Report from the Knowledge Exchange event: Pathways to open scholarship

Held in Helsinki, December 2015

Author
Nicky Ferguson

February 2016
Foreword

This report has been collated from the contributions of many. Thanks are due to those who contributed their time and expertise to the event: those who presented, chaired, facilitated and provoked during the plenary and breakout sessions, those who made contributions from the floor and during the break-out sessions and the conference organisers who ensured the event was run professionally and smoothly.

In particular to all those who made their contributions available to the author and the tweeters and collaters who made it possible to bring together a diverse and varied collection of opinions and reports.

Any misreporting and mistakes are the responsibility of the author.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to Knowledge Exchange and background to the conference</td>
<td>5</td>
</tr>
<tr>
<td>The first ten years of Knowledge Exchange</td>
<td>6</td>
</tr>
<tr>
<td>A wish list for the future: executive summary</td>
<td>8</td>
</tr>
<tr>
<td>Presentations and pressing issues</td>
<td>9</td>
</tr>
<tr>
<td>Promotion and diffusion</td>
<td>9</td>
</tr>
<tr>
<td>Risks, limitations and disincentives</td>
<td>9</td>
</tr>
<tr>
<td>Relating the success of the researcher to the success of the institution</td>
<td>10</td>
</tr>
<tr>
<td>Technology and incentives for researchers and institutions</td>
<td>10</td>
</tr>
<tr>
<td>The effect of policy changes on publication practice</td>
<td>10</td>
</tr>
<tr>
<td>Information – free or expensive?</td>
<td>12</td>
</tr>
<tr>
<td>History or anonymity?</td>
<td>12</td>
</tr>
<tr>
<td>The waist of the hourglass</td>
<td>13</td>
</tr>
<tr>
<td>Project demonstrations, prototypes and presentations</td>
<td>14</td>
</tr>
<tr>
<td>Using VIVO to pull together personal profiles</td>
<td>14</td>
</tr>
<tr>
<td>Certified repositories</td>
<td>14</td>
</tr>
<tr>
<td>Open data on atmosphere and ecosystem</td>
<td>14</td>
</tr>
<tr>
<td>Open Library for Humanities</td>
<td>15</td>
</tr>
<tr>
<td>Open Science, Research Ideas and Outcomes</td>
<td>15</td>
</tr>
<tr>
<td>What comes next?</td>
<td>17</td>
</tr>
<tr>
<td>Beyond Helsinki</td>
<td>18</td>
</tr>
<tr>
<td>Appendix – The full wish list</td>
<td>19</td>
</tr>
</tbody>
</table>
Introduction to Knowledge Exchange and background to the conference

Knowledge Exchange is a collaboration between five national organisations, DFG – the German Research Foundation, Jisc – the UK’s champion for digital technologies in education and research, DEff – Denmark’s Electronic Research Library, SURF – the ICT organisation for Dutch higher education and research and CSC – the IT Centre for Science in Finland.

These five key national bodies within Europe are working together to support the use and development of ICT infrastructure for higher education and research. Although the organisations are very different in the size and scope of their work, each has a national responsibility and influence on national policy, operates at the cutting edge level of IT development and can mobilise resources that can make a difference. Each organisation is active in Open Scholarship and supports open access to research and learning. Knowledge Exchange (KE) activities have had positive outcomes allowing partner organisations and their national policy makers to be better informed, share expertise and resources and push forward the necessary technologies to allow us to realise our shared agendas in developing and improving education and research.

Pathways to Open Scholarship was both a milestone and an agenda setting event for Knowledge Exchange. The conference briefly looked back at the first ten years of Knowledge Exchange and its achievements. But the main focus was looking forward. The goal of the event was to explore the value, the consequences and the risks of Open Scholarship, by presenting and comparing views to identify issues that require attention.

The outputs would be:

- a report showcasing the current opportunities in and progress towards Open Scholarship
- “a wish-list of what we would like to be achieved”
- an opportunity to follow these up by asking the community for recommendations for actions to take Open Scholarship forward in Europe

This report attempts to reflect the meeting. It gives a brief account of first ten years of Knowledge Exchange but in the main looks at the present and the future:

- Present opportunities, current work and resources of which we may not be aware, obstacles and frustrations of which we are sometimes painfully aware
- Future challenges, our wishes for the future and a request for you, the reader, to suggest actions which may make those wishes more attainable

Footnotes
1 Pathways to Open Scholarship
(knowledge-exchange.info/news/articles/15-10-2015)
Knowledge Exchange was founded in 2005 as a partnership to foster cooperation and the exchange of knowledge between four national organisations responsible for IT in Higher Education and Research.

It was inspired by the 2003 Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities which stated that “Our mission of disseminating knowledge is only half complete if the information is not made widely and readily available to society”. Knowledge Exchange’s initial aim was to “To make a layer of scholarly and scientific content openly available on the Internet” and its first action was to petition the EC to use its influence to ensure that publicly funded research results should be openly available to the scholarly community. In addition to sharing operational, tactical, and strategic information between decision makers within the partner organisations, Knowledge Exchange’s main activity has been to bring together experts to exchange, network, prioritise and act in various fields and on various issues. Knowledge Exchange partners have been informed and inspired by these expert networks and activities and Knowledge Exchange has gone on to promote the work of these experts through workshops, publications and responses to high-level consultations. In particular, the Knowledge Exchange Research Data expert group has built strong links with Science Europe, RDA, CNI, and EUDAT, while the Knowledge Exchange Open Access expert group is similarly connected with SPARC Europe, the European Research Council and OpenAIRE.

