Implementing Open Access Policy: First case studies

Chris Armbruster

Max Planck Digital Library, Max Planck Society, Berlin 14195, Germany

Abstract   When implementing open access, policy pioneers and flagship institutions alike have faced considerable challenges in meeting their own aims and achieving a recognized success. Legitimate authority, sufficient resources and the right timing are crucial, but the professionals charged with implementing policy still need several years to accomplish significant progress. This study defines a methodological standard for evaluating the first generation of open access policies. Evaluating implementation establishes evidence, enables reflection, and may foster the emergence of a second generation of open access policies. While the study is based on a small number of cases, these case studies cover most of the pioneer institutions, present the most significant issues and offer an international overview. Each case is reconstructed individually on the basis of public documents and background information, and supported by interviews with professionals responsible for open access implementation. This article presents the highlights from each case study. The results are utilized to indicate how a second generation of policies might define open access as a key component of digital research infrastructures that provide inputs and outputs for research, teaching and learning in real time.

The registry of open access policies (ROARMAP)[1] counts 94 institutions and 46 funders that have adopted an open access policy. This article investigates many of the most significant cases of open access policy implementation. The results show an open access rate to publications approaching 40%, while more than 50% open access is exceptional.

Till August 2010, the open access bibliography listed pieces of advocacy, commentary, reflection and descriptive analysis on open access policies, but no substantive study of policy generation or policy implementation[2]. Advocates of open access have identified suitable policy instruments, with a preference for open access mandates, which typically make provisions for the green (i.e. author's final manuscript) and gold (i.e. published version of record) route. Where policy implementation has been attempted in earnest, it seems to have taken three to four years to achieve results.

1   First-generation Open Access: Some characteristics of policy development

Open access is an unusually broad area of policy development. It spawns public policy as well as research policy, and has become an issue for a variety of academic institutions, such as research funders, national academies and universities. Policy development usually progresses in several stages, and most stages of open access policy are very well documented[3] from issue identification through an initial policy formulation and stakeholder consultation to the adoption of a specific policy solution. For this purpose it suffices to note the following:

- Open access seems driven by the idea that the rise of the Internet opens up unprecedented possibilities to make the published research output universally available and engender wider and new forms of use. The perception of a serials crisis, with steeply rising prices for an ever-increasing fleet of journals may explain why libraries invest in open access. But the serials crisis was central neither to the Budapest Open Access Initiative nor to the Bethesda Statement on Open Access Publishing and the Berlin Declaration on Open Access[3–5], all of them focus on the free use of publications and wide-ranging permissions for users, backed up by a stable digital infrastructure of open access repositories.
- With open access, new stakeholders have arisen in scholarly communication, such as digital repositories, research funders and research institutions, which have an interest in shaping the digital infrastructures of scholarly communication. The stakeholders are seeking allies among publishers and libraries, while placing on them new demands. Publishers and libraries may promote their own agenda and convert to open access, but their influence on policy development is only indirect.
- The key players shaping the policy agenda are research institutions that have passed open access mandates (which may be based on public legislation as in

the case of the National Institute of Health, or come from a public body with a wide range of functions, such as the European Commission). There is variation among the research institutions (e.g. funders and universities), in their constitution (e.g. public and private) and how the open access mandate came to pass (e.g. faculty approval, executive decision, public legislation), but it is the institution that drives forward the implementation.

- Research institutions command respect in the community. Moreover, as funders and employers, research institutions have additional leverage over researchers. Therefore, open access policies target the researcher as author, requiring the deposit of the author's final peer-reviewed manuscript and encouraging open access publishing. Since researchers are required to play along, any implementation must be sensitive with respect to the moral and legal rights traditionally associated with scholarship and authorship.

These observations justify a focus on institutional case studies, and on the involvement of authors, be it as grantees, members or employees. All policies, whether based on a mandate or not, target researchers (faculty, grantees, employees) as authors.

2   Studying cases: Sampling, interviews and validation

Theoretical sampling means selecting cases not for being representative (in any way), but as to whether they are likely to increase our understanding of the implementation process. This requires a consecutive and cumulative case study approach, in which the following case study is always used to modify, control and enlarge previous ideas and results[7,8]. The following sample would seem saturated in terms of variety, and the methods utilised define a standard for further study. These are:

- Expert interviews.   On the assumption that general incentives and parameters may be essential to policy implementation, success is highly dependent on institutional factors. Therefore, it is important to interview the professionals charged with implementation. The expert interviews, based on an open set of questions pertinent to each case, are conducted to elicit the operational know how that is so important for success[9–10].
- Qualitative content analysis.   The interviews with the experts are subjected to a content analysis by means of 1) Summation and reduction to the essential elements, 2) delineation of the major categories of meaning and interpretation; 3) explication of context and any unclear issues; and 4) structuration of text[11].
- Communicative validation.   The draft analysis is returned to the interview partners for discussion, comment and approval (internal validation). Moreover,

the analysis is discussed with outside experts in the field (external validation)[12].

2.1   Refining green open access policy: Queensland University of Technology

The Queensland University of Technology (QUT) in Brisbane, Australia, is a notable pioneer of open access,[13] being the first institution worldwide to adopt a deposit mandate in September 2003. QUT is an exemplary case of institutional implementation because one may observe the essential elements required to achieve success, namely 1) A long-term and sustained effort in terms of resources and advocacy; 2) ongoing policy monitoring and review; and 3) a service orientation in creating value for authors.

