Conference Report: OAI7 — Raising the Bar for Current Developments in Scholarly Communication, A Personal View

Paul Ayris

Co-Chair, OAI7 Organising Committee,
President of LIBER

Abstract

The purpose of this paper is to give an overview of the presentations and sessions at the OAI7 Workshop in the University of Geneva, 22–24 June 2011. The theme of the Workshop was, as always, Innovations in Scholarly Communication. The Workshop was attended by a record number of 267 delegates. Thirty-four posters were submitted for the Poster session. There were 6 tutorials, 6 plenary sessions and 6 technical sessions. The paper identifies some of the highlights of the Workshop and updates two previous review articles by Professor Raf Dekeyser and the author on the history of the OAI meetings.[1]

Key Words: OAI7; open access; scholarly communication

Attendance

Two hundred and sixty-seven delegates registered to attend the OAI7 Workshop on Innovations in Scholarly Communication from 42 countries.[2] There were 223 registered delegates in 2009 (OAI6), 215 in 2007 (OAI5), 175 in 2005 (OAI4), 143 in 2003 (OAI3); 130 in OAI2 (2002) from 20 countries; 68 registered delegates at OAI1 from 14 countries.[3] The number of attenders in 2011 is certainly a record for OAI Workshops and underlines the impact of the Workshop as an important Open Access event in Europe in the year in which it is held. The Workshop is a European meeting and is aimed primarily at informing librarians, research funders, policy makers, and publishers on

current developments in Europe in the fast-changing area of scholarly communication. The aims were set out in an introductory video, which outlined what it was hoped the Workshop would achieve.[4]

Organising an OAI Workshop

The organisation of an OAI Workshop takes 12 months to accomplish and so work on organising the academic and social programmes begins in the August of the year preceding the Workshop. This work is overseen by two committees.[5] The academic programme is overseen by the Scientific Committee. Their input is vital as they bring their expertise to shape the programme in the thematic areas, which they own and lead at the Workshop. The extensive social programme is overseen by the Local Organising Committee. No OAI Workshop could now be held without the generous financial support of our sponsors and these are chosen because they support the aims and objectives of the Workshop.[6] So successful has the format of the Geneva Workshop become, that SPARC in the USA is in active discussion with us in Geneva, as they wish to hold a mirror version of the event in North America in years when there is no meeting in Geneva.

Scientific Programme

Overview. The scientific programme for OAI7 was divided into four broad areas:[7] tutorials, plenary sessions, technical sessions, and posters.

Tutorials. The tutorials covered:

- MarcXimiL: near duplicates detection (and similarity analysis)
- CDS-Invenio
- Memento and open annotation
- Creating and managing open access journals using OJS
- Harvesters and subject-based repositories/harvesters
- Everything you always wanted to know about OA and OAI but were afraid to ask

The tutorial sessions are well-established features of OAI Workshops, lasting half a day at the start of the event. In the conference evaluation, at the end of the Workshop, respondents scored these as part of the general Workshop evaluation. One hundred and seven people said that they had attended a

tutorial and, overall, the tutorials were scored at 6.2 (out of 10), with most responses awarding a mark in the range 7–10. Respondents also made individual comments. Typical were the following:

- The Café formula for "all you wanted to know about OA" was really awesome. Having people walk around helps them to keep focus in these small group talks
- The tutorial I attended was very well run and gave the opportunity to talk to and share experiences with other attendees
- I would appreciate more.

There were suggestions for future Workshops:

- Presentation materials in a memory stick
- More interaction, less unreadable Powerpoint
- Maybe not to have too many interesting tutorial[s] concurrently :-)

Plenary sessions. The plenary sessions were devoted to the following subjects:

- Towards machine-actionable scholarly communications (Convenor: Herbert Van De Sompel)
- OAI7 aggregation (Convenor: William Nixon)
- Advocacy (Convenors: Melissa Hagemann, David Prosser)
- Open Access publishing (Convenor: Frank Scholze)
- Open Science (Convenors: Neil Jacobs, Melissa Hagemann)
- Research Data (Convenor: Raf Dekeyser)

The evaluation form asked the question: 'Overall, do you think the presentations were satisfactory?' The overall score, out of 10, was 7.9 with most respondents giving a score between 7 and 10. An identical score of 7.9 was achieved overall for answers to the question: 'Overall, do you think that the subjects covered were well balanced?'

