Scholarly communication in the context of Open Science: the need for infrastructures

Pierre Mounier
Ecole des Hautes Etudes en Sciences Sociales
OpenEdition / OPERAS

2nd Meeting of the Expert Group on Scholarly Publishing and Scholarly Communication 26/03/2018
(Introduction): starting points...
Origin

Executive Summary

The emergence of the Internet has transformed the practice of the humanities and social sciences—more slowly than some may have hoped, but more profoundly than others may have expected. Digital cultural heritage resources are a fundamental dataset for the humanities; these resources, combined with computer networks and software tools, now shape the way that scholars discover and make sense of the human record, while also shaping the way their findings are communicated to students, colleagues, and the general public. Even greater transformations are on the horizon, as digitized cultural heritage comes into its own. But we will not see anything approaching complete digitization of the record of human culture, removal of legal and technical barriers to access, or revolutionary change in the academic reward system unless the individuals, institutions, enterprises, organizations, and agencies who are this generation’s stewards of that record make it their business to ensure that these things happen.

The organized use of networks and computation for the practice of science and engineering was the subject of a 2003 report to the National Science Foundation (NSF), Revolutionizing Science and Engineering through Cyberinfrastructure. In both the NSF report and this one, the term cyberinfrastructure is meant to denote the layer of information, expertise, standards, policies, tools, and services that are shared broadly across communities of inquiry but developed for specific scholarly purposes: cyberinfrastructure is something more specific than the network itself, but it is something more general than a tool or a resource developed for a particular project, a range of projects, or, even more broadly, for a particular discipline. So, for example, digital history collections and the collaborative environments in which to explore and analyze them from multiple disciplinary perspectives might be considered cyberinfrastructure, whereas fiber-optic cables and storage area networks or basic communication protocols would fall below the line for cyberinfrastructure.
Putting down roots

Support and dissemination services assist authors and institutions in implementing OA and complying with OA policies. The most widely used are the SHERPA services, which are consistently relied on to support OA repository workflows.

OA repository services encompass a wide range of repository and related services that are crucial to the implementation of open access archiving. The critical dependency in this case is not on one or two individual services, but on the success of efforts to improve interoperability across the OA repositories landscape.

We also considered a range of OA publishing services and OA monitoring services. At the present time, policies are not highly dependent on these services but they seem likely to grow in importance in the future.

1.3. Priorities for action

The fundamental challenge for the implementation of OA policies is the need to develop a fully functioning OA infrastructure from the current disparate collection of services. This study has taken the form of a broad survey of the current OA policy and service landscape, and further work is needed to translate our findings into specific, costed recommendations. Nevertheless, the broad direction of travel is clear if OA policies are to be successfully implemented. Four priorities for action have therefore been identified:

1. Adopt sound governance structures with greater representation from funders and policy makers, promoting the wider use of crucial identifiers and standards.

2. Ensure the financial sustainability of critical services, particularly the DOAJ and SHERPA services.

3. Create an integrated infrastructure for OA repositories based on central ‘nodes’, interoperability across the broader landscape, and increased engagement with the European Commission’s OpenARE project and the work of the Confederation of Open Access Repositories (COAR).

4. Invest strategically in OA services in order to create a coherent OA infrastructure that is efficient, integrated and representative of all stakeholders.
OA monographs landscape

Summary: A landscape study on open access (OA) and monographs

Policies, funding and publishing in eight European countries

Publication date: March 2018

had significant institutional funds that were capable of supporting OA monographs. Monograph publishing models in Austria, Germany and the Nordic countries already rely on subsidies, and existing funders could consider making open access a condition of such payments in future (FWF in Austria has already done this).