In terms of resources and advocacy, QUT has built an open access infrastructure and hired personnel. By autumn 2010, the repository held more than 24,000 records, including more than 13,000 full-text open access items (1,500 of them were under embargo at that point in time, but accessible via the 'request copy' button). The deposit rate has been higher than the annual output of publications, an indicator of the backfilling of the repository. Several years of dedicated effort have resulted in an open access rate of more than 50%. A validation of this success comes from the Webometrics Ranking of World Repositories (July 2010) in which QUT ranks at 20th position among institutional repositories, but QUT is a smaller university by international standards, and ranks tenth within Australia[13].

As regards policy review, QUT refines its policy every three years. In 2003, the policy stipulated the following (excerpt):

> *"Material which represents the total publicly available research and scholarly output of the University is to be located in the University's digital or "ePrint" repository, subject to the exclusions noted* (. . .). *The following materials are to be included:*

> - *Refereed research articles and contributions at the post-print stage* (*subject to any necessary agreement with the publisher*);
> - *Refereed research literature at the pre-print stage* (*with corrigenda added subsequently if necessary at the discretion of the author*);
> - *Un-refereed research literature, conference contributions, chapters in proceedings, etc;*
> - *Theses* (*as prepared for the Australian Digital Theses* (*ADT*) *process*).

> (. . .) *Material to be commercialised, or which contains confidential material, or of which the promulgation would infringe a legal commitment by the University and/or the author, should not be included in the repository."*

A revision in 2006 clarified two aspects. Firstly, the wording was altered to indicate that authors were not required to deposit both the pre-print and the post-print of an article, and they were free to determine the appropriate moment of deposit in accordance with the prevailing disciplinary culture. Secondly, it was clarified that the act of depositing was not subject to agreement with publishers, but that making the item open access could be subject to delays (embargo) or restrictions if the publisher's terms and conditions so demand. Another revision in 2009 restated the distinction between the mandatory deposit of peer-reviewed manuscripts (journals, conferences), and an encouraged deposit of other materials, the list of which was extended to data and books.

As regards service, the repository has been positioned as empowering authors in accelerating dissemination, retaining rights over their publications and reaching a wider audience. The library collects download data from the QUT ePrints service for analysis. Some of the items downloaded most frequently are papers, which are of interest to the local community, a wider potential audience, or of relevance to industry. For the fifty most downloaded QUT authors it has been possible to collate citation data (obtained from Scopus) for the years before the authors started self-archiving (e.g. 2003 and earlier, if the author was at QUT already in 2004) and for the years in which self-archiving was comprehensive. Correlational data indicates that citation rates increased significantly after authors began making open access copies of their work available via QUT ePrints. Some of the researchers at QUT frequently deposit working papers in subject-based repositories (e.g. arXiv in Physics, RePEc in economics). For these authors, QUT is improving the interchange of items with subject-based repositories to ensure that authors must deposit only once (unless they chose to update the deposited version later). This strategy of integration maximizes the visibility of QUT authors in their scholarly community.

2.2 Refining policy to foster deposit: University of Zurich

The University of Zurich adopted an open access policy 'requiring' deposit in July 2005. However, the original wording in German had been 'erwartet', i.e. 'expected' – a weaker term. Moreover, only in October 2006 did the repository ZORA (Zurich Open Repository and Archive) go live. Initial deposit rates remained low. In May 2008, the university enhanced its policy, requiring researchers to deposit[14,15].

This is an important case study for shifting open access policy from recommendation to requirement. We may elaborate on issues examined such as 1) the resources needed to push for deposit, to retain editorial control, and to develop policy; 2) ensurement that the repository is compliant with copyright law and publishing contracts at all times; 3) the sustained advocacy required to overcome faculty

reluctance in accepting the author's final peer-reviewed manuscript as an authoritative version; 4) the advanced information technology needed to further develop functionalities for authors. The policy guidelines of May 2008 stipulated three essential elements:

- The university requires its researchers to deposit a copy of their publications in ZORA with open access unless legal obstacles prevent this;
- The university encourages its researchers to publish in open access journals whenever possible and contributes towards article processing charges;
- The annual reports of the university will be based on ZORA, and only publications registered in ZORA will be included (in practice, this refers to metadata only and does not depend on full text deposit).

During the years from 2005 to 2007, the number of available records is about 1,000. In 2008 and 2009, the number rose above 7,000, essentially capturing the entire output of the university. Researchers wanted their publications to be incorporated into the annual reports. An open access team chased authors for full texts. In 2009, excluding books, about 60% of the records include a link to the full text. However, a significant part is permanently inaccessible due to copyright restrictions. About 42% of publications are open access.

Over the past years, a task force has been assembled to deal with the spike of reporting in ZORA that occurs as the deadline for the annual report approaches. The repository is staffed for editorial control of metadata, peer review status and copyright checks. Courses are organized for ZORA-submitters.

Ensuring lawfulness and adherence to contracts was considered so important that the university commissioned a special report. While an extensive FAQ aimed at researchers was developed subsequently, the task of determining the copyright situation for every publication remains. Databases covering many of the large international publishers may facilitate this, and the open access team has compiled an additional database (covering more than 1,000 journals), but the check needs to be made for every item and the permissible version obtained from the author or publisher, with preference given to the final published version wherever it may be displayed[14].

To win over faculty, much advocacy was conducted, particularly in establishing the postprint as an authoritative version. For instance, there is confusion as to what the postprint is, and how it differs from the publisher's version. Also, scholars in the humanities were often reluctant to engage in any activity that potentially might harm the small and specialized publishers they collaborate with. The open access team responded by emphasizing that it was seeking deposit only of texts that had passed quality control (peer or editorial review), and that it would only make

accessible texts (postprint, publisher pdf) in compliance with the law and publishers policies.