Typical responses in the evaluation were:

- Excellent balance between librarian topics and scientific topics
- Brilliant: fun, informational, inspiring! And set in a wonderful town. Well organized with conversations all well wrapped up. Looking forward to OAI8.

- The programme has been satisfactory and shows the progress of open access. I think the participation of some scientists and editors has been beneficial because they are key players for the success of the movement
- The variety of the programme and sessions was good
- For a new-comer like me, the topics were very well chosen, the presentations were of a high level and the organization was excellent.

There were suggestions for future workshops:

- General sessions were a bit too long. I'd rather have 20-minute presentations instead of 45 minutes
- What I missed was (academic) books and the changes surrounding them
- The funder perspective would have [been] good to include, as well as perhaps senior administrators of institutions. These two stakeholder groups hold some of the keys to the transition to OA.

Technical sessions. There were six technical sessions:

- Next-generation OAI-PMH (Leader: Herbert Van De Sompel)
- OA Publishing: E-publishing research (quantitative and qualitative methods) (Discussion)
- Research data: Crowd sourcing and re-using research data (Leader: Victor Henning)
- Aggregating services (Leaders: Paul Walk and Adrian Stevenson)
- Advocacy: practical advocacy session (Leader: Monica Hammes)
- Open Science (Leader: David Flanders)

Ninty-eight respondents in the workshop evaluation said that they had attended a breakout session. The overall score was 5.1 (out of 10), with most scores ranging from 5 to 10. Written responses were varied:

- I did the 'open science' group ... it was a true 'workshop'
- Difficulty to choose 1 only topic and frustration of not being able to attend 2 at the same time...
- We need to formalise the reporting back. This can be done in 1 or 2 ways — (1) Through a formal reporting back session; (2) Asking the breakout groups to post their own reports to the conference website

- The layout of the venue for the breakout group was not particularly conducive for discussion during the session

Poster session. There were 33 posters registered for OAI7.[8] The poster prize was won by Pablo Iriarte, Isabelle De Kaenel, Jan Krause, and Nathalie Magnenat from the University of Lausanne for their poster 'Exploiting and completing institutional repositories for bibliometrics'.[9]

The average score for the posters in the workshop evaluation was 7.3, with most scores ranging from 7 to 10. Typical comments were:

- Was good to see examples of what people are doing around OA
- It was fine to get in contact with the people!
- I am not a big fan of the traditional poster sessions
- Too much information on many posters, i.e. needs too much time to find out if of interest or not — and whether a discussion or next poster is on the agenda.

Plenary Sessions — Detailed Content

Arguably, the heart of the scientific programme at an OAI Workshop is the series of plenary sessions. There were five such sessions:

Towards Machine-Actionable Scholarly Communications[10]

Sean Bechhofer gave a presentation on 'Research objects: towards exchange and reuse of digital knowledge'. Jon Deering talked about 'Publishing transcriptions as annotations of manuscript images', and finally Barend Mons gave an overview of 'Nanopublications'.

Sean Bechhofer showed that a scientific workflow can be seen as the combination of data and processes into a configurable, structured set of steps that implement semi-automated computational solutions in scientific problem-solving. He outlined his vision as:

- Research objects: aggregations to support sharing/publication
- Incorporating methods, data, people
- Research objects will allow us to conduct research in ways that are

- o efficient: cheaper to borrow than recreate
- o effective: larger scale through re-use
- o ethical: Benefiting wider communities, not just individuals

Jon Deering gave a detailed presentation on T-PEN, which is a transcription tool that allows users to transcribe from digitized images of unpublished manuscripts. T-PEN uses line segmentation to present the transcriber with one line at a time to transcribe, and preserves the association between the line and the portion of the image it transcribes. One of the ways users can output their transcriptions is as a set of OAC (Open Annotation Collaboration) annotations on the original source images. The talk discussed three ways in which OAC publication can enhance the user experience at the document repository.