A sample of Knowledge Exchange workshop titles would include:

- Institutional repositories (2007)
- Persistent Identifiers (2009)
- Research Data Management (2011)
- Making Data Count (2012)
- The Price of Knowledge (2013)
- Open Access Sustainable Business Models (2013)
- Research Software Sustainability (2015)
- Open Access Policy Dependency (2015)
- Research Data Management Training and Skills (2016)
While reports on issues and meetings have included:

- Interoperability between Repositories and CRIS
- Website OA success stories
- Multi-national journal licensing pilot
- Response to EC RD ambitions: a Surfboard for Riding the Wave
- Collaborative Research: VREs, Tools, Data
- Persistent Identifier project (URN-NBN, Handle, DOI at one table)
- Guidelines for interoperable Usage Statistics for OpenAIRE
- Discussion paper Open Knowledge (eco-system approach)
- Authority Files (controlled vocabularies)
- Author Identifier Summits (ISNI, ORCID at one table)
- Value, Cost, Pricing, Sharing, Funding of Research Data (Infrastructures)
- Sustainable Business models of OA services
- Research Software Sustainability

Knowledge Exchange would like to recognise with gratitude and great respect the contribution of all the experts involved with its activities. You have led the way. Thank you.

A new collaboration agreement has now been signed for 2016 and beyond which updates and formulates the Knowledge Exchange partners’ long term strategic ambitions. While continuing the important work on Open Access and Research Data, Knowledge Exchange also aims to be flexible in its response to emerging issues, urgent actions and partners’ changing needs. Formal agreements and hard work from staff and partners have underpinned the success of Knowledge Exchange, but a crucial factor has been the trust between partners which has formed the basis for creative, open and informal exchange. Knowledge Exchange activities will continue to be led and guided by experts in the partner organisations and in collaboration with others outside the partnership.

Knowledge Exchange’s new vision is “to enable open scholarship by supporting an information infrastructure on an international level”. We hope to draw inspiration from the Pathways to Open Scholarship event.

Footnotes

2 For the Netherlands, Denmark, UK and Germany, later joined by Finland
A wish list for the future: executive summary

For the tenth anniversary event, the Knowledge Exchange partners wanted to explore a new vision for the coming years, a vision which focuses on Open Scholarship and supporting the international information infrastructure that will be necessary to achieve it.

Although ambitious, Knowledge Exchange is convinced that it will make for better scholarship. Two days of discussions in Helsinki gave us a long wish list of things that we believe will move scholarship in the right direction.

Some of the priorities on our list are:

1. A new independent Open Scholarship institute to drive innovation

2. More opportunities and more encouragement for researchers to make their publications open

3. Case studies / portraits of individuals who are committed to Open Scholarship and also successful and very highly regarded in their own field

4. Case studies / portraits of institutions where Open Scholarship is part of “the plumbing”

5. A ‘selection box’ for researchers, showcasing the Open Scholarship tools, resources and infrastructure that already exist to support them

6. A ‘selection box’ for institutions, funders and other stakeholders, showcasing the tools, resources and infrastructure that they can utilise

7. Reliable metrics (or maybe a calculator) to give a realistic assessment of benefits likely from an open approach

8. Accessible, usable and realistic guidelines on handling sensitive data

9. New systems to recognise the work and merit of individual researchers, and of institutions

10. Completely transparent journal costs

11. New technology to support discovery and automate processes

Read the full wish list together with links to a number of other exciting initiatives in the Appendix.
Promotion and diffusion
Sascha Friesike, from the Alexander von Humboldt Institute for Internet and Society and author of *Opening Science* was asked to give a provocative introduction. He succeeded in provoking disagreement by comparing the promotion of Open Scholarship to the marketing of disposable diapers/nappies. The first (failed) effort to market disposable nappies focussed on waste disposal (hence the name Chucks), the successful later effort, Pampers, focussed instead on the comfort of the baby, though the product was essentially the same. A more comfortable baby was a more universal desirable attraction than a more efficient waste disposal system. Leaving aside for now the contentious issue of whether the advance of knowledge and exchange of scholarly information is comparable to selling nappies, is there anything to learn for us here? Friesike contends that this example, used by Everett Rogers in his *Diffusion of Innovation* theory, gives us some discussion points for considering how we might promote wider adoption of Open Scholarship (with Friesike’s answers in italics):

1. **Relative advantage, personal advantage:** does Open Scholarship advance the reputation of researchers? *It may do, objectively, but many researchers think not.*

2. **Observability:** can we observe or share such an advantage? *Not yet.*

3. **Complexity:** is it easy to communicate and understand? *No, the open access movement has been characterised by a proliferation of licences, colours and versions, partisan hostility amounting to religiosity in the debates and a willingness to harshly criticise anyone who isn’t “pure”. Open access itself has proved hard to define, and complex for institutions to understand and to adopt; people wanting a simple solution tend to go to publishers.*

4. **Compatibility:** is it compatible with your old practices and investments? *In some senses yes, as publishing is still putting your ideas in front of people, particularly colleagues. But the model of Open Scholarship challenges well established beliefs and values – the “leading” journals and publishers in your field, for example.*

5. **Trialability:** can we try it out? Do we get a free trial? *Yes you can have a go, but early career researchers may consider that the time and effort they have invested in their early work is not worth risking, that the overhead of learning about Open Scholarship is too great and that they might get “jumped on” (see point 3).*

Friesike concludes that we often focus so much on the detail and mechanics of Open Scholarship that we are not improving or demonstrating the incentives that already exist, the relative simplicity and multiple routes to open publication, its rapidly improving status and the possibilities of increased readership and enhanced reputation that it may offer. [Slides from the presentation](http://bit.ly/1UrpPak).