As ZORA grew, researchers began monitoring deposits, and expecting enhanced functionality of the platform. For example, authors may export metadata to other websites in multiple ways, and use them for reporting purposes. Also they track usage of their items in detail.

2.3 National platform, open collection, decentralized policy: HAL (June-October 2006)

Hyperarticle en Ligne (HAL) is a national platform[16]. HAL was conceived as a response to a temporary outage of arXiv. It was first implemented in 2001 by the newly created Centre pour la Communication Scientifique Directe (CCSD) of the Centre National de la Recherche Scientifique (CNRS). The remarkable achievement of HAL is to be a dynamic platform in a decentralized environment, compatible with a patchwork of local policies.

So far, we have looked at institutional mandates. This case enables two contrasting observations, namely on 1) what it means to develop a joint platform; and 2) how a good open access rate may be achieved without a formal mandate.

The transformation of HAL into a national platform was incremental, but one pivotal moment was the summer of 2006, when agreement was reached among the four largest research organisations, i.e. CNRS, Institut national de la recherché en informatique et automatique (INRIA), Institut national de la santé et de la recherché médicale (INSERM), Institut national de la recherché agronomique (INRA), as well as with the national association of universities and of the Grandes Ecoles. Yet, of the 49 research institutions, 89 universities and 70 Grandes Ecoles that have adopted HAL, only a few have a deposit mandate.

Nevertheless, in 2010, all French repositories, cumulatively, passed the threshold of 2 million items (much has been contributed by the digitization of nation journal archives, which often are open access or, at least, open access behind a moving wall). It is estimated that HAL holds about 450,000 records. In September 2010, the homepage of HAL indicated the availability of about 150,000 open access full texts (of which about 40,000 stem from the archive of the Journal de Physique, about 20,000 are theses, and about 10,000 are publication of the European Geosciences Union). From 2006 to 2008, it was estimated that HAL captured about 7%–8% of the French national research output[17–18]. Overall, it would seem that the virtue of HAL is to allow a whole country to move forward. Yet, what distinguishes an institution with a good open access rate?

INRIA, for example, is one of the important contributors to HAL. From 2007 to 2009 more than one third of the items logged in the repository are open access.

Among the 8 INRIA institutes (located around France), the open access rate varies from approximately 25% to 50%. As there is no uniform collection policy at INRIA, institutions and their sub-units must decide what to include. So INRIA collects scholarly output (conference proceedings, journal articles, books), technical reports, teaching materials and theses and dissertations. Among 2,964 items logged in HAL for 2007, 1,288 were conference proceedings, 763 journal articles, 430 technical reports, 142 PhD theses and 34 books (a typical distribution). For repositories overall it is estimated that about one sixth of deposits is 'grey' literature, i.e. non-published materials like a thesis or report[19].

Table 1　Percentage of open access publications of INRIA during 2007–2009

	Publications in annual report	Items logged in repository	Deposit or linked full text	Percentage of open access publications (%)
2007	4,977	2,964	1,820	36.6
2008	5,153	3,329	2,023	39,3
2009	5,868	3,423	2,154	36,7

The major reasons for the success were as follows:

- The rapid accumulation of a high number of full texts in a few months (2006) that increased the value and visibility of the repository, which is also important to the authors that have deposited. HAL INRIA is currently ranked the eighth among all repositories and the third biggest among the institutional repositories worldwide (Webometrics Ranking of World Repositories, July 2010);
- A disciplinary culture highly compatible with self-archiving and a deposit policy that encourages the deposit of a wide variety of items, including conference papers and reports, and material which, strictly speaking, is not of a scholarly nature, such as texts of popular science tutorials, posters and so on. All of this drives up the value and visibility as outlined above;
- A set of service tools for authors (e.g. author identification, publication list export, webpage creation) that provide an incentive to record and archive publications.

2.4　Maximising a funder's impact: The Wellcome Trust

The Wellcome Trust is the first mover among funders. Being a subject-specific funder of life sciences, implementation centres on a subject-based repository service (UK PMC). Its open access policy came into effect in October 2006[20]. The policy requires funded authors to make a copy of their journal article freely available through the subject-based repository UK PMC within six months. Two routes are open to grantees, either to publish in an open access journal or to publish in a journal that permits self-archiving.

The Wellcome Trust has monitored its open access policy in the following main ways:

- Compliance. As the Wellcome Trust expects deposit within six months after publication, the rate of deposit for the last six months is observed continuously. As a first benchmark, the Wellcome Trust has achieved a deposit rate of ca. 45% after three years, with four fifth of manuscripts having been submitted by publishers (green or gold), and the rest deposited through the author submission system.
- Deposit. At the end of any grant, the grantees are asked to report results and all manuscripts they have deposited. The Wellcome Trust can then request deposit of all artciles not yet submitted although this may be beyond the requested six months.
- Article-processing charges. Grantees may use their funds to pay publication charges. Additionally, block funding is provided to the home universities of grantees to cover open access publication charges. At 30 universities, at which many grantees reside, authors may draw on these funds beyond the lifetime of their grant to pay for open access publication charges.
- Finances. In the financial year of 2009–2010, the expenditure for open access was £2.9 million for an open access rate approaching 50%. Given that the Wellcome Trust's total research spend of one year is ca. £650 million, the Wellcome Trust expects that full open access would amount to 1% of the research budget.