Barend Mons talked about nanopublications, showing that nanopublications and triplets could replace classic journal publication. Why should data be buried in traditional publications, only for us to have to mine it? A triplet is, in its simplest form, a subject, predicate and object. Mons wants to turn these triplets into nanopublications, which can be weighted by assigning numbers to them [0 for uncertainty, 1 for certainty]. Mons would like all future publications to be annotated, by publishers or by libraries, in accord with his scheme to aid future science.[11]

OAI7 Aggregation

The second plenary session was devoted to the theme of aggregation services. The first talk was by Paul Walk on Building metadata aggregation services for resource discovery. A number of developments in the UK are focussed on the use of 'aggregation' services as an approach to providing services to support resource discovery. The talk described some of the issues raised by this approach, and attempted to describe patterns and 'anti-patterns' in the design of such services. Some of these issues were also explored in a technical workshop at the conference. For the interested, some of these aspects are described in three (loosely) related blog-posts.[12] Walk finished his talk by concluding that there is a need to:

- • use aggregation (applying a tool)
- • balance the creation of opportunity (building infrastructure)
- • solve problems (developing and providing services).

Ivo Grigorov spoke on An annual OA performance index: how vs. why. Grigorov's was a conceptual presentation, building on his knowledge of the

Global Climate Research community. Grigorov cited the Climategate episode of November 2009 as an example where news reporting led readers to the conclusion that this research community locked up their research and data. The way Grigorov and colleagues tried to tackle this was to address all administrators of EUR-OCEANS, a consortium of 26 top marine institutes in Europe. A *Cost benefit analysis of publishing marine ecosystem research output through Open Access 'self archiving'* was created in March 2010 and various measures of impact were identified. All this met with little response. Grigorov, however, did manage to identify a performance measure which caught the consortium administrators' attention. He showed that for one of the consortium members, IFREMER (L'Institut Français de Recherche pour l'Exploitation de la Mer) 80% of their peer-reviewed articles were freely accessible now back to 2006. This caught the administrators' attention and Grigorov's conclusion is that this sort of metric is the one to use in advocacy at an institutional level.

The final talk in this section was by Niahm Brennan on RIAN — pathways to Irish research, which was an analysis of the new Irish national portal at www.rian.ie, a showcase for Open Access Irish research from seven Irish universities and the Dublin Institute of Technology. Funder views of the research available in RIAN is one of the many notable developments in the portal, which is fully integrated with other Irish systems to form part of a comprehensive CWIS (Campus-Wide Information System). The lessons learned from RIAN were identified as:

- Keep it Simple
- Aggregate the data (& keep it clean)
- Distribute the services
- Devolve the admin
- It's about focused services
- Establish dependencies.

Advocacy

The advocacy plenary continued a theme which has been a favourite of OAI Workshops — that of advocacy. The session started with a paper by Monica Hammes entitled The Open Access conversation — more than just advocating for a mandate. Open Access is recognized as a sound concept for scholarly communication but is not yet embedded. Many evangelists who try to sell the idea at their universities have been disappointed by the lack of commitment in spite of enthusiasm. However, there are examples of Open Access success

stories. Furthermore, the impact of Open Access on scholarly communication and related issues provides an opportunity to introduce it as part of a much-needed campus-wide discourse on scholarly communication. Ownership, copyright, cost, new developments, OA publishing, e-research, data curation, research funding and institutional repositories can all be linked to Open Access. In this context Open Access makes more sense, can play a bigger role and eventually become a feature of local scholarly practice.

William Nixon spoke on Advocacy through embedding: integrating repositories and research management systems. Nixon spoke about the research publications repository at the University of Glasgow, called ENLIGHTEN, and the Glasgow Research system, which takes feeds of data and content from a number of university systems. In the Glasgow model, the time of data silos has certainly ended. Nixon underlined that repositories and research systems rely on three P's for success: people, processes, policies.

