Introduction

Wadham College in Oxford provided the historic setting for the 2016 Jisc and CNI conference. The event was designed to bring together senior practitioners from across the scholarly communications stakeholder spectrum to explore and debate many of the issues and challenges faced by the sector. These include open data, research metrics and indicators, the sustainability of open access infrastructure and analytics in research and learning. The following day a group of experts from the UK and USA were invited to consider and identify opportunities for greater collaboration through detailed discussions on the themes of first, creating, sustaining an using the scholarly record and second, enabling the scholarly record for future research and impact.

At the opening reception Richard Ovenden, Bodley’s Librarian, treated delegates to a canter through the rich history of the Bodleian Libraries. Inevitably perhaps much of the discussion during the reception was about the result of the UK’s referendum on leaving the European Union and its likely impact. This was a theme picked up by Jisc’s Chief Executive, Paul Feldman, in his conference introduction. Feldman characterised Brexit as an opportunity for change for Jisc and the sector and an opportunity to build new collaborations. Jisc has been relatively quiet while it has been undergoing a process of transformation from a funding body to a membership organisation, emerging stronger and more focused. Jisc is now doing its own development on a small number of high impact programmes, filling in the gaps where commercial providers cannot provide the necessary services. These initiatives include learning analytics, providing opportunities to undertake further research on the learning process; research management covering issues like open access, open science and open data; and the processes and data curation issues that attend the development and operation of research data repositories. In short, Jisc is acting as a catalyst for development, as a powerhouse in the research and education sector. Jisc is ambitious to do this on a global basis, a goal that underscores the importance of conferences such as this, which bring together senior stakeholders from the UK and the USA to work towards finding common solutions to common challenges.

The importance of the Jisc and CNI series of conferences was underlined by Clifford Lynch, Executive Director of CNI, during his introductory remarks. They are regarded as an opportunity to learn and lay the groundwork for collaborations. Although the situations are different in the USA and the UK there are fascinating parallels and intertwining issues making it a provocative time to revisit some of these challenges. The landscape is currently very unsettled with respect to significant political, financial and demographic shifts in both countries. The past decade has seen a profound shake-up in scholarly communications; there has been progress in open access in the journal article context with an emphasis on the sciences; there has been greater connection between funders’

The importance of the Jisc and CNI series of conferences was underlined by Clifford Lynch, Executive Director of CNI, during his introductory remarks. They are regarded as an opportunity to learn and lay the groundwork for collaborations. Although the situations are different in the USA and the UK there are fascinating parallels and intertwining issues making it a provocative time to revisit some of these challenges. The landscape is currently very unsettled with respect to significant political, financial and demographic shifts in both countries. The past decade has seen a profound shake-up in scholarly communications; there has been progress in open access in the journal article context with an emphasis on the sciences; there has been greater connection between funders’
policies and the open nature of research outputs; there has been a recognition of the need for systematic research data management strategies. Although policy is falling into place, there is still a gap between policy and sustainable implementation at scale. We are still spending a lot of time dealing with things recognisable from one hundred years ago – books and journals for example – but we might soon start the transition to newer forms that will call for different thinking and structuring. On the subject of the measurement of impact and contribution, bad measures lead to distorted systems and it is necessary to find a way through this. The landscape is challenging but also exciting. Lynch concluded his comments by noting that it is easy to be parochial about scholarly communications but that it is at heart a global enterprise and needs to be thought about in that way. We need broad scale strategies rather than national approaches.
Keynote address

Professor Sir Nigel Shadbolt, Principal of Jesus College and professor of computer science at the University of Oxford delivered the keynote address emphasising the importance of open data and developments in artificial intelligence (AI). Although the volume and ubiquity of data is increasingly rapidly the degree to which data is open is not keeping pace. The UK is top of the open data league table but persuading people of the benefits of open data is a continuing process. There seems at times to be an innate desire by public bodies to privatise data as a way of realising short-term value for the taxpayer. In reality this is invariably a short-lived gain since the true value and power of open data is its influence on innovation. Many examples can be found but one of the most compelling concerns genomics. In the race to sequence the human genome the US Supreme Court had to strike down a commercial patent on the human genome; the fact that the data is open has created a huge amount of value.

The Open Data Institute (ODI) has developed the Data Spectrum\(^1\) to help data providers understand better where their data provision sits on a spectrum ranging from “closed” through “shared” and finally “open”. The ODI also often rehearses the value of open public data because, even though the benefits are apparent, changing attitudes and behaviour is a long-term challenge. Companies that run public services through private contractors can make it difficult to get, for instance, railway fare data. It’s not just that such companies think others will not be able to understand the data or come up with better ways of extracting value from it; rather it comes down to companies wanting to charge monopoly rent to sweat the value of closed data. And yet it is known that that sharing data leads to better outcomes. The data on mortality rates in the recent Ebola outbreak enabled a lot of global effort to be focused on finding a solution. Similarly, data on death rates and obesity in England and Wales have led to positive policy discussion and changes. Analysis on prescribing data\(^2\) has shown that prescribing generic statins rather than their branded equivalents would save unnecessary expenditure by the NHS in England to the tune of £200 million. More widely, data on health trends led to the development of actuarial analyses and the health-related insurance market. And yet open access to public data is still often restricted: witness the privatisation of the Royal Mail’s postcode address data.