Different attitudes may be observed between research institutions and funders. Funders adopting open access wish to maximize the impact and re-use of research results automatically and immediately, and open access publishing delivers this; whereas self-archiving necessarily includes a delay due to the embargo that must be respected. Moreover, self-archiving by authors has its own issues (cf. the University of Zurich above and Pretoria below). A remedy may be that publishers deposit manuscripts on behalf of the author (though this is a somewhat paradoxical situation, i.e. author self-archiving becomes a publisher service). Overall, funders are able to anticipate publication charges as part of research costs and may budget reliably. This indicates that research funders are likely to prefer open access publishing over self-archiving.

2.5   Implementing open access as research infrastructure: UK PMC

Set up in January 2007, UK PMC was backed initially by 8 institutions, which are mainly public and private research funders in biomedicine and health, but including

some research performing organisations[21–22]. The British Library has delivered the repository with several partners contributing to service improvements.[1] Notable have been the deployment of a single search interface (http://ukpmc.ac.uk/); the development of software to enable citation counting and surfing; and steps towards enabling text mining of the full corpus. By 2009, it registered about 750,000 down-loads, as much as other global subject-based repositories. International funders have joined UK PMC in an effort to create Europe PMC[23].

This case study demonstrates two things: 1) If research funders are serious about open access, they must invest in infrastructure, and this makes them a new and significant stakeholder in scholarly publishing. Publications are no longer primarily a product to be purchased by libraries, but an input for new information infrastruc-tures; and 2) motivations and goals of funders will be interpreted in a new way. This shift is well stated by the Medical Research Council (MRC) in the following[24]:

> *"If an open access fee has been paid MRC requires authors and publishers to licence research papers such that they may be freely copied and re-used for purposes such as text and data mining, provided that such uses are fully attributed. This is also encouraged where no fee had been paid."*

UK PMC is a national add-on to an existing infrastructure, which was first built in the United States by the National Institute of Health. PMC Canada and Europe PMC follow the same logic, and the various portals constitute entry-points that offer differing, partly overlapping, partly competitive services based on a global and shared infrastructure. The policies of UK PMC's backers are similar. All policies mandate deposit and grant a maximum embargo period of 6 months after publication. There is a clear preference for the retention of copyright by authors. If an article processing charge has been paid, all funders require a re-use license. Publication in journals not compliant with the stated policy is strongly discouraged.

The motivation and outlook of the backers may be explored. For example, the MRC (deposit mandate October 2006) was interested in a solution corresponding to its national community of grantees, which includes 30 research units and institutes it directly sponsors. In collaborating with the Wellcome Trust, the MRC pro-actively shaped the infrastructure, UK PMC is an essential element of MRC's policy because

---

1 European Bioinformatics Institute, MIMAS (Manchester, a national data centre), and the National Centre for Text Mining.

- As a collection of outputs, MRC funded research will be more visible (in due course a complete collection of primary publications);
- Nationally, it brings all the relevant disciplinary research funders to the table. Facilitating collaboration will in turn make every investment of UK PMC more efficient;
- The repository provides a distinct interface and added value to the MRC and its community of intramural researchers and grant holders, for example, through the mining and re-use of research results.

Arthritis Research UK also joined UK PMC at the beginning, keen to collaborate with other biomedical research funders in exploring the future of digital scholarship[22]. As small specialist funder at the time being, the relationship with grantees is particularly important for Arthritis Research UK, hence there is a strong policy emphasis on persuasion by improving the functionality and service of UK PMC because

- The improved interface and additional features enable to promote the service for grantees, presenting their search results in a wider subject-based context;
- A growing corpus of content on UK PMC will support public access and open learning on the issue of arthritis;
- Publications as 'pure outputs' of research will gradually become more embedded in a wider digital research and learning infrastructure.

The Health Research Board (HRB, Ireland) joined UK PMC in January 2010[25]. This enabled HRB to implement a mandatory open access policy following the review of an HRB open access position adopted two years previously. HRB has some grantees outside academic institutions. It joined UK PMC for the following reasons:

- A funder-led subject-based repository is most effective for a direct and monitored implementation because all content is ingested and made available at a single site;
- Research outputs funded by HRB become visible internationally alongside other subject-specific outputs in a single repository with varying access portals, increasing usage and the possible uses of published outputs;
- HRB is able to become a player in the further development of subject-specific digital infrastructures by partnering with other funders;
- Joint projects of funders will enhance their bargaining power with publishers when it comes to securing green open access, buying out copyright to foster re-use and developing open access publishing solutions.

National Science Library,
Chinese Academy of
Sciences

2.6 Integrating journals with a global research infrastructure: SCOAP3 (April 2007)

The Sponsoring Consortium for Open Access Publishing in Particle Physics (SCOAP3) revives the dream that open access will enable (public) research institutions to cap the cost of scholarly publishing. SCOAP3 has promised to deliver a publishing system that could be integrated into the new digital research infrastructures seamlessly[26]. This requires collaboration on all continents to collect a €10 million budget envelope, and an inclusive international decision-making process when tendering journal publishing. This will result in open access publishing as a universal standard for a research field of high-energy physics (HEP). SCOAP3 is the first trial run for the global conversion of journals to open access.

This case demonstrates an alternative route of implementing open access, not via institutional deposit but international collaboration, and not by sponsoring articles but converting journals. Two things need to be noted: 1) strategy must initially focus on building a consortium, and this means open access is not as immediate as in the case of mandatory deposit; but 2) the innovative mechanism of tendering journal publications allows cost control, gives publishers a stake and feeds the version of record directly into scientific information infrastructures.

