The long and winding road: institutional and disciplinary repository at Saarland University and State Library

Summary

Purpose of this paper

The article describes the Open Access archiving of scientific primary and secondary literature on the institutional and the disciplinary repository of the Saarland University and State Library. The challenges, opportunities and success factors of both repositories are compared and evaluated.

Approach

The article is based on the findings and experiences of eight years of Open Access archiving practice.

Findings

Basic requirements for success and acceptance are

- quality control and standardisation; that could be proved for example by an official certificate,

- visibility and dissemination; researchers emphasize that their documents are indexed by the databases most common within their community. This applies especially to disciplinary repositories.

Value of paper

Universities should undoubtedly offer an institutional repository. Usually the university library runs this repository. The article lists some key points that should be considered when an organisation plans to start respectively run a institutional or disciplinary repository.

1.0 Introduction

The Library of Saarland University and the State of Saarland (Saarländische Universitäts- und Landesbibliothek, SULB)[1] is the central academic library of Saarland University,[2] Germany. At the same time, SULB is participating in a national program for scientific literature supply organized by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation).[3] As a result of this dual role SULB maintains and runs two Open Access repositories:

- SciDok, http://scidok.sulb.uni-saarland.de, an institutional data repository for all the scientists at the University of Saarland regardless of their discipline and
- PsyDok, http://psydok.sulb.uni-saarland.de, a disciplinary data repository for psychologists from the German-speaking countries.

By comparing both repositories this article aims to provide an insight into both the challenges and the chances of institutional and disciplinary repositories in Germany.

2.0 At the beginning of the road

Pre-print servers can be considered as the beginning of Open Access publishing. Several pre-print servers established especially for the sciences have come to enjoy an excellent reputation in the academic community. As a consequence, scientists increasingly asked for the implementation of electronic publishing platforms. For the most part, university libraries offered these services because they were able to guarantee the long-term availability of this new publishing distribution outlet. The prevailing software for digital repositories in Germany is OPUS because it is tailored to the special needs and requirements of the German e-publishing community.

OPUS[4] was developed at the University of Stuttgart (Germany). Meanwhile, several universities are collaborating in the further development of OPUS; SULB has also contributed several modules. To date, more than 50 German universities are using OPUS to run their digital repositories. The systems requirements include UNIX/Linux with a MySQL- or MSQL-Database and an APACHE/PHP-environment. OPUS is also mentioned in the Open Society Institute's guide to institutional repository software (Open Society Institute, 2004). The preferred file formats to be published are PDF-, PostScript- or HTML-files, the ideal format to date is PDF/A.

The repositories offer a variety of retrieval options such as

---

[1] For more information about Saarland University and State Library, see http://www.sulb.uni-saarland.de
[2] For more information about Saarland University, see http://www.uni-saarland.de
[3] In Germany scientific information supply is characterized by a distributed, supra-regional system of Special Subject Collections (in German: Sondersammelgebiet, SSG) which are attended by scientific libraries. There are 121 special subject collections which can be found in more than 40 libraries all over Germany. The SSG´s main function is to build a collection of subject related scientific literature which is as complete as possible; this includes domestic and foreign research literature (including electronic media, microfilms and multimedia objects) and scientific journals. Within this system the SULB is responsible for the Special Subject Collection Psychology since 1966. See http://www.dfg.de/en/research_funding/scientific_instrumentation_and_infrastructure/scientific_library_services_and_information_systems/project_funding/funding_objectives/ for more information.
[4] For more information about OPUS, see http://opus.uni-stuttgart.de

- browsing through document types, faculties and institutions as well as subject classifications (e.g. DDC, Computing Classification System (CCS), Mathematics Subject Classification (MSC), Physics and Astronomy Classification Scheme (PACS) on SciDok and APA Classification and Indexing System on PsyDok)
- meta-data search
- fulltext search (PsyDok even offers a combined meta-data/ fulltext search via the SCIRUS search engine)

In addition to this, users can subscribe to an alerting service which is realised as an RSS feed.