Although there are powerful arguments in favour of open data there remain some significant challenges. The geographic extent of datasets is one such challenge: there are many different ways to partition the UK in spatial terms, a situation made even more challenging by the process of devolution. Comparing or overlaying datasets in meaningful ways is difficult. Another challenge is establishing the authority and provenance of data. Above all there is a need for simple standards enabling the machine-readable characterisation of important variables. The community is moving towards the development of a set of registries; once a set of persistent identifier can be agreed really interesting things can be done with open data. The notion of “data as infrastructure” is gaining traction: there are some datasets that are so important they act as nucleation points for other data. Open linked

---

2. [http://www.prescribinganalytics.com](http://www.prescribinganalytics.com)
data such as land registry data and the data provided by Companies House provide really important connection points. This idea supports a wider view of data as a public good.

Looking ahead artificial intelligence (AI) continues to tantalise with the promise of future benefits. Beyond text mining and applications such as SIRI the emphasis has moved towards DeepMind\(^3\) and deep learning\(^4\) where computers observe patterns over time to produce behavioural outputs on a par with what humans can achieve. These advances can be useful for tasks like large-scale image classification. More generally, in the future information sharing economies will be more prevalent and existing canons about where information is held and shared will be challenged. Lots of start-ups are making a feature of crowdsourcing information to produce new information and new analyses. There will be pressure to move forward as a broad based community in contrast to the current situation where bilateral agreements between one commercial service provider and one institution proliferate. Along the way it will be necessary to address issues such as the cost of archiving, data silos, reproducibility, reliability and how to capture context for effective data curation. And as the data-related situation becomes more complex so the need for simplified design becomes more pressing. Simplification can be found in the development of persistent identifiers, simple citation and trackback mechanisms. There is a need to find fractal structure in the solutions adopted by the community.

The brief discussion session at the conclusion of the keynote address raised a number of interesting issues. The important role of preservation was addressed including the need for better standards, agreement on neutral formats and acknowledgement that the community needs to agree how best to decide what data is worth keeping and what might reasonably be subject to “sunsetting”. The “elephant in the room” is not just the patent system but also the role of university technology transfer officers wanting to lock up research outputs for financial gain or competitive advantage. People are thinking about the extent to which patents impede or support innovation and it was suggested that universities should think about other ways of extracting value from research. The world of protected intellectual property is unlikely to go away but the speed of innovation is such that there may be more value in not applying the brakes to this process by applying for patents. On a similar topic it was acknowledged that researchers should have the right to work out whether their data is “mad or bad” and a period of grace to get the first cut of the data and the first publications but that the data should be released after a reasonable time on the basis that others may be able to do more with the data than the creators can.

Finally, there was speculation on how to trade off motivating companies to share data but also being able to trust the data they are producing. The opaque nature of university rankings is a particular cause for concern. The answer may lie in recognising that public bodies produce much of the underlying data used for ranking purposes and that universities should club together to produce a range of open data sets as the basis for the production of transparent rankings.

---

\(^3\) [https://deepmind.com](https://deepmind.com)
Breakout sessions

Reflecting the breadth of topics covered by the conference delegates were able to choose from four different breakout sessions, morning and afternoon, encompassing 28 different short presentations in addition to question and answer sessions. Since the volume of information would overwhelm this concise report of the event as a whole, the key points, challenges and opportunities from each session have been distilled by various different note takers and are captured in tabular form below. More detailed information can be found by following the links to the presenters’ slides.

Research metrics and indicators

*Chair: Anne Trefethen, University of Oxford*

*Contributors:*

The Metric Tide. *Stephen Curry*, Imperial College; *Ben Johnson*, HEFCE

Open citation. *Cameron Neylon*, Curtin University

Open infrastructures. *Clifford Tatum*, Leiden University

| The main focus of discussion: | • A great deal of effort goes into quantifying the qualitative process of peer review including the REF despite cultural resistance to metrics and evidence that some researchers “game” the system.  
| | • A consistent and meaningful definition of research “excellence” is elusive which makes it difficult to measure.  
| | • Monitoring open access publishing activity is difficult and imprecise. |
| The main challenge: key concerns or issues raised | • There is a need to reward better collaboration but current metrics are extracted at an individual level.  
| | • The academy is becoming overwhelmed by metrics, which can be meaningless or opaque, especially when used for university ranking purposes.  
| | • Lack of interoperability is as much a problem today as it was 20 years ago.  
| | • Perhaps the question about excellence should be: what are the qualities of research that deliver desirable results (in terms of beneficial outcomes in the areas of health, environment, culture, education and so on).  
| | • Researchers are sympathetic to open access but complain about the administrative impositions especially around funder policy compliance. |
The main opportunity: was there anything to be explored, or done