For collecting the budget envelope, shares are allocated by country, based on 'peer review usage'. For example, the United States is expected to contribute by 24.3%, Germany 9.1%, China 5.6% and Brazil 2.7%. By September 2010, pledges amounting to more than 70% of its yearly budget envelope had been collected. However, implementation in the United States presented formidable challenges as pledges had to be collected individually from a large number of libraries, which purchase access to HEP titles by subscription. By September 2010, however, nearly the whole projected budget of $3.5 million had been pledged by almost 200 institutions, what remains is to extend SCOAP3 beyond the North Atlantic. The necessary contributions from Japan (7.1%), China (5.6%), Russia (3.4%), India (2.7%) and Brazil (2.7%) have been missing so far. While SCOAP3 is confident that support in these countries will materialize, the project is equally adamant that SCOAP3 will only move ahead if pledges from these outstanding countries have been made (amounting to 21,5% of the budget).

In HEP, four fifth of articles are published in 6 leading journals owned by 4 publishers, and more than 60% of the articles appear in journals owned by not-for-profit (society) publishers[27–29]. When tendering, journal publishers would be required to make their content open access in a manner that connects price to quality and volume. Assuming the capped budget envelope, this requires a ranking of journals by quality (e.g. based on bibliometrics) in combination with the

suggested price per article. Based on the proposed journal volume, SCOAP3 would award contracts until it has reached its limit. If publishing will underpin digital research infrastructures in new ways, then all articles must be pushed into a repository and be irreversibly available in open access with wide-ranging use and re-use rights, for the general public, researchers, and the machines of e-Science.

Publishers will be an integral part of the tendering and will be free to tender their journals for conversion, or continue running them on the traditional model. Publishers may also launch new journals. That said, in the past two decades, new journals aimed at the HEP community have not attracted submissions in significant numbers. This gives the existing HEP publishers a stake in pushing forward SCOAP3.

2.7   Moving towards 100% open access: Howard Hughes Medical Institute

The Howard Hughes Medical Institute (HHMI) is a noteworthy pioneer of open access. In April 2003, the Bethesda statement on Open Access Publishing was agreed upon at the HHMI headquarters (Maryland). The HHMI policy, effective since January 2008, stipulates that[30]:

> *"An Institute laboratory head is responsible for ensuring that each original, peer-reviewed research publication on which he or she is a major author is freely available and downloadable on-line within six months of publication. "Major author" normally includes both the first and last authors: if a middle author is designated in the paper as the corresponding author, then that author is also considered to be a major author."*

The case of the HHMI is significant because it differs from other institutional cases in that 1) the policy targets senior researchers; and 2) the institution has made an effort to increase the deposit rate as quickly as possible by striking bargains with publishers.

Institutional mandates usually cover all researchers. The HHMI, however, requires only principal investigators (laboratory heads) at HHMI's Janelia Farm Research Campus and at the universities to comply with its policy. On the one hand, this is easier to monitor because these are the researchers directly funded (and not just employed on projects). On the other hand, the senior researchers are expected to lead by example. In any case, one monitors compliance for a small and homogenous group of about 400.

In 2008 and 2009, more than 3,500 articles were indexed cumulatively in PubMed, for which more than 2,500 full-text versions may be found in PubMed Central. This indicates an open access rate of more than 70%. HHMI demands public access within six months and pays publishers extra. For HHMI, four main routes exist to deposit (which are the same for the NIH manuscript submission system):

- Route A. From journals that directly deposit final published articles in PMC (e.g. from an open access publisher);
- Rout B. Publisher depositing the final published article or accepted manuscript version in PMC on behalf of the author (e.g. from a hybrid OA journal);
- Route C. Self-archiving of the final or accepted version peer-reviewed manuscript in PMC via the NIH manuscript submission system;
- Route D. Verification and completion of a submission process for a final peer reviewed manuscript that the publisher has deposited.[2]

For the second route, HHMI has transferred a list of its principal investigators to publishers, and these notify the HHMI author upon acceptance of the transfer of the final manuscript as so-called 'stage two version' to PMC. The author confirms that it is her or his manuscript, and PMC prepares the manuscript for final release (after the embargo of up to six months) and the author approves this final version.

The last route of deposit is a variation on green open access, based on publisher deposit. For many journals the publishers demand longer embargoes than six months. Hence, the HHMI was to strike deals with the publishers to achieve a reduction of the embargo. For example, for articles from Cell Press (imprint of Elsevier) US $1,000 are paid to secure deposit and access within six months, whereas Elsevier and Wiley demand US $1,500 (Table 2).

Table 2   Stylized comparison of payments to publishers to achieve green or gold open access

	HHMI	Wellcome Trust
Elsevier journals	Charged US $1,500	Charged US $3,000
	Author manuscript version — after peer review — deposited by Elsevier in PMC	Final version of paper deposited in PMC by Elsevier
	Paper available in PMC/UK PMC at six months after publication	Paper available in UK PMC/PMC at time of publication
	No re-use licence	Full re-use licence
Cell Press (imprint of Elsevier)	Charged US $1,000	Charged US $5,000
	Author manuscript version — after peer review — deposited by Elsevier in PMC	Final version of paper deposited in PMC by Elsevier
	Paper available in PMC/UK PMC at six months after publication	Paper available in UK PMC/PMC at time of publication
	No re-use licence	Full re-use licence

---

2   Verification is often part of the process for route B) and C) as well.

A difference may be observed between the Wellcome Trust and HHMI. The Wellcome Trust prefers to pay for open access publishing, while HHMI prefers to pay for repository deposit. The contrast clearly indicates the difference between open access publishing and green open access (at six months).