Nixon gave an interesting case study of a mini-REF (Research Excellence Framework) exercise which used the ENLIGHTEN repository as the platform for this activity. In the UK, the REF will actually assess UK University research outputs in 2014, and the results will dictate how millions of pounds sterling will be awarded to universities to support their research.

Heather Joseph, Executive Director of SPARC, presented on Open Access advocacy on the national — and international — level. Joseph reminded the audience that the reason for the formation of SPARC was to act as a catalyst for action. She stressed that by Open Access, SPARC sticks closely to the full definition of the concept codified in the Budapest Open Access Initiative — this is 'the whole ball of wax', as she put it. SPARC has three programme areas, which Joseph described:

- Education of stakeholders about change
- Incubation of new business/publication models
- Advocacy for new policies.

In terms of advocacy, Joseph stressed that it is important to use the language of policy makers and to remember that their concern is always 'the bottom line'. As well as work with policy makers, SPARC has also reached out to taxpayers with their 4 Principles for Taxpayer Access, work enshrined in the Alliance For Taxpayer@ccess. Joseph emphasised that for a national advocacy

campaign, there is a need for a clear 'Ask' and pointed to SPARC's work in influencing the US Consolidated Appropriations Act of 2008. As a result of that work, 2.2 million full-text articles are now available via PubMed Central. Joseph finished her talk by pointing to two new areas of development in the field of openness — the work of the Open Educational Resources community, with its emphasis on permissions; and that of the Open Data community, with a more management policy approach.

The advocacy session concluded with a panel discussion of this theme.

Open Access Publishing

Open Access publishing is a new addition to the stable of topics offered by OAI Workshops. Again, there were three papers in this session. The first was by Salvatore Mele from CERN entitled <u>Open Access publishing today: what publishers offer and what scientists want</u>. The SOAP (Study of Open Access Publishing) project investigated the supply of and demand for Open Access journals, as well as the experiences of scholars with this new publication paradigm. On the supply side, SOAP compiled data on existing Open Access journals, finding they publish around 8% of the yearly total of scholarly articles, in many cases with extremely high-quality content, but with a somewhat confused licensing landscape. On the demand side, SOAP ran a large-scale survey of researchers with 40,000 answers across disciplines and around the world. The results show an overwhelming support for the idea of Open Access, while highlighting funding and (perceived) lack of quality as the main barriers to publishing in Open Access journals. The talk captured an image of the SOAP project in four numbers:

- 89% of scholars are positive about Open Access journals
- 8% of articles each year are published in Open Access (120,000/1.5 million)
- 39% of scholars quote funding as a barrier to Open Access
- 31% quote 'prestige' and 'quality' as a barrier.

Christoph Bruch and Barbara Kalumenos spoke on <u>The PEER project: observing the impact of green Open Access</u>. PEER (Publishing and the Ecology of European Research), supported by the EC eContent+ Programme, investigates the effects of the large-scale, systematic depositing of authors'

final peer-reviewed manuscripts on reader access, author visibility and journal viability, as well as on the broader ecology of European research. The project is a collaboration between publishers, repositories and researchers and will last from 2008 to May 2012. The first part of the talk covered the technical challenges which were overcome by the project and the participating publishing houses in setting up the PEER observatory. The second part of the talk focused on presenting the lessons learned from setting up the complex project infrastructure (PEER observatory) and the preliminary findings of the three PEER research projects. These addressed behavioural research, usage research and economic research.

An electronic survey of researchers was conducted in Summer 2009. The headline findings were these:

- More than two-thirds of respondents indicated that they understood Open Access to mean free electronic access to the fulltext of articles
- 46% of respondents preferred placing their material in subject-based repositories
- Only 37% of responding authors thought that they had access to a suitable subject-based repository
- In terms of barriers, copyright concerns are the main problem.

