY (2001).

Finally, an aspect of particular importance is the systematic support for multiple languages in parallel and for trans-lingual operations. It is simply impossible to impose the use of English or of any other single language upon our world wide community—and even if there was a technical choice of doing so, such a choice would annihilate a big part of our scientific potential: One of the very foundations of WWW community is the linguistic richness and the variety of cultural and epistemological diversity that comes with it. This aspect requires a major technical investment as a consequence, and a costly one too: language engineering components as typically required in such scenarios remain very expensive resources—still, they need to be part of the design of any technical solution for our journal, even though the initial investment may create major problems and thus the actual implementation of such components may not be possible right from the start. [15]

3. GAP—The Publisher's Perspective on Open Access

The overall goal of the DFG-funded project *GAP* is to stimulate and support open-access based scientific communication and to "return science to the scientists," as stipulated in the [Budapest Declaration](#). [16]

GAP does so by building an open co-operation framework for federating academic e-publishing institutions. This framework includes shared/distributed technical facilities, e.g. a shared WWW-based workflow implementation, supporting tools for open, standard-based object modeling and a generic, standard-based authentication layer that can be integrated in SSO architectures. On the community building and marketing side the network offers/develops common organizational/exploitation components, such as a common business model (one of the harder tasks and far from being resolved yet), a common framework of legal regulation for enabling a fair and transparent use of author's rights in an open digital context, and portal services for marketing the *GAP* partners' publications and making the related metadata accessible via the protocol for metadata harvesting proposed by the [Open Archives Initiative \(OAI\)](#). [17]

GAP is currently working on a business model designed to make this framework sustainable and scalable beyond Germany. We also plan to further investigate new models of article publishing (post-journals) and of quality assurance (public peer reviewing) in order to offer novel functional options to our community, which in turn can serve as valid arguments for requesting financial contribution in return for such novel services. [18]

GAP thus is an attempt to build a non-profit community service model around the open-access principle which would nevertheless have to be economically self-sustainable. Participants in *GAP* are a continuously growing number of major German Universities as well as independent scientific publishing initiatives, using *GAP* as a facilitating and federating instance as well as a technical back office. [19]

Even if a standard-based perspective aiming at the use of open-source solutions wherever appropriate has been at the very roots of the *GAP* project, the necessity

of applying this perspective very consequently has been stressed as a consequence of our collaboration with *FQS*. This co-operation started in 2003 and resulted in the proposal submitted to the DFG by *IZ*, *FQS* and *GAP* and which we referred to at the beginning of this paper. [20]

Standard-based technical solutions are omnipresent in this proposal, such as for instance the OAI-PMH-protocol or the use of metadata standards such as Dublin Core. These technical choices are currently widespread and probably more or less mainstream, in the sense of being adopted by all initiatives oriented towards open access, which are not essentially commercially driven. Therefore we perceive an essential need to build on open, affordable technical components that avoid any proprietary development, which communities would not be able to sustain in the end. Even though mainstream to some extent, these choices cannot be truly termed "innovative" anymore. [21]

There are, however, other aspects of our approach, which are based on emerging standards for open information and knowledge management, and which deserve the attribute "innovative." The remaining part of this paper will concentrate on three such aspects as the "pièce de résistance" of this contribution. [22]

4. *DBCclear*— Generating the Interactive Environment

For the different persons and groups involved in publishing an open-access journal like *FQS*—namely authors, editors, reviewers and readers—an environment has to be established in which predefined sequences of tasks can take place and are controlled and supervised. This includes the original submission of a journal article by the author, the reviewing process, and finally the publication on the Internet, where readers have access to the complete journal and can read and comment on the single articles. [23]

Within our project, from the submission of an article up to the point where the final version of the article is marked as "published," the *GAP* workflow system (see Section 3) handles all the intermediate steps and keeps track of the status of every publication (the journal) and entity (the articles and other parts of the journal). The ultimate publication step is carried out by the *DBCclear* system, which was originally developed for building clearinghouses for Internet resources, but can be used to maintain and publish different types of hierarchically structured data. [24]

4.1 Modeling an e-journal with *DBCclear*

[DBCclear](#) is a flexible system for cataloging Internet resources and describing them with metadata (HELLWEG et al. 2002). It can handle different metadata schemas in parallel and combine them for searching and browsing, as long as some common attribute is used in every single schema. In the case of clearinghouses, this is often a classification, a hierarchical view on the relevant aspects of a domain, which can be used to classify objects like publications, institu-

tions, data or methods. For journals and books, the table of contents builds up a similar structure, where different parts of a publication are arranged hierarchically and can be browsed by the reader. Each part of the publication is described by a special set of metadata attributes, which are provided by the *GAP* workflow system through an OAI-MPH interface (OAI 2003). The metadata describes the state of every part of a publication, e. g. that it has been submitted to the system, is being peer reviewed or ready to be published on the Internet. [25]