2.8   Open access to all publications, internationally: Austrian Science Fund

The Austrian Science Fund (FWF) is the research funder with the most comprehensive open access publishing policy: Publications and data, articles and books, nationally and internationally[23].

This case study allows us to anticipate the road a research funder is likely to travel, particularly in the European Union. The FWF recognizes open access as a fundamental shift in scholarly communication, leading to new ways of presenting, reading and using research results. Transition costs are anticipated. As main benefit, advances in the mode of knowledge production are expected (e.g. increased speed, improved peer review, broader impact, interlinkage of text and data). Additional benefits are a strengthened relation between scholars and the public (through public access), and a reinvigoration of the market for scientific publishing.

In line with the aim of aiding system change, the FWF stipulates in detail that

- Open access publishing charges may be covered from FWF grants;
- Additional support for publication charges may be requested up to three years after the end of any project;
- Monographs, proceedings and collections are also covered by the policy;
- In the life sciences, the preferred solution is for all papers to be made available through UK PMC regardless whether published in open access, deposited by publishers or self-archived by researchers themselves;
- Research data is also to be made available via repositories within two years after the end of any project.

The FWF policy reflects some notable challenges facing funders. Since 2004, the FWF has been supporting open access journal publication charges. A first deposit request was issued in 2005. The integrated policy, as outlined above, was in force from March 2008. In a consolidated version of 2010, the FWF addresses the following issues. Firstly, research papers are normally completed only after the lifetime of a funded research project, but the FWF extends the deadline for funding publication charges. In 2010, the FWF funded about 350 open access publications with a total cost of approximately €400 k. Secondly, generic research funders support disciplines that publish lots of books, but FWF supports open access to books with up to €6,000 for immediate open access and €4,000 if embargoed for no more

National Science Library, Chinese Academy of Sciences

than 12 months (between 30 and 40 books have been funded in 2010 at a cost of approximately €150 k). Thirdly, European national research funders have been seeking international collaboration, the FWF is willing to support international solutions like UK PMC (with several hundred articles ingested in 2010). Internationalisation is backed by the FWF policy of soliciting international peer review for project applications as well as the review of books submitted for an open access subsidy.

Compliance is monitored by FWF through project reports, in which grantees have to provide justification if any research result is not open access. In 2009, the FWF spent about €350 k on open access from a budget of €150 million, but expenditure is expected to rise due to cooperation with UKPMC and the funding of book publication charges to €550 k in 2010, which is ca. 0.3% of the total FWF budget.

2.9   Open access, applied research and industry: Fraunhofer-Gesellschaft

The Fraunhofer-Gesellschaft is Europe's largest application-oriented research organization with 60 thematically diverse institutes, some of which have shared infrastructure concepts, while others have highly customized information technology infrastructures and specialized libraries. The close cooperation with industry partners calls for careful publication planning with respect to confidentiality and patent registration.

This case introduces further variation, namely by showing that 1) while most funders and institutions focus on basic research, Fraunhofer demonstrates that open access may be equally relevant for applied research; and 2) a coherent policy may be implemented in a country that has a strong legal and moral tradition of 'scientific freedom'. This is often understood to imply that researchers cannot (and should not) be mandated to do something. Fraunhofer's board of directors adopted an open access policy for all publications in July 2008, with an embargo up to 12 months as acceptable, and stated[31]:

"The Fraunhofer-Gesellschaft is committed to providing the necessary financial, organizational and non-material support that will allow the concept of open access to be implemented under optimum conditions."

Fraunhofer-Publica (http://publica.fraunhofer.de), a bibliographic database with 130,000 records, has been supplemented by a server providing open access to full texts (http://eprints.fraunhofer.de). A system of incentives and rewards is being developed to foster adoption by leading researchers. Senior researchers are requested to support junior colleagues. Authors are encouraged to retain the rights to their

National Science Library, Chinese Academy of Sciences

publications. A service infrastructure is offered to support authors in publishing and depositing. Open access advocacy, institute-by-institute, complements this strategy. A lesson learned by Fraunhofer's central publication support is that from a researcher's perspective the notion of open access is encountered somewhere towards the end of the publication process. Increased efforts are required to help authors consider open access as part of their publication strategy. The idea is to introduce management routines that establish open access as a natural part of the publishing process for each institute.

Users from industry are interested mainly in direct access to the full text. Hence Publica now includes links to relevant external content, for example, 6,800 URLs (Uniform Resource Locator) point to patents granted to the Fraunhofer-Gesellschaft, the details of which are accessible through the database of the European Patent Office. Furthermore, 17,000 DOIs (Digital Object Identifiers) point to free down-loads for documents that, for example, have been provided through national or campus licenses.

Since 2000, researchers at Fraunhofer have published about 18,000 journal articles and 23,500 conference papers (till September 2010), of which 5,300 are available as full text. However, Fraunhofer estimates that the amount available in open access could and should be tripled to about 40% of publications. Furthermore, value has been added to the repository by systematically making the so-called 'grey' literature of working papers, presentations and posters available.

2.10   Second Generation Open Access: The University of Pretoria

The University of Pretoria is an important case because its open access policy development has been reflexive, i.e. the policy was drafted by comparing policies and their implementation globally, and its policy implementation is aligned with the national research assessment scheme[32].

For a second generation open access policy, three things are worth mentioning, namely that 1) policy formulation and adoption may be enhanced and eased by a reflexive approach; 2) faculty members nevertheless must be persuaded extensively; and 3) open access may aid research evaluation but not the other way round.