2.1 Open Access and long term accessibility and data preservation

Both SciDok and PsyDok are Open Access servers, which means publishing and document use are free of charge. Many different document types may be published and found in these repositories: diploma and master theses, dissertations, professorial dissertations, preprints, teaching materials, psychological scales, grey literature, monographs, book chapters, journals and of course articles from journals.

The SULB Library ensures permanent and long-term access, long-term usability and **citeability** for all documents published. In order to provide a permanent and stable electronic location for its documents, both repositories offer a consistent identifier in the form of a Uniform Resource Name[5] (URN) for each document. Each URN is registered by the German National Library[6] (Die Deutsche Nationalibliothek, DNB). As a State Library with the duty of archiving electronic documents, SULB uses the results and findings of relevant German projects such as NESTOR[7] or KOPAL[8] in order to guarantee the long-term access to and availability of its repositories. Technologies and practices developed and evaluated within these projects will be applied to PsyDok and SciDok. Further features to be implemented as soon as possible include multilingual user interfaces, cross-lingual search and interoperable download statistics for each document.

2.2 Quality and standardisation of the repositories

When speaking of standardisation and technological and organizational quality control of digital publication services in Germany, the German Initiative for Network Information (German: Deutsche Initiative für Netzwerkinformation e.V., DINI)[9] has to be mentioned. DINI developed a

---

[5] For more information about the URN, see http://www.persistent-identifier.de/?lang=en
[6] For more information about DNB, see http://www.d-nb.de
[7] For more information about nestor, see http://www.langzeitarchivierung.de/index.php?newlang=eng
[8] For more information about kopal, see http://kopal.langzeitarchivierung.de/
[9] For more information about DINI, see http://www.dini.de/

variety of standards digital repositories have to comply with in order to become certified DINI-Servers. DINI strives to lend greater credence and reliability to electronic publishing with digital repositories; the DINI Certificate is a mark of quality. The criteria checked by DINI-experts include organizational matters (e.g. workflow optimization), a dedicated policy for the repository, questions of data integrity and authenticity, guaranteed archival storage periods, detailed description of the service, data export into long-term archival systems, questions of content indexing and question of meta-data management and export. PsyDok and SciDok meet all these criteria and are certified DINI-Servers. Another important proof of the adherence to standards is of course the compliance of these two repositories with the Protocol for Meta-data harvesting of the Open Archives Initiative.[10]

2.3.1 SciDok: An institutional repository for scientists from Saarland University

The SciDok digital repository was launched in 2001. Initially, SciDok - just like most German repositories - was used for the electronic publication of dissertations. At the beginning, the task of winning support for SciDok was fraught with obstacles: visiting scientists on-site to introduce SciDok to them, staging campaigns for electronic publication, obtaining modifications of internal university regulations governing the publications of the university's members, giving Ph. D. students the possibility to publish their dissertations in electronic form etc. At the same time, an awareness of the advantages inherent in depositing other publications (e.g. research papers) in the repositories had to be developed. This is the reason why so many dissertations can be found on German Open Access Repositories: in German repositories, grey and white literature is usually mixed.

The scientists responded to SciDok in two different and opposing ways: natural scientists, computer scientists and physicians on the one hand received the new offering positively, while researchers from the Humanities on the other hand were hampered by a lack of confidence in the "immaterial" character of online technology.

2.3.2 PsyDok: A disciplinary repository for the psychological community

In the fall of 2002, SULB and the Institute for Psychological Information (Zentrum für Psychologische Information und Dokumentation ZPID, Germany)[11] launched a joint project named „Digital Psychology Information (DPI)" which was funded by the DFG. The aim of DPI was to improve the availability and accessibility of electronic psychological information. One of the work packages defined within DPI was the implementation of the subject-specific repository PsyDok operated by SULB. By establishing a service like PsyDok, SULB responded to the ephemeral

---

[10] http://www.openarchives.org
[11] ZPID creates electronic databases on literature, tests, audiovisual media, and Web resources. For more information,

nature of most online information and at the same time made a contribution to the availability and preservation of electronic documents. In accordance with the principles laid down for special subject collections (SSG-), PsyDok acts as a supra-regional, disciplinary repository for digital scientific documents. It is a central point for psychological fulltext publishing under Open Access policies, not only for Germany but also for other German-speaking countries like Switzerland and Austria.