The final talk in this session was by Mark Patterson on Re-engineering the functions of journals. The use of online media is allowing the processes of scholarly communication to be re-invented and re-engineered. Open Access to research information is a key first step because it removes all barriers for access to and re-use of the information, thereby maximizing its impact. At PLoS (Public Library of Science), the initial focus has therefore been to establish a successful and sustainable Open Access publishing operation. Having achieved this goal, PLoS is now exploring new ways to enhance scholarly communication through online publications that publish new findings more rapidly, and new products that facilitate the evaluation and organisation of content after publication.

Patterson identified two functions of journals which need to be re-engineered: dissemination and organisation of content. In terms of the organisation of content, PLoS One's key innovation is the editorial process. PLoS One is the largest of all peer-reviewed journals with 50,000 authors and 1,500 academic editors. Patterson talked about the development of PLoS Hubs, which will

aggregate Open Access content wherever it is published, adding value to the content by connecting with data and building communities around that content. Patterson suggested that the new models of scholarly communication will:

- enable more rapid communication
- accelerate moves to Open Access
- lead to a consolidation in fewer 'journals'
- provide article-level research assessment
- innovate with post-publication content enhancement.

Open Science

Open Science was again a new topic for the OAI Workshop. The first paper was The rise of citizen cyberscience and its impact on professional research by François Grey, who argued that the rapid growth in the last decade of direct public participation in science via the Web — citizen cyberscience — has important implications for the Open Science agenda. Leading researchers now routinely use citizen cyberscience to tackle large-scale scientific computing and data analysis challenges, in areas as diverse as climate science, epidemiology and molecular biology. In aggregate, millions of volunteers are contributing to such projects by donating spare time on their computers, participating directly in data analysis via the Web, or even collecting data from the field using smart phones.

Grey identified seven myths of cyberscience:

- It doesn't produce real science
- It won't work for my science
- Nobody will be interested in my science
- You can't trust results from ordinary people
- It is energetically hugely wasteful
- It doesn't really engage people in science
- One day we will run out of volunteers.

In exchange for their effort, volunteers typically want more openness and better communication about what scientists are doing. Some even want to have an influence on the direction of the research. This raises new issues for how open and participative the scientific process can and should be.

The second paper in this session was by Cameron Neylon, entitled Technical, cultural and legal infrastructure to support effective open science communication. The technical challenges in sharing data and other artefacts of scientific research beyond the traditional paper remain formidable. There are clues as to how to proceed, and big positive steps are being taken. These will require significant investment in trustworthy, workable, and flexible infrastructure. Legal tools and systems are also required to ensure that these outputs are freely useable and re-usable. Creative Commons and the work of the Open Knowledge Foundation are part of this — to ensure that people are confident about their ability to re-use material. For Neylon, the real problem is that individuals are not involved in the process, and so a new cultural infrastructure is needed. Mandates, research evaluation and the like can change practices, as long as they are embedded in the workflow. However, there is a need to build real foundations and pillars around this scaffolding. Re-use and sharing should be the norm of open research. What is *new*, however, is the need to be open for input from the outside world. This is a challenge, which is being grappled with by society at large. Thinking of these different components as part of the scientific research infrastructure will be crucial in building a viable platform for modern research communication and exploitation.

The final paper in this session was by Victor Henning: <u>Mendeley as a component in the open science infrastructure</u>. Henning was one of three PhD students, collaborating together via skype, who found themselves overwhelmed by the number of pdf's they had to manage. The company was started in 2008 and now employs 40 people in London and New York. In essence, Mendeley extracts research data and aggregates it in the cloud. Ninety million documents have been uploaded to date (42 million unique documents) by one million users. Users can read documents, highlight portions and set up shared groups to work collaboratively. All the data are anonymously aggregated into a central database. Newsfeeds are possible, as is a recommendation system. Open URL's link through to commercially published versions of texts. Readership statistics are available at document level. Henning described the JISC DURA project (Direct User Repository Access), whose aim is to bring repository deposit into the workflow. Through a link to the Symplectic system, Mendeley can deposit relevant publications in institutional repositories. Henning also described API development work, the aim of which is to make information in Mendeley available for re-use.

Finally, the speaker described an API competition which Mendeley had recently organised to drive forward the concept of Open Science.