- An independent review by James Wilson led to the publication in 2015 of the Metric Tide report that backed peer review and recommended research communities better understand the role of metrics. There is work to be done helping researchers understand the issues.
- There are many more signals beyond simply citations and we could usefully be measuring them.
- We need to focus more on the science of evaluating knowledge.
- The CWTS Leiden Ranking seeks to measure the proportion of an institution’s published output resulting from collaboration with other institutions.
- Suggested that persistent IDs form the thin, open middle part of the hourglass model of the Internet’s interoperability layer.
Open access progress and sustainability

*Chair:* Neil Jacobs, Jisc

*Contributors:*

UK and US positions on Open Access. *Steven Hill,* HEFCE; *Sarah Thomas,* Harvard University

University of California and university digital library costing models. *Mackenzie Smith,* University of California, Davis

Total cost of ownership and flipped OA. *Liam Earney,* Jisc

| The main focus of discussion: | • Policy driving open access (OA) adoption. Government policy drives funder policy.  
• But US is different to Europe. The common theme of difference is US: green, Europe/UK: gold/APC preference.  
• The economics and costs of moving to OA. Looking at potential models/scenarios of flipping in the US, and looking at current move to APCs and offsetting deals in the UK. |
| --- | --- |
| The main challenge: key concerns or issues raised | • If flipped to the full OA model, APCs costs are more than subscription costs for a research-intensive university/library. i.e. total APC spend would be more than current library resource budget. How to address this? How to introduce market forces to make this sustainable?  
• Questions between the benefits and issues of green and gold still remain. US/Europe are going down different paths, which may lead to issues; it would be better to have common approach (and a common approach between different national funders).  
• APCs could rise substantially from traditional publishers.  
• Green: is it sustainable? Still paying for the increasing journal costs while also paying for green infrastructure.  
• How can we reduce the costs of publishing? Can we transform scholarly publishing while keeping the same players?  
• What are the constraints on publishers in a sub free world? |
| The main opportunity: was there anything to be explored, or done | • Giving authors discretionary funds (using funds traditionally used for journal subs) to help introduce APC competition without interfering author choice.  
• Offsetting agreements have potential, also flaws, to drive adoption to OA in sustainable way, transferring spend to publishers from subs to APCs.  
• Aligning international and funder policies to help achieve consistent approach |
Data sharing and analytics in research and learning

Chair: Martin Hall, consultant

Contributors:

Learning analytics: progress and solutions. Niall Sclater, Jisc; Michael Webb, Jisc

Reading analytics. Clifford Lynch, CNI

Sharing data safely and its re-use for analytics. David Fergusson, The Francis Crick Institute

| The main focus of discussion: | • Ethics of analytic practice  
|                             | • Convergence of approaches  
|                             | • Data security and access  |
| The main challenge:         | • New and emerging applications and possibilities  
|                             | • Relationship between academic applications and commercial innovations  
|                             | • Role of students  |
| The main opportunity:       | • Convergence of different approaches to analytics  
|                             | • New approaches to informed consent  
|                             | • Continuing priority of data security and access  |
Equipping the researcher – patterns in the UK and USA

*Chair: Louisa Dale, Jisc*

*Contributors:*

UK and US academic practices. *Christine Wolff*, Ithaca S+R; *David Prosser*, RLUK

Digital scholarship centres. *Harriette Hemmasi*, Brown University; *Joan Lippincott*, CNI

Software carpentry and software skills and practice. *Neil Chue Hong*, Software Sustainability Institute

| The main focus of discussion: | • Skills for digital scholarship - not just for postgraduates/research students/staff but also undergraduates/students below postgraduate level.  
• Roles: curation is important including for research data; the role of the librarian in supporting the researcher.  
• Collaboration: digital scholarship involves mixed, often-virtual teams including librarian roles. |
| --- | --- |
| The main challenge: key concerns or issues raised | • Skills: need to prepare researchers at an early stage to handle the challenges of discovery, how to cope with ‘digital doubt’ and take care of their own wellbeing.  
• Visibility: the changing nature of the library space (not just a building but services to support scholarship) and its relationship to digital space. Digital scholarship is not easily visible.  
• Openness: need to consider “findability” of research by the non-affiliated researcher: when something is ‘open’, who is it open to? |
| The main opportunity: was there anything to be explored, or done | • Policy: need to drill down into the question of how the policy environment shapes digital scholar behaviour (differences in US and UK).  
• Collaboration: need for librarians and others to work together to make data management work in different disciplines (“we need to co-design findability”).  
• Skills: need to focus on equipping researchers at all levels with the capabilities required to flourish, with attention paid to the human dimension. |
# Tracking research and research systems

*Chair: Catherine Grout, Jisc*

**Contributors:**

- Digital scholarship and identifiers. *Geoffrey Bilder, CrossRef*
- SHARE update. *Elliott Shore, ARL*
- Jisc Monitor update. *Neil Jacobs, Jisc*

**Infrastructure and services to track research activity. *Daniel Hook, Digital Science***

<table>
<thead>
<tr>
<th>The main focus of discussion:</th>
<th>Research information infrastructure and standards in research data.</th>
</tr>
</thead>
</table>
| The main challenge: key concerns or issues raised | • Sustainability of standards.  
• Organisational IDs.  
• Capturing key research events in a standardised way, who decides which events to capture. |
| The main opportunity: was there anything to be explored, or done | • Building consensus around identifiers.  
• Getting international infrastructure sustainable.  
• More collaboration and exploitation of and collaboration within existing networks (e.g. ARL). |
## Repository and preservation systems