**Risks, limitations and disincentives**
Elke Greifeneder, from the Berlin School of Library and Information Science, pointed out that it was widely known and widely resented that researchers seem to pay (or provide free labour) for everything in the publication process: producing, reviewing and reading content together with producing, formatting and validating data and then reproducing and defending it. And yet … “I used to publish everything I can in Open Access, but I stopped”. Why? - because I cannot control the reviewing and assessment processes where my (unknown or at least not selected by me) reviewers may attach more

---

**Footnotes**
importance to where I am publishing than whether or not it is open. The time I spend on ensuring the quality of my data and making it reusable is not rewarded or even recognised. Greilfeneder concludes that the realities of Open Scholarship do not yet reflect the rhetoric.

**Relating the success of the researcher to the success of the institution**

Alenka Prinčič from The Netherlands Delft University looked back in history to attempt to learn lessons about how to produce an environment which allows researchers to produce excellent research. She postulates that the environment should allow researchers to focus on their research and not be bothered with a lot of extraneous responsibilities. So, in this phase of technology change and transition, this puts a very large responsibility on support staff. We must also recognise how the different stages of researchers’ careers affect their activities. She suggests that instead of judging researchers on the number of publications they produce, institutions should instead set a maximum number of publications during a project (and take time to get high quality peer and user review of that small number of publications). This would be a radical shift away from the demand for quantity of publications towards an emphasis on their quality.

**Technology and incentives for researchers and institutions**

Oscar Corcho from the Universidad Politécnica, Madrid, believes that technology will quickly provide us with tools for better collaboration, better data manipulation, reproduction and sharing. But he agrees with earlier speakers that we must provide clearer and better understood incentives for researchers and their institutions. For example, Data Management Plans (DMPs) are widely discussed but still not universally applied and rarely followed carefully and comprehensively. Implementing them can take a lot of time and effort and, understandably, they are not popular amongst researchers who don’t see many if any benefits to themselves. Corcho suggests that institutions might include in their hiring criteria for staff their record in data management and their participation in Open Scholarship (and formally make these criteria at least as important as which journal they have published in).

From the panel’s “open chair” reserved for a member of the audience, Björn Brembs, responded to the previous speakers that researchers are no less or more motivated by self-interest than anyone else. He contends that mechanisms for Open Scholarship could and should be provided for researchers by their institutions, in the same way as the telephone line, electricity, photocopier etc. People recycle more if the boxes are picked up from the doorstep rather than making them drive ten miles to the tip. So the best way to get a service used is to provide it for the researchers in an easy to use fashion. He concludes that if the institutions made the mechanisms for Open Scholarship easily available to their staff by default then this would remove a significant barrier.

**The effect of policy changes on publication practice**

Tony Hey from the Science and Technology Facilities Council (STFC), UK was more optimistic about the progress towards Open Scholarship and particularly the cumulative effect of mandates and policy shifts on cultural change. Hey uses the particle physics community as an example of early adopters who made a change, put papers up on arXiv.org first and did not wait for approval from institutions or publishers. Hey talks about “designing for adoption”, using the Sketchplanation sketches as a visual aid; he then moves on to “understanding the chasm” based on the work of Geoffrey A Moore. We are at the edge of the chasm, the early adopters are on one side, can we leap across it so that it becomes the norm?

Hey says that the US white house memorandum of February 2013, requiring the major Federal Funding agencies “to develop a plan to support increased public access to the results of research funded by the Federal Government” was a huge change and will probably be widely influential outside the U.S. After the National Instutes of Health (NIH) in the US implemented policy changes with a change in the law requiring open access compliance in 2007, publication in PubMed Central (PMC) went up from 16% to 72%. With their subsequent policy of holding back “processing of non-competing continuation awards if publications arising from grant awards are not in compliance with the Public Access Policy” the compliance rate increased by 0.5% per month and reached 86% by November 2014. Hey is similarly optimistic about the role of institutions, pointing out that the University of California (UC) requires all 8000 faculty to deposit full text copies of their research papers in the UC eScholarship repository unless they specifically choose to opt-out.
This institution is the largest public research university in the world and its faculty members receive roughly 8% of all research funding in the U.S. UC produces 40,000 publications per annum corresponding to about 2 – 3% of all peer-reviewed articles in the world each year. As for the creation of metadata and data management plans, things are not perhaps moving so quickly, but the biggest change Hey puts down to a funding agency mandate: the National Science Foundation (NSF) in the U.S. insisting on a Data Management Plan for all proposals. Hey notes that metadata creation should be automated as much as possible and that researchers need to work with subject specialist librarians when creating metadata by hand.

Several contributions from the floor followed up with concerns about data, its quality and reliability. There were claims that the old way of scholarship is dying and classical scholarship is broken (several speakers from the floor claimed that 50% of biomedical research results cannot be reproduced). What will replace it? From the platform, speakers responded with questions of their own:

- Will there be too much data for its quality to be assessed as useful?
- If 50% of data is unreliable surely the only practical way to respond is to make it public and allow the wisdom of the crowd to expose its flaws? It cannot be feasible to “fix” all this data before making it public
- Why, in spite of funding from Europe, have we not been able to produce a widely used repository competitive with commercial offerings?