From the inception of the project on, special emphasis was put on grey literature. In contrast to SciDok, where different faculties and their attitudes towards online documents have to be considered and taken into account, PsyDok has only one single clientele: the psychological community. Hence, PsyDok was advertised at psychologists' conferences and meetings and a general acceptance was easily achieved. This is the reason why PsyDok was and is more frequently used than SciDok.

3.0 The road splits and turns to green

Scientists cannot and do not want to abandon the publication of their articles in scientific journals. The main reason for this is the peer review process guaranteed by the publishers. This quality control and the attendant gain in reputation linked to this process provides an incentive for academic authors to publish preferably in such journals. Repositories – especially the multidisciplinary repositories of universities – cannot offer a comparable mechanism for quality control. As a result, the desire to collect journal articles in open access repositories in order to make them freely available arose soon after. Initially, academic publishers tried to prevent this by referring to copyright issues. However, sustained pressure on the publishers by universities, libraries, funding organisations and, last but not least, the authors themselves is gradually effecting a change, so that it is now getting easier over time to deposit digital copies of journal articles in the repositories. SHERPA creates both public awareness and puts political pressure on the publishers by using an system of colour codes: it assigns publishers a certain colour according to their copyright and self-archiving policies. Publishers allowing parallel depositing of postprints or preprints of journal articles are rated as 'green' publishers. Hence the metaphor "green road". Meanwhile, all the important international publishers like Elsevier or Springer have become green publishers. Unfortunately, most of the smaller German publishers still have difficulties with adopting a green copyright policy.

The SULB library is aiming to get as many journal articles into SciDok as possible. The starting point is the annual bibliography ("Jahresbibliographie") of Saarland University. Here, all publications of all university members have to be collected. Initially, contributions were voluntary

---

see http://www.zpid.de

and the cooperation from the scientists was poor. But since the university administration is now using the figures of this bibliography as the basis for the evaluation of the faculties and departments, all researchers at Saarland University are taking pains to contribute all their publications.

A unique feature of SciDok is its dovetail with the annual research bibliography of the Saarland University:[12] Scientists can simultaneously contribute publications for the annual research bibliography and publish the corresponding documents in SciDok in one single step. If scientists register only the metadata in the annual research bibliography, SULB staff members check the publisher´s copyright policies for self-archiving in the SHERPA database.[13] If a parallel publication is possible, a SULB staff member will ask the author for the full text and publish it in the repository using the metadata already stored in the bibliography. Metadata in the annual research bibliography and full texts in SciDok are interlinked.

3.1 Dissemination and Visibility

The primary beneficiaries of these repositories are the scientists themselves. Depositing documents in Open Access repositories offers impressive advantages especially when compared with the still customary but less-than-ideal practice of publishing documents on web servers of academic institutions or faculties. Such information cannot be searched (or retrieved) systematically; if a document is found at all, it is found accidentally by click-by-click-retrieval. In addition to this, publications like these are of a very limited value in the scientific discourse: both the life-span and availability of such documents are short (or at least unpredictable) and permanent access to these data, their immutability and quotability, are doubtful.

By offering a rapid and global flow of scientific information, PsyDok and SciDok exemplify the motto "maximizing research impact by maximizing research access". Numerous studies have proved that the additional publication of documents in an Open Access repository increases the number of its citations (Lawrence 2001, Brody et al. 2005, Brody et al. 2004, Hajjem et al. 2005, Eysenbach 2006). The number of citations also depends on the visibility of the documents; that is another incentive for authors to publish their documents in these repositories.