*Chair: Chris Keene, Jisc*

*Contributors:*

Research data management shared service for the UK. *John Kaye, Jisc*

Hydra. *Tom Cramer, Stanford University; Chris Awre, University of Hull*

Addressing the preservation gap at the University of York. *Jenny Mitcham, University of York*

Emulation developments. *David Rosenthal, Stanford University*

| The main focus of discussion: | • Progress towards the provision of a research data management shared service; currently building the supplier framework.  
• The flexibility offered by Hydra to help create sustainable repository infrastructure to serve digital content management needs.  
• Articulating the challenges faced by institutions looking to preserve research data.  
• The possibilities offered using emulation as a preservation strategy. |
|---|---|
| The main challenge: key concerns or issues raised | • There are no neat solutions to the RDM problem; messy and partial solutions area pragmatic response but tools are needed to help people address the messy components.  
• Droid, the database used to identify file formats to help with preservation planning, needs more content to improve its utility.  
• Legal issues are the biggest barrier to the use of emulation in preservation. |
| The main opportunity: was there anything to be explored, or done | • Hydra-in-a-box brings the best of the Hydra elements that have successfully been deployed into a single package.  
• York's Archivematica system is providing preservation functionality as part of an RDM service and could be developed and rolled out further.  
• Emulation technology will continue to improve and developments in container technology may help simplify the preservation process.  
• The UK’s HE community is invited to continue to engage with the RDM shared services development process; quarterly events and a Research Data Network wiki are planned. |
Incentives for modern research

Chair: Steven Hill, HEFCE

Contributors:

Incentives for sharing research data. Veerle Van den Eynden, UKDS

Incentives to innovate. Joe Marshall, NCUB

Incentives in university collaboration. Tim Lance, NYSERNet

Giving researchers credit for their data. Neil Jefferies, BDLSS

<table>
<thead>
<tr>
<th>The main focus of discussion:</th>
<th>• Incentives for researchers and facilitation of working with these drivers.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Incentives for research from business and economic need.</td>
</tr>
<tr>
<td></td>
<td>• The relationship of policy to research norms.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The main challenge: key concerns or issues raised</th>
<th>• How to link researcher incentives to the incentives, or the socioeconomic rationale, for research.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• The continuing divide between industry and academia - different drivers, different language, finding the right contacts and communication channels.</td>
</tr>
<tr>
<td></td>
<td>• Matching policy to research norms, when these vary by age, discipline, country: matching the vision or pace of policy change to researcher appetite and emerging research process needs.</td>
</tr>
</tbody>
</table>

| The main opportunity: was there anything to be explored, or done | • Collaborate with researchers and policy makers to identify and facilitate engagement with incentives: |
|                                                               |   - match incentives to all actors |
|                                                               |   - analyse pinch points in research cycle |
|                                                               |   - spot opportunities to do the heavy lifting - from single-function support through to infrastructure creation |
|                                                               |   - facilitate through technology for scale |
|                                                               |   - support as service |
|                                                               |   - collaborate and engage to achieve buy-in |

|                                                       | • Ensure policy needs and research expectations are satisfied by infrastructure before expecting compliance: |
|                                                       |   - close knowledge of researcher work-flows (existing and emergent) to maintain understanding of feasibility of policy development and gap-spotting for investment/support |
Active research management and sharing

Chair: David De Roure, OERC

Contributors:

Open science framework. Jeff Spies, Centre for Open Science

Active research from lab to publication. Simon Coles, University of Southampton

Managing active research in the university. Robin Rice, University of Edinburgh

Making research available – FAIR principles and Force 11. David De Roure, Oxford e-Research Centre

| The main focus of discussion:             | • Ecology of data management: a framework offers flexibility allowing multiple tools to interact.  
|                                         | • Openness: is not just an ideal, it’s about increasing research quality and efficiency  
|                                         | • FAIR: Findable – Accessible – Interoperable – Reusable (Force 11 principles) |
| The main challenge: key concerns or issues raised | • Not all data is highly structured; systems need to accommodate this.  
|                                         | • How to manage researcher concerns relating to the sharing of data and if/when data should be declared ‘dead’.  
|                                         | • Disciplines understand differently what it means for research data to be ‘active’ or ‘archived’. |
| The main opportunity: was there anything to be explored, or done | • There needs to be a focus not just on sharing data but on process.  
|                                         | • Need to take a lifecycle approach to managing research data in the pursuit of open innovation.  
|                                         | • Edinburgh University offer a MOOC on RDM as part of their training and support for researchers. |
Closing plenary

Infrastructure for US research and scholarship

John Wilkin, Librarian at the University of Illinois, Urbana Champagne (UIUC), described the ways in which the infrastructural situation is really rather different in the USA in comparison with the UK. In the USA there is a predisposition away from shared infrastructure tied into notions of federalism and individualism. Going it alone is a deeply embedded trait in the USA. There are no organisations similar to Jisc (the Association of Research Libraries doesn’t really do the same job) and there are no US government organisations set up to provide or facilitate the development of shared infrastructure. The lack of shared infrastructure is a result of the choices institutions have made, but it is a choice that eschews scale and efficiency. In fact where shared efforts can be identified, mostly at the local institutional level, they add to rather than substitute function and cost. It appears that the prospect of efficiency and cost savings are insufficient incentives for sharing.