Research Data

The final plenary session in OAI7 concerned research data, again a topic which was given a heightened prominence in the Workshop compared to previous years. Two papers were presented in Geneva. The first was by Anja Jentzsch entitled <u>Linked Data — towards a Web of data</u>. Jentzsch outlined the vision of Linked Data as the creation of the Web as a single, global dataspace. She highlighted a set of principles from Berners-Lee to outline what, in practice, needs to be done for publishing structured data on the Web:

- Use URI's as names for things
- Use HTTP URI's so that people can look up those names
- When someone looks up a URI, provide useful RDF information
- Include RDF statements that link to other URIs so that they can discover related things.

The speaker then went on to characterise the properties of the Web of Linked Data:

- Global, distributed dataspace built on a simple set of standards
  - RDF, URI's, HTTP
- Entities are connected by links
  - creating a global data graph that spans data sources and
  - enables the discovery of new data sources
- Provides for data-coexistence
  - Everyone can publish data to the Web of Linked Data
  - Everyone can express their personal view on things
  - Everybody can use the vocabularies/schema that they like.

The goal of the W3C Semantic Web Education and Outreach group's Linking Open Data community project is to extend the Web with a data commons by publishing various open datasets as RDF on the Web and by setting RDF links between data items from different data sources. In October 2007, datasets consisted of over two billion RDF triples, which were interlinked by over two million RDF links. By September 2010 this had grown to 25 bil-

lion RDF triples, interlinked by around 395 million RDF links. There is also an interactive visualization of the linked data sets to browse through the cloud.[13]

An RDF (Resource Description Framework) file should parse down to a list of triples. A triple consists of a *subject*, a *predicate*, and an *object*. But what do these actually mean? The *subject* is, well, the subject. It identifies what object the triple is describing. The *predicate* defines the piece of data in the object we are giving a value to. The *object* is the actual value.[14]

The speaker described the DBpedia project, which is a joint project with the following goals:

- extracting structured information from Wikipedia
- publish this information under an open licence on the Web
- setting links to other data sources.

In the library community, Jentzsch described a sizeable uptake in Linked Data implementations:

- Library of Congress (subject headings)
- German National Library (PND dataset and subject headings)
- Swedish National Library (Libris — catalog)
- Hungarian National Library (OPAC and Digital Library)
- The Europeana project has just released data for about 4 million artifacts
- W3C Library Linked Data Incubator Group
- OKFN Working Group on Bibliographic Data, with goals to:
  o Integrate Library Catalogs on global scale
  o Interconnect resources between repositories (by topic, by location, by historical period, by ...).

How is Linked Data published? The author gave a concise answer:[15]

- Make data available as RDF via HTTP
- Set RDF links pointing at other data sources
- Make your data self-descriptive
- Re-use common vocabularies.

Jentzsch concluded by saying:

- The Web of Data is growing fast
- Linked Data provides a standardized data access interface
- Linked Data allows for the development of a variety of tools to integrate, enhance and view the data
- Web search is evolving into query answering
  - Search engines will increasingly rely on structured data from the Web
- Next step: Linked Data within enterprises
  - Alternative to data warehouses and EAI (Enterprise Application Integration) middleware
  - Advantages: schema-less data model, pay-as-you-go data integration.

The final paper in the Workshop was by Peter Wittenburg (who presented the paper in Geneva) and Krister Linden, entitled <u>Riding the wave. After the EU HLEG Report: vision (and reality) about accessing research data</u>. The purpose of the presentation was to take the debate on accessing research data forward in the wake of the EU High Level Expert Group's Report *Riding the Wave*.[16] *Riding the Wave* pointed out what has increasingly been felt for years: there is a need to take urgent measures with respect to scientific data, if the disaster of not being able to access it anymore is to be averted. However, the Report also emphasizes the opportunities of the information hidden in the increasing amount of data which are available. There is a need to increase the awareness of all stakeholders such as researchers, data scientists, research organizations and even the public about the huge relevance of such data, to extract the knowledge that is needed in the coming decades and beyond. A concerted action is required that will amount to a Collaborative Data Infrastructure (CDI) consisting of three layers: the researchers as data generators and users, research infrastructures offering community-specific data services and a data-oriented e-infrastructure offering common data services. Since the creator-consumer relationships become more and more anonymous, there is a need for new ways to establish trust relationships. Because knowledge about the stored data objects is distributed vertically, the responsibility for data curation is shared — yet there are not yet proper mechanisms in place to synchronize decisions. The Report makes a number of suggestions to identify a vision for 2,030 for data management and access.