Infrastructures for open access vary across countries. Some have taken steps towards legislation that encourages open access – Germany, for example, has included provisions for OA within federal copyright legislation, while France has enacted a Law for a Digital Republic which strongly encourages OA. The Netherlands has negotiated national deals with publishers to cover OA for all Dutch academics. Some countries have also made financial investments in publishing infrastructure to support OA; the most significant example is perhaps OpenEdition in France, a digital platform and ‘freenium’ business model to support digital and OA publishing, but the OAPEN platform in the Netherlands and the Finnish model of state-funded presses are also important.

Publishers and institutions also play an important role in providing the infrastructure for OA. Across all countries, most universities have a well established network of institutional repositories for academics pursuing the self-archiving route. Some, notably in Germany, also

Finally, there are some important disrupters stimulating new conversations about open access for monographs. In the UK, new university and academic-led presses are experimenting, not just with new business models but with new publishing formats, and modes of marketing, sales and distribution. In Germany, Language Science Press, a born-OA publisher, also offers open review and a comments option post-publication for monographs. Institutions from most countries in the study are signed up to Knowledge Unlatched, using crowd-sourced library funding to pay publication costs of OA monographs. These initiatives shift the conversation about what is possible.

Monograph publishing

Established monograph publishing practices affect the way that open access monographs are seen and the practical steps that must be taken to achieve them.

Funding models vary: in countries like Austria, Finland, Denmark, Norway and, to an extent, Germany, where many monographs are funded via subsidies, open access business models can repurpose the funds already associated with an individual book in order to make it OA. In very commercial environments like the UK and the Netherlands, where publishing costs are recovered primarily via sales, business models may require a more significant change. Although publishers
Open Access in Humanities and Social Sciences: a milky way
OPERAS-D Project

About OPERAS-D

The OPERAS-D (Design) project supports the 9 main partners ("core group") of the OPERAS network in the development of a European e-infrastructure for open access publications in the SSH. The project addresses long-term requirements for the development of the e-infrastructure and community building, as well as seeks to expand other interested parties within and beyond Europe, and in diverse fields of the SSH. To achieve this goal, the key objectives of the OPERAS-D project are to prepare a design study that defines governance models, structures and technical concepts for future services that the infrastructure will provide, and

Design for Open access Publications in European Research Areas for Social Sciences and Humanities

De 2017-01-01 à 2018-06-30, projet en cours

Détails concernant le projet

Coût total: EUR 408 586,25
Contribution de l’UE: EUR 408 586,25
Coordonné à/su(x)/en: France

Sujet(s): INFRASEPP-03-2016 - Support to policies and international cooperation for e-infrastructures
Appel à propositions: H2020-INFRASEPP-2016-1 See other projects for this call
Régime de financement: CSA - Coordination and support action

Objectif

OPERAS-D (Design) project will support the 5 main partners ("core group") of the OPERAS network in the development of a European e-infrastructure for open access publications in the Social Sciences and Humanities (SSH). The project will address long-term requirements for the development of the e-infrastructure and community building, as well as seek to expand other interested parties within and beyond Europe, and in diverse fields of the SSH. To achieve this goal, the key o...

Coordinateur

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS
RUE MICHEL ANGE 3
75794 PARIS France

Contribution de l’UE: EUR 100 932,50

Activity type: Research Organisations

Contact the organisation

Participants

ETHNIKO IDRYMA EREVINON

Greece

STICHTING OPEN*OPEN ACCESS PUBLISHING IN EUROPEAN NETWORKS

Netherlands

MAX WEBER STIFTUNG DEUTSCHE GEISTESWISSENSCHAFTLICHE INSTITUTE IM AUSLAND

Germany

KNOWLEDGE UNLATCHED CIC

United Kingdom
In examining all emerging trends in journal and monograph publishing, the report outlines key challenges and potential issues to be addressed by future initiatives. Recently introduced and experimental models (such as scholar-led publishing bodies, and new university presses) share common orientations towards increased participation of researchers in the publishing process and overcome certain deficiencies of the commercial publishing model.