Although there is very little relevant infrastructure for national collaboration and no sign of investment (even intra-institutional research teams are not supported sufficiently, with more than 60% having to rely on local or institutional infrastructure) it seems that successful collaboration can be achieved when it is necessary to do so. The prime example given is the Hathi Trust5 which emerged in the wake of Google’s digitisation efforts and which was born out of necessity. The shared infrastructure underpinning the service provided by Hathi Trust offers in the region of 14.5 million digital volumes in over a hundred institutions, 500,000 of which have been deposited by UIUC and 5.6 million of which are public domain volumes. The cost to UIUC is minimal. SHARE Notify6 is sometimes cited as shared infrastructure, designed to build on the sunk costs of institutional repositories. The key point to note, however, is that although SHARE contributes to the ecosystem it is an added layer of technology at added cost. It is the same situation with the Digital Preservation Network7, a large-scale digital preservation service: it adds functionality but also cost to the system. The final example provided is Unizin8, a collaboration of a number of US universities designed to support digital learning including the creation of common standards. Founding institutions contributed USD 1 million but there is insufficient public information to be able to determine the financial payoff. It appears to be, however, an efficient substitute for Blackboard and Moodle.

---

5 https://www.hathitrust.org
6 http://www.share-research.org/projects/share-notify/
7 http://dpn.org
8 https://en.wikipedia.org/wiki/Unizin
Efficient infrastructure for UK research

Providing a UK perspective, David Maguire, Vice Chancellor of the University of Greenwich and Chair of Jisc, started his presentation by noting that research in the UK is highly efficient: 76% of research submitted to REF 2014 was deemed to be world leading or internationally excellent. Policy developments on open access to journal articles and advances towards open data make the UK a leader in this respect supported by a range of Jisc’s open access services. The UK’s approach to and use of metrics has been studied carefully and the outcomes disseminated in the Metric Tide report. An independent review of the work of the UK’s Research Councils has been recently concluded and an independent study of the REF has been under way. Finally, a Higher Education and Research Bill is working its way through Parliament, which will impact on the current higher education and research landscape.

It has been suggested that there are a number of challenges that attend the REF. First, it is expensive to collect, manage and analyse the huge amount of data collected in each assessment; second, there is limited re-use of existing information; third, the same information may be collected multiple times; fourth, a lack of standards and access to suitable tools can make it difficult to analyse information and compare results; finally, building a clear picture of the state of national research strengths and priorities is challenging. In addition, national infrastructure has been developed piecemeal from the bottom up, resulting in a patchwork of infrastructure services including disciplinary repositories and data centres, open access infrastructure such as CORE and Jisc Monitor, RCUK’s Gateway to Research, the Janet network, cloud computing services such as JASMIN the Alan Turing Institute (the National Institute for Data Science) and a registry of research equipment. At the level of individual institutions the patchwork theme continues with CRISs, repositories and various other internal and external systems typically interoperating poorly if at all.

In response to the somewhat chaotic nature of the current information landscape Jisc has been developing a plan for a National Research Information Infrastructure (NRII). The key strategic goals of the NRII would be to reduce the burden on researchers, lower research costs, the enabling of better analysis and reporting and the facility to support the development of national priorities. The proposed NRII would comprise five elements. First, an information model would encapsulate the terms and definitions associated with all the main research information objects, describing how they relate to one another. This will enable the research community to

---

9 http://www.ref.ac.uk
10 http://www.hefce.ac.uk/pubs/rereports/Year/2015/metrictide/Title,104463,en.html
11 http://www.rcuk.ac.uk/media/news/151119/
14 https://core.ac.uk
15 https://jiscmonitor.jiscinvolve.org/wp/
16 http://gtr.rcuk.ac.uk
17 https://www.jisc.ac.uk/janet
18 http://www.jasmin.ac.uk
19 https://turing.ac.uk
20 http://equipment.data.ac.uk
develop a common language and will provide the foundation for the following four elements. Second, machine-readable access to existing information and systems will be enabled by a set of open protocols. These will also facilitate re-use of information and will help reduce the costs of collection, reformatting, integration and submission. Third, a national warehouse would store and provide open access to information likely to be submitted to the next REF while also becoming a permanent archival record of the state of the UK’s research, capable of being mined for useful insights. Fourth, the provision of capability for analysis and reporting will include a set of standard methods for interrogating information in the warehouse, providing thereby the foundation of an online national research analytics platform. Finally, a small expert team will maintain the infrastructure and develop and share best practice on the collection, storage, management and analysis of research information.