Open access advocates at the University of Pretoria realized that an institutional policy would be more effective if:

- A series of experiments result in demonstrators, which convey the impact of the repository while minimizing any extra work for the scholars;
- Extensive advocacy on the benefits to the institution and scholars is matched by a readiness to listen to concerns and misgivings about the method and proc-ess of open access (e.g. deposit process, copyright addenda);

- A mandatory policy is only instituted once the necessary authority and resources for implementation are secured.

Awareness has been raised among faculty members, but the Open Scholarship Office nevertheless encountered the following challenges:

- Faculty members have queries and reservations about the post-print. As publications are part of an assessment system, the publisher's version seems all-important. Faculty voice reluctance because the green version of any article (i.e. author's final peer-reviewed manuscript) is not identical to the publisher's version.
- Faculty members tend to focus on the extra work, not the extra benefits of green open access, which makes every conversation about the (potential) benefits of green open access difficult. Notoriously, faculty members will post the publisher's pdf on a personal website to foster dissemination of their work, but be reluctant to provide a green version to the repository, even if the later offers superior possibilities for discovery.
- Librarians and the Open Scholarship Office increasingly assume responsibility for submissions: in the interest of building an open access collection and maintaining momentum, it becomes necessary for the Open Scholarship staff and the librarians to archive on behalf of the authors. Thus at the University of Pretoria,, self-archiving hovers below 5% for 2010, meaning that more than 95% of the papers are deposited in a library assisted deposit system.

Since 1997, South African research institutions must submit annual research reports in recognized national and international databases, for which the institution receives (extra) funding. At the University of Pretoria, each faculty has a Research Information Systems coordinator. The Open Scholarship Office offered support and additional services in exchange for a chance to capture the full text. Repository staff was able to trace published items that had been missed earlier and to add a Uniform Resource Identifier to many publications, thus enhancing their visibility.

The effort to improve the position of the university (i.e. more funding) via open access came before the move to mandate deposit. Yet there are some difficulties: Firstly, every evaluation connected to the distribution of life chances (e.g. funding, tenure) must be based on authoritative documents, and the system must be considered very fair. Currently, only the published version is seen as authoritative (no comparable green open access version has yet been standardized) and only the systems offered by large commercial firms (e.g. Thomson Reuters Journal Citation Report or Elsevier's Scopus) are trusted. Secondly, any distribution of life chances by systemic evaluation has resulted in much negative news (e.g. rejection letters, less

funding than competitors), which means that participants may comply, but often comply unwillingly. Hence, research assessment schemes would not seem to lower faculty reluctance vis-à-vis the deposit of the green version of manuscripts.

3 Suggestions for open access policy development

The incremental approach in studying cases should foster the emergence of second-generation open access policies by enabling a systematic appraisal of earlier success and failure. The case studies mentioned above offer numerous examples of revised and refined policy, and also indicate how research institutions and funders have learned from each other.

Digital publishing is changing within an emerging system of cyber-infrastructures and data-driven science[31]. Particularly, the funders are directly contributing to its emergence. A certain homology may be observed between the emergence of digital research infrastructures and the call for open access to published results. It is a homology implicitly assumed in the Bethesda statement and the Berlin Declaration, based on a fundamental compatibility between the norms and economics of cyber-science, open access and the knowledge economy[33]. In my opinion, open access advocates might centre their vision on integrating open access with research infra-structures. Publishing research results will then become a service whose rationale is to feed that infrastructure with valid and reliable results that may be used in research, teaching and learning.

In most regions of the world, major digital infrastructure initiatives are under way. Scholarly communities recognize their importance. Funds are being mobilized. Research institutions and national governments are the major players[34]. Policy makers should ask themselves how to articulate open access as an essential part of the new infrastructure that merits institutional investment (in repositories, publication charges etc.). My suggestions are as follows:

- If open access is articulated as an essential and founding element of digital research infrastructures, the issue of 'embargoed' open access will go away. Digital infrastructures, by definition, operate in real time and therefore any sort of delayed access will become unacceptable to researchers. Of course, this has consequences for open access policy, particularly for green open access policy, and policy makers will reflect the implications;
- As published research results become an essential element of digital infrastruc-tures, the balance of bargaining power shifts considerably. Rather than having to try converting publishers and journals to open access (or invest time and money to start new ones), research institutions and national governments

National Science Library,
Chinese Academy of
Sciences

would be in a strong position because publishers and journals would be required to seek access to these new infrastructures, or face marginalization;

- With a focus on scholarly communities, it may be suggested in what areas research institutions and open access policy makers might concentrate their efforts. Open access publishing, integrating itself into everyday research is most advanced in the life sciences, economics, physics and computer science. In any of these areas large-scale projects could be envisaged, integrating published research results in community-specific infrastructures;

- With a focus on infrastructures, it is possible to envisage contracts, not just for open access publishing, but also for licensed content, which would tie these outputs to emerging research infrastructures. This would be particularly interesting if data and publications could be tied into the same infrastructure. Of course, there is a strong element of continuity in scholarly communication, which is based on communities. However, these communities, and their scholarly societies, are not capable of building and maintaining the new digital infrastructures. Research institutions and national governments will be organizing funding and governance. Hence, policy makers would need to understand if and how real time open access to the scholarly literature could be integrated into the emerging national and international research infrastructures.