To accommodate for many different structures—the actual sequence of volumes and issues down to the single editorials, articles and other parts of publications—*DBCclear* is able to automatically and dynamically generate the structure of an electronic publication by taking a sequence of attributes (e. g. volume, issue, page number and article title) and generating table of contents by using the actual values of these attributes. With every change in the metadata provided by the workflow system (e. g. publishing or withdrawing an article), *DBCclear* recalculates the browsing structure and presents the reader with an updated view on the publication. [26]

4.2 Interactive aspects of e-journals

The full benefits of electronic publications can only be realized if they follow the publishing paradigm of the Internet—the presentation of content in the HTML format—and combine this approach with online communication facilities. Currently, discussion boards or mailing lists are used to allow readers to exchange their thoughts and to support (informal) communication between them. Inherent in this interaction model is the separation between the original content—an article published online in a defined version by the editor in charge—and the new content being produced by the readers of the article over time. The implications are twofold. On the one hand, the persistence of the original work is preserved, so that it can be referred to by others. On the other hand, new information entered by the readers might correct, modify or extend the original content leading to new versions of the publication, but since it uses a separate media—the discussion board instead an authoring system—this new version is never actually created. [27]

One of our goals is to connect the original publication more tightly with the annotations made by the readers so that smaller entities within a publication (e. g. paragraphs) can be subjected to comments etc. Taking the current structure of the *FQS* into account, paragraphs could be the proper level of detail as they are consistently numbered across the different language and file versions of the articles (currently PDF and HTML format). How to handle annotations for a specific paragraph in multiple languages, e. g. whether to mix them or to restrict annotations to a single language, has to be discussed with authors and readers. [28]

Essential for such an interactive publishing model are technical standards which ensure consistency and integrity. Regarding the document structure, open formats like the [OpenOffice](#) XML format (EISENBERG 2004) or the [DocBook](#) standard (WALSH & MUELLNER 2003) have to be evaluated. Since both are

based on XML, transformations to different output formats, like HTML or PDF, can be realized easily. To link between different documents, Uniform Resource Locators (URL), Uniform Resource Names (URN), [Digital Object Identifier System \(DOI\)](#), or the [XML Linking Language \(XLink\)](#) and its related standards have to be watched closely. And last but not least, authentication and authorization are important as they define the identity and roles of actors in the context of the publishing process (again covering everyone from author to reader) and determine the operations (e. g. reading, writing, deleting) they are allowed to execute on specific entities (e. g. the whole document or only parts of it, like paragraphs, annotation or metadata). Open standards like [Shibboleth](#) or Lightweight Directory Access Protocol (*LDAP* [Broken link, FQS, December 2004]) will be considered for authentication and authorization as they allow us to link to already established systems. [29]

4.3 Overview of the publishing process

The publishing process consists of discrete, sometimes iterative steps from article submission to publication on the Internet. The overall workflow is divided into two parts, the editorial process—handled by the *GAP* workflow module—and the presentation process, which is handled by *DBCclear*. Figure 1 gives an overview of the steps involved in each process.

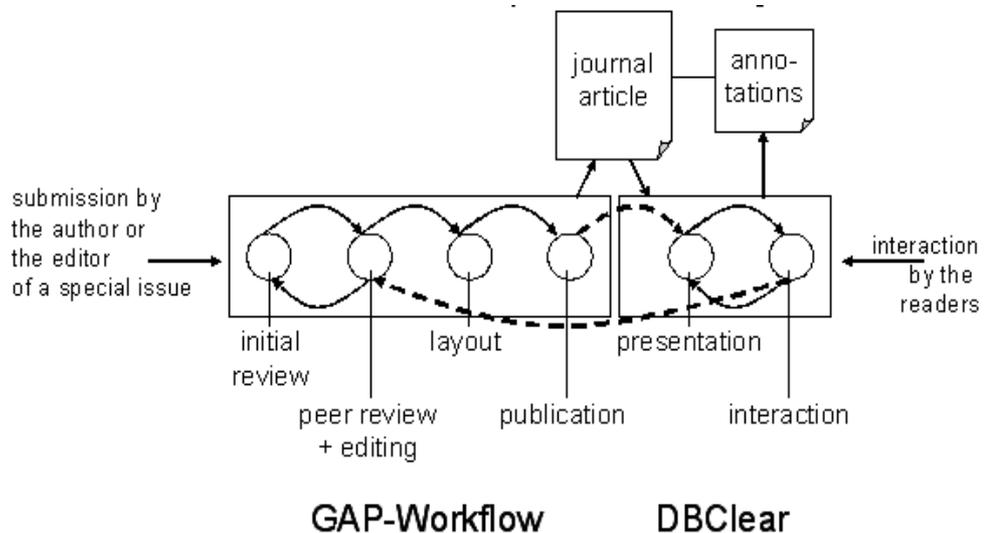


Figure 1: The overall workflow for an interactive e-journal [30]