There will of course be challenges in terms of persuading people to agree protocols, getting the funding required to build the infrastructure and overcoming the fears of institutions and researchers about sharing data and deeper scrutiny. On the upside, however, government and funders should like the scalability of the solution as well as the provision of richer, more reliable information about the UK’s research base. Institutions would benefit from improved data quality and reliability, better benchmarking and business intelligence, simpler REF returns and a simpler policy compliance process. Institutions without a CRIS could use the NRII for that purpose and researchers are likely to welcome reduced administrative burdens and improved business intelligence to support career development.

For the NRII concept to become reality there is much groundwork to be done but to move the agenda forward three phases of development are envisioned. Phase one will focus on detailed analysis and planning to produce an appropriate design and identify existing and required standards. Reference would be made to national systems designed for similar purposes that are already in place such as the ANDS/RD Switchboard in Australia. Phase two would see the creation of the central Research Information System and Data Warehouse and during phase three NRII’s functionality will be enriched.

During the discussion period it was stated that the proposed cost of £4 million was an estimate at this stage but that a more detailed view of likely costs would be one of the outputs from the detailed planning process in phase one of development. In answer to a question about how to guarantee that a large national development would work, Jisc’s national analytics framework was referenced as a good parallel, one that provides confidence that the NRII can be a success with the help of the community. Finally, in line with the theme of transatlantic collaboration, the question was posed as to whether the UK and the USA could pursue collaborative development of the NRII, not least given that together both nations produce 40% of the world’s research output. Since the issues being tackled by NRII are rarely institution or country specific this could certainly be a catalyst for collaboration; the UK would be very interested in sharing technology, ideas and expertise.

---

21 [http://www.rd-switchboard.org](http://www.rd-switchboard.org)
Closing remarks

Clifford Lynch and Rachel Bruce, Deputy Chief Innovation Officer for Jisc, offered their observations and reflections to round off the day. There was a particular emphasis on the need for transparency not just in terms of data and metrics in particular but also the need for "accountable algorithms"; additional work is required to build on these themes of accountability and transparency. There is also a view that in terms of open access things have become too complicated and that a process of aligning workflows and policies across funders and countries would be a worthwhile and beneficial endeavour.

Turning to the situation in the USA it is thought that the way forward is not to think of directly supplanting the institutional investment but instead to think about infrastructure and add on services that lever institutional investments and make them more effective. On a note of caution many scholars do not feel strong institutional affiliation; they form teams with other scholars regardless of institution. In terms of NRII it is important to have a vision, something thought to be lacking in the USA at present. Finally, the theme of simplification has resonance. In the USA the funder compliance environment is complex; complex responses like SHARE exist but such responses can be burdensome and simplification for researchers would be welcome.
Jisc and CNI supplementary meeting

Following the main conference on 6th July a more focused meeting was convened on 7th July. The meeting was structured to consider four key questions around the scholarly record and its sustainability and how to enable the scholarly record for future research and impact.

Introduction

The meeting was opened by Rachel Bruce of Jisc who welcomed delegates to a day of concentrated discussion. It was noted that there are lots of forums in the area of digital scholarship but that the partnership between Jisc and the CNI allows for a sharper focus on the various issues of concern to a national infrastructure provider and an organisation that engages in new practice and policy around scholarly communications. In contrast to the previous day’s conference, the meeting was designed to facilitate discussion around four specific topics that are troubling to people on both sides on the Atlantic with the hope that interests will converge and stakeholders become mutually supportive. The first of two sessions was designed to look at the scholarly record and sustainability. There is a key distinction between STEM and some social sciences and humanistic disciplines that are not so chained to journals for research output dissemination. Although it seems that the paper form will persist there are fabulous examples of what can be done in the digital environment. The second session, in the afternoon, considers all the information in the scholarly record including research data, software and evidentiary material (newspapers, political advertising and so forth) and how this might be looked after for future research.

Key question 1

*For monographs and new types of long form publication, how far away are we from having infrastructure that reduces the risks to the persistence of the scholarly record? Where are the major challenges and opportunities?*

Don Waters provided insights based on his experience at the Andrew W. Mellon Foundation to seed the discussion. The system of monograph publishing is well entrenched. The relationship between author and university is critical with regard to tenure and promotion and the relationships between publisher and author and publisher and reader are also well established. While it is tempting to avoid disrupting this web of relationships things are changing, not least the fact that the academic library market for publishers is diminishing. Other changes, determined during an extensive consultative process, include the following: first, use of the digital medium is pervasive in the student and general audiences that faculty in the humanities are trying to reach; second, primary sources in the humanities are increasingly digitized or born digital; third, the use of digital methods and tools for analysing textual, audio visual and geospatial sources is spreading rapidly in humanities research and teaching. In conjunction with these changes, scholars in the humanities are having difficulty disseminating the results of their digitally based research to modern readers and getting the credit for it.