What is the vision for 2,030? Wittenburg outlined it as follows:

- All stakeholders, from scientists to national authorities to the general public, are aware of the critical importance of preserving and sharing reliable data produced during the scientific process.
- Researchers and practitioners from any discipline are able to find, access and process the data they need. They can be confident in their ability to use and understand data and they can evaluate the degree to which the data can be trusted.
- Producers of data benefit from opening it to broad access and prefer to deposit their data with confidence in reliable repositories. A framework of repositories work to international standards, to ensure they are trustworthy.
- Public funding rises, because funding bodies have confidence that their investments in research are paying back extra dividends to society, through increased use and re-use of data.
- The innovative power of industry and enterprise is harnessed by clear and efficient arrangements for exchange of data.
- The public has access to and can make creative use of the huge amount of data available; it can also contribute to the data store and enrich it.
- Policy makers can make decisions based on solid evidence, and can monitor the impacts of these decisions.
- Global governance promotes international trust and interoperability.
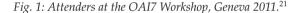
Social Programme and Social Networking

The Workshop had strong a social programme in The Globe at Cern Laboratories. On 22 June, the Workshop was treated to a buffet supper there, a formal address by Emmanuel Tsesmelis, a member of the Directorate Office, a Physics demonstration entitled *Fun with Physics* (Drôle de Physique) by Sebastien Pelletier, a CERN user coming from the Fachhochschule Wiener Neustadt, and optional tours of the ATLAS Visitor Centre.[17] The evening also included the now-traditional Drinks Sharing, where participants bring a traditional drink from their home country to share with everyone else. On 23 June, there was a rooftop aperitif in the University, generously sponsored by Microsoft Research.

The Workshop has also generated strong coverage in social networks. Many informal photographs of the event can be found, contributed by partici-pants, on the Workshop website.[18] The Workshop also had a formal Twitter account,[19] which was regularly used during the event to initiate discussions and comment on presentations in real time, and also an account on Facebook.[20] Participants regularly commented on the content of the presentations in real time by sending tweets with the #OAI7 hash tag.

Conclusions

The OAI7 Workshop at the University of Geneva on 22–24 June 2011 was important in a number of respects. There was a record attendance at this Workshop (267 registered attenders, Figure 1), which suggests that the Workshop is becoming embedded in the European information landscape. From the formal evaluation forms returned by attenders, it is clear that they both enjoyed the Workshop and found it successful. The evaluation form asked the question: 'Overall, do you think the presentations were satisfac-

Fig. 1: Attenders at the OAI7 Workshop, Geneva 2011.[21]

tory?' The overall score, out of 10, was 7.9 with most respondents giving a score between 7 and 10. An identical score of 7.9 was achieved overall for answers to the question: 'Overall, do you think that the subjects covered were well balanced?'

On a personal level, three things struck me in OAI7. First, it was clear that Open Access is becoming embedded in institutional research management systems. Several papers showed how Open Access had become part of their CWIS (Campus-Wide Information System). Second, it was clear that commercial offerings are playing an increasing role in providing Open Access platforms, infrastructure and services. A number of papers showed that this was the case and the success of these developments shows that such offerings are being taken up by the scholarly community. Third, the topics of research data and open science show how new modes of research and evolving. Data-driven science, with its emphasis on empirical evidence, collaboration and re-use, is empowering. This is a topic to which future OAI Workshops must return.