From the floor, Cliff Lynch, from CNI, pulled the threads of the discussion together by noting that touting of papers, reliance on journal prestige rather than article content etc. are all symptoms of a lack of sensible evaluation of the contribution of scholars. arXiv\(^1\) has an important lesson, the publishers had to follow the consensus of the community because it was united and strong. Can we come up with lists of problems in a discipline over which there is a consensus (on the importance of the problem not necessarily its solution)? If not then how do we come up with a solution to dissemination of scholarly information?

A series of five minute presentations followed looking at specific parts of the Open Scholarship landscape. Like the preceding talks, they each linked to one of the four central themes of the event:

- Theme 1 – Benefits, risks and limitations of Open Scholarship
- Theme 2 – Success as a researcher
- Theme 3 – Technology
- Theme 4 – Publishing and publication services

Footnotes
\(^6\) arXiv (http://arxiv.org)
\(^7\) Designing for adoption (http://bit.ly/1UFjwj3)
\(^8\) Understanding the chasm (http://bit.ly/1o60Lar)
\(^9\) Jono Hey - sketchplanations.com (http://bit.ly/1UFjwj3)
\(^10\) Jono Hey - sketchplanations.com (http://bit.ly/1o60Lar)
\(^11\) arXiv (http://arxiv.org)
Information – free or expensive?
Neil Jacobs from Jisc took as his starting point the aphorism that “Information wants to be free … information wants to be expensive”. This is attributed to Stewart Brand and first recorded in 1984:

On the one hand information wants to be expensive, because it’s so valuable. The right information in the right place just changes your life. On the other hand, information wants to be free, because the cost of getting it out is getting lower and lower all the time. So you have these two fighting against each other.

Jacobs used several examples to illustrate the huge value to society, governments, individuals and corporations of scholarly research. On the one hand commercial companies operating in the areas of e.g. health (big pharma), environment (mining, energy), etc, are willing to pay for access to valuable research; on the other hand we can point to arXiv where over 1m preprints are freely available at an annual running cost of far less than a dollar per article. As for research data, which Jacobs describes as “incomprehensibly massive”, the benefits to science and industry of open data are obvious but the costs of making data “intelligently open” beyond those immediately working on it are potentially huge, and there are few incentives for researchers to do so in some disciplines. He points out that:

we do have national and international data centres such as EMBL/EBI, where the research community itself curates its data. Is this a model for research publications? If not, why not?

A look at data on individual behaviour, citations, activity and usage illustrates the conflicts and tensions in this area. Debates rage about open access policies, while fewer people seem to worry about the terms of use and privacy policies of services such as ResearchGate and Academia.edu. For them, information is valuable, expensive and closed. While still very small in comparison, PLoS ONE, with the largest journal ever, is also very successful and its staff understand the importance of transaction data – for them, this information should be free, with the research community in control of how this works. Calls for open citation data and open metrics are repeated later in this report.

Jacobs contrasts two sometimes conflicting views of Open Scholarship:

- The purely economic model in which the internet enables the long tail to be freely available while the most commercially valuable aspects of research will be marketable commodities, and the giants, e.g. ResearchGate and Elsevier, will fight to control them
- The model of enlightened and ethical progress where difficult issues such as open transaction data, open research data, reproducible research at scale, research assessment etc. need solutions owned by academics and their community

He notes examples of where these two “world views collide” – amongst which are issues already raised by several contributors:

- Researcher incentives and career progression
- Investing scarce resources into opening up research data for scrutiny and reuse

and asks can these two world views coexist? And if not which will prevail amongst academics?

History or anonymity?
Heidi Laine from University of Helsinki raised other questions for Open Scholarship to address. What answers does Open Scholarship provide to ethical problems such as author misbehaviour and plagiarism? Is it unreasonable to expect this since the issues of review, peer review, authenticity and quality judgements are essentially separate from access and openness? She points out that historians have always reused data and asks Can I declare with certainty that no harm will come to my interviewees? She answers: No, in some cases, interviewees can easily be identified even if I anonymise. If we completely hide the identity of respondents then we lose valuable historical data. She echoes previous speakers in contrasting the care and concern taken with data in academic institutions with the wide and some would say intrusive terms, conditions and usage policies of everyday commercial web services. While academia destroys and anonymises data, commercial companies collect huge amounts of data from reluctant consumers who see no other option but to agree to their terms and conditions, usually without even reading them.
The waist of the hourglass
Clifford Tatum from Centre for Science and Technology Studies, U of Leiden draws on work by Zittrain (2009) and Akhshabi and Dovrolis (2011) to suggest a model for developing an infrastructure for research information. He gives analogies of the development of the Internet and personal computing:

As the internet has developed, the relatively stable and freely available Internet Protocol (IP) has given us a layer we can rely on. Innovation can take place “above the waist” in terms of competition and multiplicity of delivery options (email, http, voip etc.) and can also take place “below the waist” in terms of provision options (fibre, copper, PPP, radio etc.) Similarly in the development of personal computing the three major relatively stable and available operating systems, Windows, Unix and Mac (with their offshoot mobile versions) can be seen as a middle layer. Below the waist are multiple hardware developers providing multiple kinds of devices; above the waist are multiple software developers supplying browsers, word processors etc. The key in each case is that (ideally if not always completely in reality) competition for applications depends on, but does not disrupt, the interoperability layer and, similarly, new technologies and infrastructures can be implemented without disruption to interoperability between layers. So the relatively stable interoperability layer remains consistent while innovation is pushed to the upper and lower layers.