In response to these drivers the Foundation has made in the order of 30 grants over the past 18 months totalling USD 12.7 million. The main areas of focus have been to better understand the costs of monograph publications, to explore alternative business models and to develop guidelines for promotion and tenure. There is also a drive
to enhance university press infrastructure with an emphasis on shared services, digitally enhanced print publications and publications that can only be produced digitally. Evidence is emerging that the rate at which university presses are being reorganised under the university library is accelerating and that institutions are becoming increasingly involved in supporting faculty – as a community of authors – in terms of support for design, developmental editing and placing work with suitable publishers. The emphasis on support for developmental editing for digital works represents growing faculty awareness of the need to learn how to prepare works that address broad audiences at the intersection of the digital and public humanities.

It has been argued that in recent times there has been little understanding of the needs of readers with respect to research monographs. The recent structural alignment of libraries with university presses signals in part an attempt to address a fundamental question: are the discovery and other services provided by Amazon and Google sufficient for readers to find and use digital works? It is thought that the digital supply chain from academic press to reader is under-developed, especially for open access works, and requires investment in digital marketing. The library, however, has a pathway to the reader community and there is an increasing reliance on library-mediated discovery services. The Foundation has made a few minor grants to investigate the reader-related landscape in greater detail and to see what might be done to strengthen the services readers need.

Discussion

There is some caution about disrupting the traditional monograph publishing model as it can cause new problems. The Open Humanities Press22 for example is free but the “affordances” of the traditional system are lost – you can’t put an open access book onto Amazon, for instance. When it comes to the library supplanting the traditional roles of other stakeholders in the monograph publishing process it was noted that the library does not have any role in the tenure or promotion process. Despite such worries, there was support for the notion that the library is the right place to make monograph discovery and publishing happen. The potentially transformative effect of having universities serve as an author’s agent, realigning alliances and incentives, is seen to be appealing. It was suggested that this shift is not revolutionary: libraries have been working to move beyond being providers of content to helping to create content. Many libraries have staff whose role it is to work with faculty and students to create content.

Discussion about the role of copyright in the publication of new long form digital publications was wide ranging. In the USA Yale has determined that in most cases the use of an image embedded in a scholarly analysis transforms that image such that it cannot be seen to be destroying a market; rather, it exists in relationship with the scholarly text. If an image is under license the rights holder still needs to be paid but this is, nonetheless, a shift in presumption. The Internet Archive takes a similar approach in terms of presumption and it is thought that this development may lead to a general shift in peoples’ views of copyright in relation to digital monograph

22 http://openhumanitiespress.org
publishing. It was noted that a number of university libraries and museums in the USA have issued statements affirming that their materials can be freely reproduced especially in the context of scholarly communications, a move that begins to reduce uncertainty in the system. Some UK institutions have also done this including the Victoria & Albert museum in London. Work done by Jisc has highlighted that rights, particularly with reference to images, has been highlighted as being a major issue of concern in the UK. It was also pointed out that the Jisc survey on the use of monographs indicated that researchers are dissatisfied with electronic monographs and that people prefer print. Progress has stalled in the past five years; it was asserted that the Amazon ebook pdf model is not the way forward and the limitations of pdf might be the source of the reported dissatisfaction with electronic monographs.

On the subject of technological developments, some projects funded by the Andrew W. Mellon Foundation operate around the print based model but with digital enhancements. The University of Minnesota is working on a renewable book – one that is fixed at a certain point in time and then reworked at a later date in line with feedback. One view on future developments is that it is not useful necessarily to ask people what they want from new long form publications but that new ideas will emerge naturally and patterns will emerge over time. For example, the history department at the University of Oxford got their idea for new long form monographs from thinking about what they communicate via a website. It was suggested that peer reviewers be invited to review the technology being used to publish a monograph to try to move away from the bespoke nature of many new form digital monographs. These comments in combination with other discussions will be helpful in stimulating further thought about Jisc’s work in this area.²³

²³https://scholarlycommunications.jiscinvolve.org/wp/2016/01/13/open-access-books-jisc-work-in-the-area/
Key question 2

What roles do local, regional, national and global actors have in ensuring the persistent and reliable availability to researchers of the evidence base on which research relies? This evidence base includes both research data and digital material emanating from outside the sector. How are these roles defined and trust achieved for all?

Ensuring the persistence and reliability of the research evidence base can successfully be achieved by the research community itself, as Simon Coles, Associate Professor of Chemistry at the University of Southampton, explained. The crystallographic community is coherent, relatively small and acts as a body in its own right through a professional society, the International Union of Crystallographers (IUCr). In 1991 the community devised a format that described the data coming from their instruments and by 1992 nearly all the software researchers used was producing data in that format. Two years later the Crystallographic Information File format was being used for writing and publishing papers. The CCDC\(^{24}\) repository, a well-known resource in the world of chemistry, transformed its business to accommodate the new standard format since when the database has grown by several orders of magnitude. At present mainly results data are kept but the community recognises the need to also keep the raw data. The volumes of raw data are much greater so there is continuing discussion about how best to curate it.

Perhaps the key point to note is that scientists working in the discipline are part of the infrastructure. The data standard is maintained by the IUCr, which funds a committee for that purpose. The scientists involved with this form of maintenance get explicit recognition for that effort from both their own institution and the community.