References

1   Registry of Open Access Repository Material Archiving Policies (ROARMAP). Retrieved on December 3, 2010, from http://www.eprints.org/openaccess/policysignup/

2   Bailey, C.W. Transforming scholarly publishing through Open Access. A Bibliography. Self-publication, Retrieved on December 3, 2010, from http://digital-scholarship.org/tsp/w/tsp.html

3   Suber, P. SPARC Open Access Newsletter. Retrieved on December 3, 2010, from http://www.earlham.edu/~peters/fos /newsletter/archive.htm

4   Budapest Open Access Initiative. Retrieved on December 3, 2010 from http://www.soros.org/openaccess

5   Berlin Declaration on Open Access. Retrieved on December 3, 2010, from http://oa.mpg.de/

6   Glaser, B., & Strauss, A.L. The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine De Gruyter, 1967.

7   Strauss, A.L., & Corbin, J. Basics of qualitative research: Grounded theory procedures and techniques, 2nd edition. Thousand Oaks: Sage, 1998.

8   Hertz, R., & Imber, J.B. (Ed.) Studying elites using qualitative methods. Thousand Oaks, CA: Sage, 1995.

9   Meuser, M. & Nagel, U. ExpertInneninterviews — vielfach erprobt, wenig bedacht. In Garz, D., & Kraimer, K. (Eds.), Qualitativ-empirische Sozialforschung: Konzepte, Methoden, Analysen. Opladen: Westdeutscher Verlag, 1991: 441–471.

10   Mayring, P. Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim: Deutscher Studien Verlag, 2000.

11   Seale, C. The quality of qualitative research. London: Sage, 1999.

12   Queensland University of Technology. Retrieved on December 3, 2010, from http://www.mopp.qut.edu.au/F/F_01 _03.jsp

13   Ranking Web of World Repositories. Retrieved on Aug. 30, 2010, from 14 http://repositories.webometrics.info/top800_rep_ inst.asp.

14   University of Zurich. Retrieved on December 3, 2010, from http://www.oai.uzh.ch/

15   SHERPA RoMEO list. Retrieved on December 3, 2010, from http://www.sherpa.ac.uk/romeo/

16   Hyperarticle en Ligne (HAL). Retrieved on December 3, 2010, from http://hal.archives-ouvertes.fr/ & http://www.ccsd.cnrs.fr/

17   Schöpfel, J. & Prost, H. Développement et Usage des Archives Ouvertes en France. Rapport. 1e partie: Développement. HAL Deposit, 2010. Retrieved on December 3, 2010, from http://archivesic.ccsd.cnrs.fr/sic_00497389_v1/

18   Barruch, P. Open Access developments in France: The HAL Open Archives System. Learned Publishing, 2007, 20(4): 267–282. Retrieved on December 3, 2010, from http://hal.archives-ouvertes.fr/ & http://www.ccsd.cnrs.fr/

19   Schöpfel, J., & Stock, C. Grey literature in French digital repositories: A survey. Deposit at HAL, 2008. Retrieved on December 3, 2010, from http://archivesic.ccsd.cnrs.fr/sic_00379232/en/

20   The Wellcome Trust. Retrieved on December 3, 2010, from http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Open-access/index.htm.

21   UK PubMedCentral. Retrieved on December 3, 2010, from http://ukpmc.ac.uk/About

22   Arthritis Research UK. Retrieved on December 3, 2010, from http://www.arthritisresearchuk.org/research/our_ research_policies/open_access_policy.aspx

23   Austrian Science Fund (FWF). Retrieved on December 3, 2010, from http://www.fwf.ac.at/de/public_relations/oai /index.html

24   Medical Research Council UK. Retrieved on December 3, 2010, from http://www.mrc.ac.uk/Ourresearch /Ethicsresearchguidance/Openaccesspublishing/index.htm

25   Health Research Board, Ireland. Retrieved on December 3, 2010, from http://www.hrb.ie/research-strategy-funding/policies-and-guidelines/policies/open-access/

26   Scholarly Consortium for Open Access Publishing in Particle Physics. Retrieved on December 3, 2010, from http://scoap3.org/

27   Mele, S. et al. Quantitative Analysis of the Publishing Landscape in High-Energy Physics. Journal of High Energy Physics, 2006, (12): 1. Retrieved on December 3, 2010, from http://iopscience.iop.org/1126-6708/2006/12/S01

28   Mele, S. SCOAP3: Where libraries, scientists and publishers meet repositories, peer-reviewed journals and Open Access. Serials Review, 2009 35(4): 264–271

29   Gentil-Beccot, A., Mele, S., & Brooks, T.C. Citing and reading behaviour in high-energy physics. How a community stopped worrying about journals and learned to love repositories. ArXiv, 2009. Retrieved on December 3, 2010, from http://arxiv.org/abs/0906.5418

30   Howard Hughes Medical Institute. Retrieved on December 3, 2010, from http://www.hhmi.org/about/research/ policies.html

National Science Library,
Chinese Academy of
Sciences

Research Papers

31 Fraunhofer-Gesellschaft. Retrieved on December 3, 2010, from http://www.fraunhofer.de/en/ Images/OpenAccess Policy_tcm63-828.pdf

32 University of Pretoria. Retrieved on December 3, 2010, from http://www.ais.up.ac.2a/jour-nals/journalsopen.htm

33 Hey, T., Tansley, S., & Tolle, K. The fourth paradigm: Data-intensive scientific discovery. Seattle: Microsoft Research, 2009.

34 Armbruster, C. Cyberscience and the knowledge-based economy, open access and trade publishing: From contradiction to compatibility with nonexclusive copyright licensing. International Journal of Communications Law and Policy, 2008, Issue 12. Retrieved on December 3, 2010, from http://www.ijclp.net/issue_12.html

(*Copy editor*: *Ms. Jing CAO*)