Researcher profiles exist in many different places, most are commercial and it is not possible to export data and reuse it. But open and persistent IDs could be at the waist of the Open Scholarship hourglass in a similar way to how the IP address is at the waist of the internet protocols hourglass. Above the waist, competition and innovation (e.g. among service providers) is pushed to the outer layers; below the waist similar innovation in authentication, rights management, ID resolution and analytics could happen. Tatum’s vision being that as the system evolves, the thin layer of (core) interoperability remains relatively stable, giving us adaptability without disrupting interoperability or future development.

Footnotes
14 ResearchGate ([researchgate.net/about](http://researchgate.net/about))
15 Academia.edu ([academia.edu/about](http://academia.edu/about))
16 PLoS ONE ([http://journals.plos.org/plosone](http://journals.plos.org/plosone))

The waist of the hourglass: Research info

- competition and innovation (e.g. among service providers) is pushed to the outer layers
- adaptability without disrupting interoperability
- as the system evolves, the thin layer of (core) interoperability remains stable

Project demonstrations, prototypes and presentations
Using VIVO to pull together personal profiles
Lambert Heller from TIB Hannover (German National Library of Science and Technology) showed prototypes using VIVO which will potentially allow services to “open up and connect scholarly profile pages”. VIVO is a current research information system (CRIS) and the prototype integrates CVs from multiple institutions and databases. In the background, everything is stored in RDF triples e.g. person-authorship[role]-article. In a TIB student project, VIVO has also been enhanced to display interactive visual representations of co-authorship and thematic networks; this second prototype can be accessed on GitHub. Heller discusses further whether this kind of development might form the basis for a viable alternative to ResearchGate, Academia et al in his article on the LSE blog.

Certified repositories
Ingrid Dillo from DANS (Data Archiving and Networked Services) Netherlands, told us that Open Scholarship requires open data and that if we want to keep our data and ensure and demonstrate that they are fair, then they should be stored in certified repositories. Certification is fundamental in guaranteeing the trustworthiness of digital repositories, and thus in sustaining the opportunities for long-term data sharing and corresponding services. There are several certification schemes but the two basic and accessible ones that Dillo describes are the Data Seal of Approval (DSA) focused on humanities and social sciences and the WorldDataSystem (WDS) focused on earth and space sciences. They have agreed to work closely together as part of the Research Data Alliance (RDA) – and formed a working group under the umbrella of the RDA/WDS Interest Group (IG) on Certification of Digital Repositories. The remit of the group is broad, extending from harmonised procedures and requirements to workflows, security, licensing and confidentiality. Dillo believes that reliable and trustworthy repositories are crucial for the future of scholarship and therefore certification is an essential element of the EU “open science cloud”.

Open data on atmosphere and ecosystem
Jenni Hyppölä from CSC Finland presented an example of open science and open data. The Finnish atmospheric science community is internationally recognised and Open Science and Research is an initiative funded by the Ministry of Education and Culture with the aim of making Finland one of the leading countries in open science and research by 2017. The community runs measurement stations across the country, monitoring atmosphere and ecosystems and taking a wide range of measurements including temperature, concentration of various compounds, fluxes etc. A long time series is collected continuously. Using the AVAA open data publishing platform, the initiative, together with the researchers themselves, has developed access tools, applications and interfaces for data download, analysis and visualisations and also case studies of the use of open research data. The principles adopted for data management and storage are that the data will be:

- safely stored in the long term
- available with suitable licence to scientific and other use
- furnished with necessary metadata and other documentation
- made available for metadata searches internally and externally, with a commonly accepted vocabulary
- available in totality, including data sets produced by visiting scientists, students and other transient staff
- attached to persistent standardised and unique identifiers, which connect the data set, metadata and any citation

The SMEAR (Stations for Measuring Ecosystem-Atmosphere Relations) project makes extensive datasets available. The data are licensed under Creative Commons 4.0 Attribution (CC BY) licence.

Open Library for Humanities
Caroline Edwards from Birkbeck, University of London, introduced the Open Library of Humanities (OLH), which is a charitable organisation dedicated to publishing open access scholarship with no author-facing article processing charges (APCs). It is funded by an international consortium of libraries which has joined in the mission to make scholarly publishing fairer, more accessible and rigorously preserved for the digital future. OLH uses a new business model – the “Library Partnership Subsidy” – not a subscription model, to subsidise world-wide use. Partners get a stake in governance of the service with 155 libraries signed up so far and a target of 300 participating libraries at an average contribution of $850 per library. The OLH publishing platform supports academic journals from across the humanities disciplines, as well as hosting its own multidisciplinary journal. The OLH platform mitigates the social and economic challenges to new OA publishers by hosting pre-existing journals which bring reputations and reader bases with them, builds OLH prestige and financially benefits participating partners. This makes “flipping” journals from subscription models to Open Access more economically and technically feasible.