The technical linkage between the processes will be realized by exchanging messages over the OAI-MPH protocol. *DBCclear* polls the work flow's OAI interface in regular intervals to learn about newly published articles and to insert them into the structure of the journal. Depending on the role of the reader and the actions he or she executes (e.g. reading, annotating, discussing), new events will be generated and inserted into the workflow if necessary. In the case of annotations, an additional step of reviewing and moderating annotations could be initiated, so that authors could response before they are generally Since the

concept of having annotations made by readers as integral parts of a publication is not new (<http://www.wiki.org/>), but lacks proven and accepted models in the context of scientific publications, feedback from all actors will be needed to develop such models and in trying to establish them in different scientific communities. [31]

5. Conclusion

We hope to have shown how open-access scientific journals can create challenges as well as offer attractive opportunities. We had a closer look at some technical and standardization aspects of open-access publishing in order to present a technical way of managing the communication and publication chain. This in turn creates a potential for transforming these hitherto rather static information resources into dynamic components of scholarly communication and publishing environments to be modeled in future research information systems. [32]

References

- Fielding, Nigel (2003). *Qualitative Research and E-Social Science: Appraising the Potential*. ESRC consultative study. Available at: <http://www.esrc.ac.uk/esrccontent/DownloadDocs/Fieldingreport.doc>.
- Harnad, Stevan (2003). Open Access to Peer-Reviewed Research Through Author/Institution Self-Archiving: Maximizing Research Impact by Maximizing Online Access. In Derek Law & Judith Andrews (Eds.), *Digital Libraries: Policy Planning and Practice*. Ashgate Publishing. Available at: <http://www.ecs.soton.ac.uk/~harnad/Temp/digital-libraries.htm>.
- Eisenberg, J. David (2004). *OpenOffice.org XML Essentials—Using OpenOffice.org's XML Data Format*. Preliminary version. Available at: <http://books.evc-cit.info/>.
- Hellweg, Heiko; Hermes, Bernd; Stempfhuber, Maximilian; Enderle, Wilfried & Fischer, Thomas (2002). DBClear: A Generic System for Clearinghouses. In Wolfgang Adamczak & Annemarie Nase (Eds.), *Gaining Insight from Research Information: Proceedings of the 6th International Conference on Current Research Information Systems*, University of Kassel, August 29-31 (pp. 131-139). Available at: http://www.uni-kassel.de/CRIS2002/programme_abstract_ghk?id=78.
- Mruck, Katja & Mey, Günter (2001). FQS—From the Shop Floor (43 paragraphs). *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [On-line Journal], 2(2), Art. 8. Available at: <http://www.qualitative-research.net/fqs-texte/2-01/2-01hrsg-e.htm>.
- Mruck, Katja & Mey, Günter (2002). Peer Review: Between Printed Past and Digital Future. *Research in Science Education*, 32(2), 257-268.
- Mruck, Katja; Gradmann, Stefan & Mey, Günter (2004). Open Access: (Social) Sciences as Public Good [32 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [On-line Journal], 5(2), Art. 14. Available at: <http://www.qualitative-research.net/fqs-texte/2-04/2-04mrucketal-e.htm>.
- OAI (2003). *The Open Archives Initiative Protocol for Metadata Harvesting*. Open Archive Initiative, Protocol Version 2.0. Available at: <http://www.openarchives.org/OAI/openarchivesprotocol.html>.
- Walsh, Norman & Muellner, Leonard (2003). *DocBook: The Definitive Guide*. Updated: Wed, 31 Dec 2003. Available at: <http://www.docbook.org/tdg/en/html/docbook.html>.
- Woolgar, Steve (2003). *Social Shaping Perspectives on e-Science and e-Social Science: the case for research support*. ESRC consultative study. Available at: <http://www.sbs.ox.ac.uk/downloads/E-SocialScience.pdf>.

Authors

Stefan GRADMANN, Head of the Hamburg
University Information Management Unit

Contact:

Dr. Stefan Gradmann

Hamburg University Computing Center
Schlüterstr. 70
20146 Hamburg, Germany

E-mail: stefan.gradmann@rrz.uni-hamburg.de
URL: <http://www.rrz.uni-hamburg.de/RRZ/S.Gradmann/index.html>,
<http://www.gap-c.de/>

[Katja MRUCK](#), FQS Main Editor

Contact:

Dr. Katja Mruck

FQS / qualitative-research.net
Freie Universität Berlin
FB 12, WE 09
Habelschwerdter Allee 45
14195 Berlin, Germany

E-mail: mruck@zedat.fu-berlin.de
URL: <http://www.qualitative-research.net/fqs/impressum/mruck-e.htm>

Maximilian STEMPFHUBER, Head of Research &
Development Department, Social Sciences
Information Centre (IZ)

Contact:

Dr. Maximilian Stempfhuber

Social Science Information Centre
Lennéstr. 30
53113 Bonn, Germany

E-mail: Stempfhuber@iz-soz.de
URL: <http://www.gesis.org/IZ/>

Citation

Gradmann, Stefan; Mruck, Katja & Stempfhuber, Maximilian (2004). Refining the Semantics of Open Access [32 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 5(3), Art. 9, <http://nbn-resolving.de/urn:nbn:de:0114-fqs040395>.

Revised 6/2008