Notwithstanding the importance of such initiatives, as the Report concludes, fragmentation (both in terms of the size and nature of publishers and of their business models) is a key characteristic in the academic publishing landscape. In this context, the main challenge in adopting effective open access publishing practices is to identify and assess current needs and limitations that permeate the academic publishing landscape, in operational as well as communicational terms.

The landscape study confirms that successful research relies primarily on unrestricted access to high quality scientific output and cross-disciplinary, international collaboration. Shared and remotely accessed digital infrastructures constitute an important feature towards the realisation of the European Research Area, and OPERAS aspires to be actively engaged in the implementation of a new mode of science that overcomes fragmentation and enables unrestricted access to high quality scientific output.
# Survey on publishers

## OPERAS Design Study

### July 2017

## Publishing software

<table>
<thead>
<tr>
<th>Software</th>
<th>Spring Survey</th>
<th>Autumn Survey</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>XMLmind</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>Wordpress</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>RUA</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>pub2web</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Pressbooks</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>PDF creator</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>OJS</td>
<td>25%</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td>MS Word</td>
<td>19%</td>
<td>19%</td>
<td>19%</td>
</tr>
<tr>
<td>Lodel</td>
<td>29%</td>
<td>29%</td>
<td>29%</td>
</tr>
<tr>
<td>Joomla</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>ipublishcentral</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>InDesign (Adobe)</td>
<td>12%</td>
<td>12%</td>
<td>12%</td>
</tr>
<tr>
<td>Hyperwave</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>Highwire</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Dreamweaver (Adobe)</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>Acrobat (Adobe)</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
</tbody>
</table>
OPERAS partners technical mapping

<table>
<thead>
<tr>
<th>PARTNER</th>
<th>Huma-Num (Isidore)</th>
<th>EKT</th>
<th>IBL PAN</th>
<th>MWS</th>
<th>Open Edition</th>
<th>OAPEN</th>
<th>OAPEN (DOAB)</th>
<th>Ubiquity Press</th>
<th>UCL press</th>
<th>Univ. Coimbra</th>
<th>University of Turin</th>
<th>Univ. Zadar</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB Size (GB)</td>
<td>1024GB</td>
<td>35GB</td>
<td>Open Edition [60GB]</td>
<td>100GB</td>
<td>60GB</td>
<td>60GB</td>
<td>200MB</td>
<td>270 GB</td>
<td>OAPEN [60GB]</td>
<td>419GB</td>
<td>8GB</td>
<td>78GB</td>
</tr>
<tr>
<td>Storage Size (TB)</td>
<td>400TB</td>
<td>50TB</td>
<td>Open Edition [40TB]</td>
<td>1,5TB</td>
<td>40TB</td>
<td>5,5TB</td>
<td>0,7TB</td>
<td>5TB</td>
<td>OAPEN [5,5TB]</td>
<td>25TB</td>
<td>2TB</td>
<td>1,25TB</td>
</tr>
</tbody>
</table>

- **MySQL**
- **MongoDB**
- **AFS**
- **Handle.net**
- **Virtuoso**
- **Exist**
- **BaseX**
OPERAS partners technical mapping

| PARTNER     | OPENEDITON  | Huma-Num (Isidore) | EKT  | Ubiquity Press | MWS  | Unito | OAPEN | OAPEN (DOAB) | Univ. Coimbra | Univ. Zadar | IBL PAN | UCL press |
|-------------|-------------|---------------------|------|----------------|------|-------|-------|---------------|----------------|-------------|----------|---------|-----------|
| CPU cores   | 724         | 192                 | 72   | 31             | 12   | 12    | 4     | 4             | 6              | 4           | OpenEdition | OAPEN    |
| Servers (nb)| 21          | 8                   | 6    | 10             | 1    | 1     | 4     | 4             | 4              | 6           | OpenEdition | OAPEN    |
| RAM (GB)    | 3000        | 1000                | 1100 | 115            | 24   | 32    | 32    | 32            | 28             | 8           | OpenEdition | OAPEN    |
| Provider    | CC IN2P3    | CC IN2P3            | EKT  | Amazon         | BSB  | Cineca| UvA   | SemperTool    | Univ. Coimbra | SRCE        | OpenEdition | OAPEN    |
Visibility of OA monographs