Comments about OLH from the floor and in breakout sessions indicated a widespread support for this achievement in a difficult area. There is by no means consensus amongst humanities and social science scholars about Open Scholarship and the UK government's policy announcements and stipulations have been met with scepticism and hostility. Those present in Helsinki from many scientific and other disciplines recognised the significance of the OLH and suggested it as an exemplar for others, agreeing with David Armitage, Lloyd C. Blankfein Professor of History at Harvard, that:

The Open Library of Humanities is a transformative venture on the leading edge of open-access initiatives on both sides of the Atlantic. As ambitious as it is well-planned, it presents a cogent vision of the future with well-designed pathways to its realisation. There is hardly a more important project in train for scholarship in the humanities today.

Open Science, Research Ideas and Outcomes
Daniel Mietchen from the National Institutes of Health in the US spoke about publishing and wikimedia, the #openscienceprize and infrastructure. Mietchen seems to echo some of the concerns raised in earlier presentations when he and a number of co-authors say in PLoS Computational Biology Meets Wikipedia that for computational biology:

our profession seems to gain [from Wikipedia] more than it gives. We suggest a principal reason for this limited breadth and depth of coverage of topics in computational biology is one that affects a number of disciplines: reward. Authors in the biomedical sciences get academic reward for publishing papers in reputable journals that are indexed in PubMed and have associated digital object identifiers (DOIs). In contrast, contributions to Wikipedia can be anonymous and do not count for much in the current system of academic advancement. We hope to help to resolve this disparity in PLoS Computational Biology.

Mietchen also mentions the Wikiversity Journal of Medicine, which is an open access scientific journal in medicine and biomedicine and which seeks to gather and provide free access to medical knowledge and make such knowledge available to be used to improve Wikipedia articles or other projects that are operated by the Wikimedia Foundation.

Footnotes
20 Data Seal of Approval (http://datasealofapproval.org/en).
21 WorldDataSystem (icsu-wds.org).
22 working group (http://bit.ly/1Mm3Z5a).
23 Open Science and Research (http://openscience.fi/services).
24 AVAA (http://avaa.tdata.fi/web/avaa/etusivu).
26 Open Library of Humanities (openlibhums.org).
The Open Science Prize is a partnership between the Wellcome Trust, the US National Institutes of Health (NIH) and the Howard Hughes Medical Institute to “unleash the power of open content and data to advance biomedical research”. The prize provides funding to encourage and support the prototyping and development of services, tools or platforms that enable open content – including publications, datasets, codes and other research outputs – to be discovered, accessed and re-used in ways that will advance discovery and spark innovation. It also aims to forge new international collaborations that bring together open science innovators to develop services and tools of benefit to the global research community.

The DOI event tracker (DET) seeks to help understand better how scholarly research is used. Scholarly content has always been discussed by scholars outside the formal literature and by others beyond the academic community. DET is a new service which tracks activity surrounding a research work from potentially any web source where an event is associated with a DOI. Following a successful pilot run started in Spring 2014, the service has been approved to move toward production and is expected to launch in 2016.

Mietchen is particularly concerned that it is recognised that the research process is more than the “final results” – he asks shouldn’t the rest of the research process be published, shared and openly accessible? The Research Ideas and Outcomes (RIO) journal publishes all outputs of the research cycle, including: project proposals, data, methods, workflows, software, project reports and research articles together on a single collaborative platform, with the most transparent, open and public peer-review process. The scope encompasses all areas of academic research, including science, technology, humanities and the social sciences.

Footnotes
32 The Open Science Prize (opencodeprize.org).
33 DOI event tracker (bit.ly/1q2hzAT).
34 Research Ideas and Outcomes (http://riojournal.com).
Many questions were posed during the breakout sessions and plenaries in this event. Below, we select a few which we feel represent significant themes which recurred several times and are questions which need to be answered. Knowledge Exchange alone cannot answer these questions, but it is very important to Knowledge Exchange that answers are forthcoming in order to define the direction and nature of future work and the actions that we should be taking and suggesting that others take. More than that, the answers to these questions will be important for our community to fully realise the potential of Open Scholarship.

To all researchers: will your data still be available to others even when your interests shift or you move job? Can you guarantee that your data will not be interesting, useful or even vital to another researcher in 5, 10 or 100 years’ time?

To institutions: will you consider offering by default to all your research staff the opportunity to make their publications and data openly available? In the words of one speaker: “it should be so easy that it is seen as part of the plumbing”. There are many ways to do this – through local or national or international repositories, consortium agreements or Open Access publishers.

To senior researchers: will you reconsider the advice you give to your postgraduates and early career researchers? Even just a few years ago the Open Library of Humanities described earlier in this report would have seemed highly unlikely. What changes will have affected the world into which your postgraduates will be reporting and/or publishing in 5 or 10 years’ time?

To early career researchers: how might Open Scholarship benefit you, your colleagues and your audience? How many more eyes (including those outside of your immediate field) are likely to see your work if it is Open? A perspective on these questions by a young researcher and Scholarly publishing: a perspective from an early career academic - with a focus on publishing the research outputs relevant to early-career researchers.

To funders, EU and national governments: how might Open Scholarship benefit the interests of your citizens and your economies? Can you find ways to collaborate to set up an exemplar Open Scholarship Institute? How will you participate in and promote initiatives like the European Open Science Cloud for Research? How will you guide your researchers so that they properly handle sensitive data while not restricting legitimate research into that data, now and in the future?

To Knowledge Exchange: how can we best help answer some of these questions and achieve at least some of our wish list? Throughout the sessions at the conference it was noticeable that there was great and widespread concern at the costs of access for academics to publications and data produced by their colleagues, outputs whose production had often been publicly funded. Opinions on actions which should be taken in this regard broadly fell into two camps.