Perhaps the final word should be directed towards the members of both the Scientific Committee and the Local Committee. The Scientific Committee organises the academic sessions of the Workshop and the Local Committee makes all the local arrangements for the three days of the Workshop. It is tribute to the work of the Local Committee that they achieved the highest score in the formal evaluation returns, scoring an impressive average of 8.8 out of 10.

The Organising and Local Committees unanimously agreed to carry on the work by organising an OAI8 Workshop in two years' time. The OAI Workshops have become embedded into the European information landscape. They are community-owned and the evaluation feedback shows that they are meeting the community's needs. Perhaps this is the most notable feature of the workshops. They succeed because the community feels part of them and wants them to succeed. It is a good omen for the future success of OAI8 in 2013.

Notes

[1] See R. Dekeyser, 'The LIBER Workshops on the 'Open Access Initiative' at Cern, Geneva' in *LIBER Quarterly*, vol. 16 (2006), no. 1; and P. Ayris, 'Embedding Open Access into the European Landscape — the contribution of LIBER' in *LIBER*

Quarterly, vol. 17 (2007), no. 2. *LIBER Quarterly* is available in Open Access at http://liber.library.uu.nl/.

[2] See http://indico.cern.ch/confRegistrantsDisplay.py/list?order=down&sessionShow NoValue=1&sortBy=Country&confId=103325#results.

[3] For these comparisons, see http://indico.cern.ch/internalPage. py?pageId=5&confId=48321; Dekeyser (2006); and http://www.dlib.org/dlib/ november02/vandesompel/11vandesompel.html.

[4] See http://www.ucl.ac.uk/stream/media/swatch?v=a53f790bd14e.

[5] Listed at http://indico.cern.ch/internalPage.py?pageId=6&confId=103325.

[6] See http://indico.cern.ch/internalPage.py?pageId=7&confId=103325.

[7] The splash page for the OAI7 Programme can be found at http://indico.cern.ch/ conferenceTimeTable.py?confId=103325#20110622. The slides and audio recordings of the plenary sessions are linked from the relevant Programme page.

[8] See http://indico.cern.ch/sessionDisplay.py?sessionId=10&tab=contribs&confId= 103325.

[9] See http://indico.cern.ch/contributionDisplay.py?contribId=70&sessionId=10&co nfId=103325.

[10] In transcribing the titles of the talks, I have used the powerpoint slides presented on the day and the audio of the presentation on the Cern website. These are linked to the Programme at http://indico.cern.ch/conferenceTimeTable. py?confId=103325#all.

[11] I am grateful to Laika's MedLibBlog for providing additional context for me to describe Mons's paper at Geneva; see http://laikaspoetnik.wordpress. com/2010/06/23/will-nano-publications-triplets-replace-the-classic-journal- articles/.

[12] See 'Institutions and the Web done better' at http://blog.paulwalk. net/2010/09/21/institutions-and-the-web-done-better/; 'Aggregation and the Resource Discovery Taskforce vision' at http://www.ukoln.ac.uk/jisc-ie/ blog/2010/08/19/aggregation-and-the-resource-discovery-%20taskforce-vision; 'An infrastructure service anti-pattern' at http://blog.paulwalk.net/2009/12/07/ an-infrastructure-service-anti-pattern/.

[13] See http://en.wikipedia.org/wiki/Linked_Data.

[14] Taken from http://www.robertprice.co.uk/robblog/archive/2004/10/What_Is_ An_RDF_Triple_.shtml.

[15] See Tom Heath and Christian Bizer: Linked Data: Evolving the Web into a Global Data Space at http://linkeddatabook.com/.

[16] Riding the Wave. How Europe can gain from the rising tide of scientific data. Final Report of the High-level Expert Group on Scientific Data, October 2010, http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf.

[17] See http://atlas.ch/atlas_tour/index.html.

[18] See http://www.ucl.ac.uk/stream/media/swatch?v=a53f790bd14e.

[19] See http://twitter.com/#!/oai7.

[20] See http://www.facebook.com/group.php?gid=26157314629&v=wall.

[21] This is the official Workshop photograph, taken from the CERN website, and provided separately for publication by the Local Organising Committee.