The Visibility of Open Access Monographs in a European Context

A report prepared by Knowledge Unlatched Research

30 January 2018

Cameron Neylon,†2 Lucy Montgomery,†2 Alkim Ozaygen,†2 Neil Saunders* and Frances Pinter†2

1. Centre for Culture and Technology, Carrie University, Kent St, Bernley, Western Australia
2. Knowledge Unlatched Research, United Kingdom

This project received funding from European Union’s Horizon 2020 research and innovation programme under grant agreement 731031.

institution). This means that we clearly need to understand what the usage data is telling us before we have any use for it.

A number of respondents expressed a desire for a “dashboard” or other visualisations that could bring multiple data sources together. The consequent need for data integration and standardisation to achieve this was mentioned in one or two responses but awareness of the challenges of comparison across sources appeared to be limited. There was some evidence of a conflation of visualisation with data integration.

Respondents are small organisations with limited capacity. There is a desire for coordination and shared services, infrastructures, standards. A common thread in the responses was that the publishers and platforms who are engaged in Open Access scholarly book publishing are relatively small. This is both a challenge and an opportunity. They have limited capacity to develop internal processes and systems are looking for shared services and platforms to assist in developing usage data capabilities.

It would be of great help if we could have a main service from where we could manage all the information related to statistical usage data.

[to engage more effectively with usage data we would like a]...consortium agreement with Google on how to gather and assess usage data.

We would like to see a usage aggregation service that consolidates usage data from different hosting partners into one standardised report in an automated way. In turn, this should translate into a usage dashboard that can be embedded into platforms and allow customers to use different filters to analyse usage by publisher, region, etc.

[one of our biggest challenges is...optimizing workflow, how to do more work with small resources.

What emerges overall is a picture in which platforms and publishers are implementing tools and approaches locally and using what they are provided with to some degree. There is generally a good technical awareness of the tools being deployed, but less apparent awareness of data curation and quality assurance issues.

Many of the challenges arise from issues of data integration and standardisation. Small, and even medium-sized, players have limited capacity to engage with detailed standards or technical development. Equally there are limitations on what capacity a small organisation can provide to investigate the meaning and context of the data being generated. The majority of data use seemed to be in pro-
KU Research findings

- The metadata held and managed by OPERAS partners is inconsistent and variable in quality.
- The visibility of OPERAS partner books in catalogues varies by publisher.
- Evidence can be obtained that books relevant to specific regions gain interest and attention in that region.
- The variable quality of book metadata creates challenges in analysing visibility consistently.
- The variable quality of book metadata creates challenges for downstream data aggregation and analysis providers.
OPERAS proposition: a distributed infrastructure for open scholarly communication in SSH
To provide a pan-European infrastructure for open scholarly communication

Integration of the long tail into Open Science
OPERAS is led by OpenEdition (France)

Consortium

- **35** Partners
- **11** Countries
- **9** Core group Members
- **2 H2020** Projects
OPERAS members serve the researcher’s needs **all along the research cycle**
A federation of publication platforms...