One group feels that the steps already being taken across the sector are bringing education and research to a better place regarding Open Scholarship. They see evidence that a step-by-step approach involving both top-down change (e.g. mandates and institutional and funders’ policies) and bottom-up change (communicating to, motivating and incentivising early career researchers) will bring stakeholders along on the journey.

The other group is deeply frustrated with what it perceives as a lack of transparency on the costs and subscription models of journals. This group wants to see a faster pace of change and is advocating some radical disruptive measures.

So our final wish is that Knowledge Exchange can continue to act as a bridge between the two schools of thought.

Footnotes
35 A perspective on these questions by a young researcher (http://bit.ly/1VoOjpz).
36 Scholarly publishing: a perspective from an early career academic? (youtube.com/watch?v=L65TG9xzGfY).
Beyond Helsinki

We would like to build on our past successful activities to ensure that gradual change continues and necessary developments are progressed but combine this with more radical pilots in coalitions of risk-takers willing to try something significantly new and different. Bridge building is something that Knowledge Exchange does well – it is ideally placed to help stakeholders come together and to have a real influence on shaping developments.

Discussion in the breakout sessions, plenaries and presentations in Helsinki all made it clear that there are indeed important issues to be raised and bridges to be built. After internal and external consultations, the issues discussed in this report will be reflected in Knowledge Exchange’s activity and engagement plans. We will explicitly ask our expert groups, partner organisations and selected stakeholders in the Open Scholarship arena to respond to this report and to prioritise the key and urgent issues for Knowledge Exchange to take forward. We look forward to turning ideas into actions and continuing to act as a catalyst for change.
Appendix – The full wish list

The event took place in Helsinki in December so it seemed appropriate to write a wish list for Santa. But behind this jolly fancy was a serious intention to focus on what was needed to promote and improve Open Scholarship. So this wish list, compiled from discussion in the breakout sessions and the plenaries, is “for life and not just for Santa”.

1. An independent Open Scholarship institute should be launched

2. Institutions to provide Open Scholarship publication options as part of the “academic plumbing” like the telephone and the academic email address; this will also need more librarians and tenure committees to encourage and support researchers to make their publications open as preprint or postprint

3. Show researchers, and encourage them to use, the presents that are already waiting for them to open: e.g.
   - the Open Access button
   - Contributor badges and Open Research Badges
   - 6000+ researchers telling about their research practices
   - ACUMEN Portfolios for individual academics
   - Science in Transition
   - Joint Declaration of Data Citation Principles
   - The Open Library of Humanities
   - Research Ideas and Outcomes
   - ReScience
   - Wiki publishing
   - Data Management Plan tools such as: UC’s DMPtool and the Digital Curation Centre’s DMPonline

4. Show Institutions, funders and other actors, and encourage them to use, the presents that are already waiting for them to open: e.g.
   - The NISO Alternative Assessment Metrics (Altmetrics) Initiative
   - Article-Level Metrics like Lagotto
   - Declaration on Research Assessment (DORA – written for funders and institutions)
   - The Metric Tide: Final Report and executive summary of the independent review of the role of metrics in research assessment and management
   - The Leiden Manifesto for research metrics (original article in Nature)
› **SCOAP3** – a partnership of thousands of libraries, funding agencies and research centres in more than 40 countries – has converted key journals in the field of High-Energy Physics to Open Access at no cost for authors

› **FORCE11** – a community of scholars, librarians, archivists, publishers and research funders aiming to bring about a change in modern scholarly communications toward improved knowledge creation and sharing

› **How Open is Your Research?** – a tool for institutional managers to help them visualise how open is the way they manage and disseminate their research and teaching output

› The **PASTEUR4OA** project is developing a similar tool for individual researchers, plus a simple summary for institutions of the advantages of Open Data, plus a briefing for research managers on openness as a key element of the European Commission’s Responsible Research and Innovation (RRI) agenda

5. **Researchers of tomorrow** – *The research behaviour of Generation Y doctoral students* – showed us that we need to actively make Open Scholarship something a young career researcher will want to do. One possibility is to create some sort of Early Career Researcher Protection Programme which takes a Science in Transition approach (protecting early career researchers in the current system whilst also changing to evaluation which supports Open Scholarship). And a handbook for young researchers to navigate the current system and to help them take action to influence a transformation

6. Reliable metrics or a calculator for likely revenues coming from an open approach (who might create this?). Plus an extension of the Open Access button to a “monthly statistic sent to authors detailing how many people ran into a paywall for your articles and subsequently neither read nor cited it

7. Simple to understand guidelines (from whom, suggestions please?) on sensitive data handling; and an easily-used and widely recognised and adopted approval process declaring data “safe to publish”

8. Encourage all of us in our working lives to challenge any use of proprietary metrics and to suggest replacement by public/open metrics, using our data and our building power

9. New systems for reflecting merit, contribution and worth of individual researchers (and, aggregated, institutions). Less measurement by (individual) citations, more by (collective) research progress

10. Provide incentives to ensure Open Scholarship makes sense to an individual’s career as well as to society in general (see above for badges providing Reputation Mechanics) and give researchers credit for the everyday activities they perform / contribute in their (open) research endeavour