Unfortunately, most publishers expressly allow the depositing of a parallel version of the articles only in *university* repositories. Disciplinary repositories like PsyDok are seldom mentioned as another publication outlet in the SHERPA list. Thus, another grey area has grown as authors are uncertain whether to contribute articles to PsyDok in addition to the conventional academic

---

[12]   http://jahrbib.sulb.uni-saarland.de
[13]   http://www.sherpa.ac.uk/about.html

publishing outlets or not. This divergence between institutional and disciplinary repositories is likely to disappear, however, because both authors and publishers stand to benefit from a parallel publication: documents deposited in Open Access repositories achieve greater impact and gain greater exposure because of the dissemination via the Open Access and scientific search and retrieval infrastructure, e.g. the Open Archive Initiatives Protocoll for Metadata Harvesting (OAI PMH).[14]

Hence, the SULB library makes every effort to integrate both repositories in disciplinary and multidisciplinary retrieval systems.

3.2.1 SciDok

The documents on SciDok are indexed by numerous search engines, databases and catalogs:

- SULB-OPAC,[15]
- Karlsruher Virtueller Katalog (Karsruhe virtual catalog, a meta search engine for library catalogs),[16]
- OPAC of the Südwestdeutscher Bibliotheksverbund (SWB),[17]
- Bielefeld Academic Search Engine (BASE),[18]
- Open access to scientific literature (OASE),[19]
- OAI based search engines,[20]
- Google Scholar,[21]
- Scirus[22]/SCOPUS[23] of Elsevier,
- Thomsons Web Citation Index,[24]
- partly indexed in subject databases (e.g. CiteSeer[25] or IOPort[26])

3.2.2 PsyDok

PsyDok offers all psychologists – regardless of their university or institution – the opportunity to make their scientific output freely available and accessible. Disciplinary repositiories are very attractive for database producers and search engine operators because the entire information in such

---

[14]   http://www.openarchives.org
[15]   http://opac.sulb.uni-saarland.de
[16]   http://www.ubka.uni-karlsruhe.de/kvk.html
[17]   http://swb.bsz-bw.de
[18]   http://base.ub.uni-bielefeld.de/
[19]   http://www.ubka.uni-karlsruhe.de/kvvk.html
[20]   e.g. OAISTER, http://oaister.umdl.umich.edu/o/oaister/
[21]   http://scholar.google.com
[22]   http://www.scirus.com
[23]   http://www.scopus.com
[24]   http://scientific.thomson.com/press/2005/8298416/
[25]   http://citeseer.ist.psu.edu/cs

a repository can be used. For the same reason, disciplinary repositories are also appealing for scientists: a highly-regarded repository very likely paves the way into a particular community´s well-established databases.

In addition to the information sources mentioned above, PsyDok is integrated in the following services:

- Diploma and matser theses database[27] of the ZPID,
- PsychSpider[28] of the ZPID,
- the database Psyndex[29] of the ZPID (the most comprehensive psychological database in the German speaking countries, corresponding to PsycINFO in the Anglo-American region),
- PsychLinker of ZPID and SULB,[30]
- Current Web Contents,[31]
- PsychCRAWLER[32] of the American Psychological Association (APA).

And nota bene, SCIRUS ranks PsyDok among its preferred web sources (like ArXiv or BioMed Central et al.).[33]

3.3 Issues of quality and standardisation

Usually, operators of 'green road' repositories do not have any influence on the quality of the submitted and archived documents. Hence, SULB has little chance to ensure any kind of quality control for SciDok or PsyDok. Nevertheless, quality principles can be upheld implicitly because they are bound up with the type of document. With institutional series or working paper series, an editorial review should be assumed. With dissertations and professorial dissertations a sufficient quality can be assumed because of the type of document. Diploma and master theses will only be published if the referee explicitly states a recommendation. In the case of the growing body of preprints, the authors' reputation can be based on former full publications; and the quality of a pre-print document can be judged through a comparison with the later formal publication. Postprints are controlled by the publishing houses' expert staff, mostly in the form of a peer review.
Each department or faculty can implement a peer review for their own documents.