... to deliver Open Science services...
Coordinated by the Core Group
OPERAS Working Groups

1. **Advocacy**: MWS (contact point), AISA, UGOE, IBL PAN, KU Research, OpenEdition, University of Turin, The Arctic University of Norway

2. **Publishing Tools**: OpenEdition (contact point), Hypothesis, IBL PAN, C²DH, Roma Tre University, Stockholm University Press, Ubiquity Press, University of Turin, University of Milan

3. **Standards**: EKT (contact point), OAPEN, OpenEdition, University of Milan

4. **Business Models**: UCL Press (contact point), AEUP, IBL PAN, KU, KU Research, Lexis, OLH, UC Digitalis

5. **Best Practices**: OAPEN (contact point), AEUP, Hypothesis, LingOA, OpenEdition, OLH, QOAM, Lexis, Stockholm University Press, Ubiquity Press, University of Milan, University of Zadar

6. **Multilingualism**: UC Digitalis (contact point), UGOE, Huma-Num, EKT, ISCTE-IUL

7. **Platforms and Services**: OAPEN (contact point), OpenEdition (contact point), Huma-Num, LingOA, Open Books Publishers, QOAM, University of Zadar
2. Introduction: SSH and Open Science

SSH and OS are today inextricably linked with each other. OS practices are the next logical step for SSH, and the needs and insights of SSH are important for the development of OS.

This bond is not an accidental development of the SSH, their becoming open is the coherent fulfillment of the modern idea of knowledge as a process grounded on the public use of reason and confrontation with the widest possible audiences. However, the transition of SSH to OS has still to overcome many barriers.

The variety of tools and services currently available for research and publishing requires a mediation between two essential aspects of the intellectual work, namely, on the one hand, the individual, autonomous and self-responsible activity of the single scholar and, on the other hand, the cooperation between the scholars themselves and between scholars and institutions providing tools and services for publishing. The scholarly primitives of SSH, i.e., the basic activities underlying each research process, can only be carried out if supported by publishers who are able to provide the most advanced tools for organizing knowledge, editing, annotating, disseminating, etc. More than in other academic disciplines, in the humanities and Social Sciences, the publishing process is, all the more so in the era of digital humanities, an essential part of the research process itself.

These different needs of individuality, autonomy and cooperation in the research process so as the development of digital tools enabling a more effective realization of the scholarly primitives

---

Aeya Diarage
1 PM Tuesday

I suggest to rewrite this as: "The beginning open of the SSH is not an accidental development, but a coherent fulfillment..." I think you had to like the idea to begin with.

Elizabeth Heimann
1 PM Tuesday

I agree.

Andrea Bertini
3.30 AM Today

ok

Aeya Diarage
1:1 PM Tuesday

I took the liberty of moving this paragraph down. I would like to have a small introductory part in the first paragraph.

Andrea Bertini
3.30 AM Today

ok

Aeya Diarage
2:2 PM Tuesday

I am not aware of...", and, on the other hand, cooperation—both between the scholars themselves and...
HIRMEOS: a proof of concept for OPERAS collaboration
(Conclusion): to be continued...
ESFRI launched the 2018 Roadmap Update process on 17 January 2017 during a dedicated Info Day in Malaga. New proposals for the ESFRI Roadmap can be submitted by National ESFRI Delegations and EIROforum members until 31st August 2017, 18:00 CET.

Roadmap 2018

The Roadmap 2018 submission process will remain open until August, 31st and welcomes submissions of new proposals. Following the official presentation of the ESFRI Roadmap 2018 update process at ICR2016 in Cape Town, and the Roadmap 2018 Info Day in Malaga, ESFRI published important documents with the aim to provide a complete description of the requirements, the procedure and the methodology of selection of proposals for the Roadmap Update and thus facilitate new submissions by ESFRI Delegations and EIROforum Members.

The ESFRI ROADMAP 2018 GUIDE offers support to proposers preparing a submission and to the Projects and Landmarks involved in the update procedure. It contains the definitions, models and methods, and describes the procedures applied for this update. It represents ESFRI's best effort in road-mapping methodology and may thus serve as reference to complementary national exercises. ESFRI invites and welcomes the engagement of the research and innovation communities, as well as of the stakeholders from Europe and beyond, to identify potential new Projects and the ways to strengthen the running ones, and to maximise the return from the pan-European RI investment in terms of science, international collaboration and innovation.
THANK YOU

open access in the european research area through scholarly communication