11. Gaming or manipulating metrics can be a problem. Rather than looking for one magic metric, maybe encourage a large number of different ones; then gaming becomes far more difficult and researchers are stimulated to just do good research (“we want to be human again”). Other opinions on this also expressed – some want one number that indicates impact of research to society

12. A broad international coalition of libraries could collaborate to push for completely transparent journal costs. By working together they could act boldly for Open Scholarship and move away from the current economic and reward system …

13. Pursue innovative and open models for access to research articles. This could help build the marketplace for Open Scholarship publishing and be a significant step towards construction of a new economic and reward system. See the Quality Open Access Market, which provides quality scoring of the journals in QOAM, based on academic crowd sourcing

14. Work with the Learned Societies to take a new look at Open, to reconsider current financial models and consider other than traditional approaches:
a citation-sabbatical to experience how it can work in an alternative open environment

look for practical, pragmatic strategies and avoid the aggressive evangelism and past religious debates of Open (which may not be palatable when under severe economic pressure)

15. Presents in the stockings for those who do practice open research, and those who do counter problems with solutions:

- the ability to draw information and expertise from a number of different places in the research lifecycle. Multinational publishers can draw from many places in their “walled gardens” – can we replicate this by drawing from a number of Open Scholarship sources and services? Who can bring this together?

- an integrated approach to evaluation/judgement/ portfolios which includes peer review plus tools which deal with a multiplicity of quality criteria. Who is likely to bring this together?

- lots of copies of everything in different places. LOCKSS is a good example – we need to go further, to protect against the danger of depending on one storage location / one service provider / one source of funding / one government. So all services must demonstrate that the data they hold and the data structures on which the services depend will not disappear if their funding or their business model or the politics of their government suddenly changes.

- astronomers, human genome and particle physics communities have used their history of collaboration to establish control over and access to their publications. Whether by circumstance, good luck or good management, these are examples of good practice – are there lessons for us to learn here? How might this spread to other communities?

16. Some more technology presents:

- long term archiving and preservation technology allowing Rediscovery

- automation … technology that automatically adds metadata (which I can review) and minimises the effort needed to deposit data

- simplification … technology that aids and assesses creation of metadata and contextual data (e.g. provenance, error estimations) and estimates likely quality, giving credit to the author for their efforts to ensure data quality

- technology which seeks to ensure that the experiments described in my papers are reproducible and preserved over XX years and allows me to share and reuse methods described by others

17. A reward system for senior staff to incorporate and reward Open Scholarship practice in their research group. Encourage them in their supervision and mentoring of early career researchers to be part of the transformation and not to assume that past experience is necessarily the key to current and future career progression

18. Make it more attractive to publish the research process underlying those final outputs. For humanities in particular, make it easy to publish the story around the article along with the article, a set of attached things around the process – also include multiple authors and collaborators, moving away from concept of individual authorship which is still predominant in humanities

19. Different disciplines will require different approaches. The scale and cost of data has an effect (is it true that some data – e.g. at CERN – is so expensive and so huge that it is widely recognised that it must be shared to make progress, while other much smaller data is sometimes closely guarded?). Help us to understand better the disciplinary differences. Allow us to accept complexity, while striving for simplicity – help us to remember there are parallel worlds to Open Scholarship

20. Let’s take this wish list beyond wishing and turn it into a to-do list in 2016! Help us build an action plan beginning to point to what actions we should be taking and who else we should approach to take actions – e.g. Knowledge Exchange partners, learned societies, high profile champions, institutions, transnational collaborations
Footnotes
37 Open Access button (https://openaccessbutton.org).
38 Contributor badges (https://badges.mozillascience.org).
40 6000+ researchers telling about their research practices (http://bit.ly/1ME8ovc).
41 ACUMEN Portfolios for individual academics (http://research-acumen.eu/portfolio).
42 Science in Transition (scienceintransition.nl/english).
43 Joint Declaration of Data Citation Principles (http://bit.ly/1MpAo6p).
44 The Open Library of Humanities (openlibhums.org).
45 Research Ideas and Outcomes (http://riojournal.com).
46 ReScience (http://rescience.github.io).
47 Wiki publishing (http://compbiolwiki.plos.org/wiki/Main_Page).
49 DMPonline (dcc.ac.uk/dmponline).

In the PASTEUR4OA project, Alma Swan is working on a briefing on the current picture on Data Management Plans – who requires them, what are the trends, see also below for more on PASTEUR4OA. Also promised but not yet fully delivered is the fledgling Open Citations Corpus (http://bit.ly/1UOg1pv) (https://is4oa.org/services/open-citations-corpus/).
50 The NISO Alternative Assessment Metrics (Altmetrics) Initiative (niso.org/topics/tl/altmetrics_initiative).
51 Lagotto (lagotto.io).
52 Declaration on Research Assessment (ascb.org/dora).
54 The Leiden Manifesto (leidenmanifesto.org).
56 SCOAP3 (https://scoap3.org/what-is-scoap3).
57 FORCE11 (force11.org/about).
58 How Open is Your Research? (http://sparceurope.org/howopenyourresearchis).
59 PASTEUR4OA (pasteur4oa.eu/home).
60 Researchers of tomorrow (jisc.ac.uk/reports/researchers-of-tomorrow).
61 Science in Transition (scienceintransition.nl/english).
62 Quality Open Access Market (qoam.eu/about).
Knowledge Exchange Office
C/O Jisc,
One Castlepark,
Tower Hill,
Bristol, BS2 0JA

t: +44 203 697 5804
e: office@knowledge-exchange.info