4.0 Conclusion

---

[26] http://www.io-port.net/
[27] http://www.psyndex.de/index.php?wahl=products&uwahl=frei&uuwahl=thesesinfo
[28] http://www.psychspider.de
[29] http://www.psyndex.de/
[30] http://www.psychlinker.de
[31] http://www.scientific.thomson.com/free/essays/selectionofmaterial/cwc-criteria/
[32] http://www.psychcrawler.com

The situation at Saarland University reflects the status of Open Access at most German universities. Unfortunately, the university has yet not adopted a voluntary Open Access policy, not to mention a mandatory one. At the same time, the idea of Open Access does not have a lot of currency among many scientists. This is consistent with the findings of a study conducted by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG, 2005), which proves that Open Access is hardly known among German scientists. However, those scientists who became familiar with Open Access services are using them to an increasing extent.

The development of PsyDok clearly demonstrates that. After getting familiar with PsyDok, many scientists have come to appreciate the benefits of Open Access publishing, which then leads to an increasing number of publications. As a result of good experiences[34] with this publishing outlet, a number of authors have asked SULB for extension of its publication services. By now, several primary publications such as conference proceedings or monographs can be found in the two repositories.

Apart from the electronic Open Access document, authors and learned societies are asking for printed versions of those documents:

- In 2006, a theological Open Access journal[35] was founded at Saarland University and is published electronically at SULB. The editors plan to collect the articles of one full year in a printed volume.
- Several authors and institutions using PsyDok are asking to have monographs published.
- Psychologists from the University of Klagenfurt (Austria) already publish their proceedings *Beiträge zur Qualitativen Inhaltsanalyse* on PsyDok and are planning to edit a printed version additionally.

These encouraging experiences confirm SULB in its efforts to improve and expand the green road services and also to develop alternatives for golden road publishing. This means that SULB would have to take on the function of a university press.

Literature

Brody, Tim and Harnad, Stevan (2004). Comparing the Impact of Open Access (OA) vs. Non-OA

---

[33] http://www.scirus.com/srsapp/aboutus/#sources
[34] "PsyDok can be considered as an alternative to conservative publication modalities: high visibility, long-term citability and free accessibility are the platform´s advantages." (Mayring, Brunner, 2006)
[35] http://aps.sulb.uni-saarland.de/theologie.geschichte/

Articles in the Same Journals. Dlib Magazine, 10, Nr. 6,
http://www.dlib.org/dlib/june04/harnad/06harnad.html

Brody, Tim, Harnad, Stevan and Carr, Les (2005). Earlier Web Usage Statistics as Predictors of
Later Citation Impact. Journal of the American Association for Information Science and
Technology (JASIST). http://eprints.ecs.soton.ac.uk/10713/

Eysenbach, Gunther (2006). Citation Advantage of Open Access Articles. Public Library of Science
Biology 5(4), http://biology.plosjournals.org/perlserv/?request=get-
document&doi=10.1371/journal.pbio.0040157

Hajjem, Chawki, Harnad, Stevan and Gingras, Yves (2005). Ten-Year Cross-Disciplinary
Comparison of the Growth of Open Access and How it Increases Research Citation Impact. IEEE
Data Engineering Bulletin 28(4), S. 39-47
http://eprints.ecs.soton.ac.uk/11688/

Lawrence, Steve (2001). Free online availability substantially increases a paper's impact.
Nature, 411, S. 521-522 http://www.nature.com/nature/journal/v411/n6837/full/411521a0.html

Open Society Institute (2004). A guide to institutional repository software, 3rd edition.
http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf

Deutsche Forschungs-Gemeinschaft (DFG) (2005). Publikationsstrategien im Wandel? : Ergebnisse
einer Umfrage zum Publikations- und Rezeptionsverhalten DFG-geförderter Wissenschaftler unter
besonderer Berücksichtigung von Open Access / Ed. Deutsche Forschungsgemeinschaft (DFG),
Weinheim, Bergstr : Wiley-VCH, 2005

Bios

Ulrich Herb studied sociology at Saarland University. He is member of the electronic publishing
group of the Saarland University and State Library.

Matthias Müller studied biology at Saarland University. He is deputy director and member of the
electronic publishing group of the Saarland University and State Library.