In the wider realm of chemistry the picture is less rosy mainly because it is a less coherent community lacking a culture of developing and maintaining standards around the data they create; data is typically given to publishers. It was reported that chemists are disenchanted with the services provided by their institutions, arguing that their library doesn’t do much for them in terms of developing new ways of capturing and maintaining their research data outputs. Chemists want to comply with funder mandates so are currently looking for alternatives to simply giving outputs to publishers. This situation exemplifies the challenges facing researchers, individual institutions and national service providers.

To provide a more general perspective on the question, Cameron Neylon, Professor of Research Communications at Curtin University, presented a range of options for funding those parts of scholarly communications infrastructure that have historically been difficult to resource. First, groups with just a small number of members are able to survive financially but this solution is not scalable. Second, the membership model is commonly found in the scholarly communications space and can be successful (Jisc itself is currently moving to this model) though it is not universally applicable. Third, infrastructure can be funded by taxation – not necessarily taxation that takes the form of direct government funding but some form of “compelled payment”; it could be akin to the IUCr model where, for example, part of the Union members’ fees goes towards maintaining the Crystallographic Information File format. Although this is the model favoured by many, the

\(^{24}\) https://www.ccdc.cam.ac.uk/Community/initiatives/
problem is that taxation without representation creates difficulties and we do not currently have adequate governance structures or models in place to overcome those difficulties. Continuing the sustainability theme it was argued that grants are not helpful because of their short term and finite nature and that sustainable organisations that provide infrastructure should be built on a model of charging for services rather than content with the goal of achieving a financial surplus to avoid brittleness. The community needs to be able to hold such organisations to account.

Discussion

During the discussion there was general support for the taxation model. This is the funding mechanism generally employed for physical infrastructure (which are typically rivalrous because access is based on merit) so why not for more open forms of infrastructure? The sticking point may be that there are too many interested parties to get around a table to agree what is needed, at what price and how it should be governed. It was pointed out that it is hard to convince the public of the need to invest public funds in scholarly communications and that Jisc’s move from a taxation model to a membership model is evidence of this. Not everyone agrees central government taxation is an appropriate model on the basis that it has been shown to be an unsustainable and unreliable source. Library budgets, on the other hand, are relatively stable. On the subject of governance, is it possible to devise a governance model that will accommodate necessary competition for scarce funds? If just one solution is selected it will eventually become inefficient and, in any case, picking winners in the technology arena is extremely difficult to do. It was also mentioned that, particularly in the USA, the endowment model is both valid and useful. In reality solutions are messy and tend to be a mash up of the various different funding models highlighted in this session.

With a number of funders present it is unsurprising that there was debate on the subject of grants. Proponents argued that grants are usually given to recipients willing to take a chance on innovation (normally with terms requiring them to move towards sustainability, away from reliance on grants). It is thought that grant funding provides a useful element of competition for resources and it was said that valuable services generally do find a way to survive. In response, the audience was reminded of the demise of the Arts & Humanities Data Service in the UK. It was also argued that the grant system makes it hard to build interlocking infrastructure since projects tend to be one-offs. One suggestion is for a certain amount to be top-sliced from the “grant bucket” to fund infrastructure. Questions were raised as to why grant funders seem not to fund projects for the long term. In response it was said that grants are available for up to, say, 10 years to provide stability and that in the USA grants are available for the life cycle of large pieces of equipment – 20 years for a telescope, for instance. Foundations do provide ongoing support for some entities but they see their role primarily as one of providing seed capital rather than trying to pick winners and fund them in perpetuity.

There was some debate about what “sustainability” means when it comes to infrastructure. For some it means more than simply maintenance or breaking even in financial terms. It is continuous investment and development, retaining good staff by offering job security and providing the ability to grow and adapt to changes in the environment. It was also noted that the community needs to be more specific when talking about funding and sustainability, explicitly recognising disciplinary differences.
Finally there was a call for self-confidence and action. The people in the room and others not present have control over enough of the pieces of the puzzle that makes up the scholarly communications world to move to a maintaining phase rather than continuing to innovate. There is a lot of “plumbing” that requires maintenance such as registries, catalogues, protocols and routers. That is, it would be better to select some existing technologies and reduce the number of bespoke projects, thinking like a community rather than as smart individuals, using money collectively. This may involve concentrating on standards rather than trying to pick winners.

In his summing up of the morning’s discussions Clifford Lynch welcomed the insights and encouraged people to consider now how high-level ideas might be converted into specific projects that can be taken forward. There is a need to better understand where curation fits with respect to innovation and maintenance and a need to come up with projects that address issues around succession agreements and disaster recovery. It is desirable also to figure out what ideas or projects can be pruned if they are not looking promising. There needs to be more room for the smaller, facilitating projects that provide the “glue” that binds together infrastructural components and more work on standards is essential.

Key question 3

Extensive and reliable data are needed to underpin the responsible use of metrics for research and learning. These data arise as a result of a variety of interactions with the scholarly record by researchers, learners and others. Under what conditions are these data and associated algorithms available to the academic community? If the data and algorithms are held in “walled gardens” what challenges does that pose and how might they be addressed?

Introducing the afternoon session Clifford Lynch expressed the community’s interest in measures of impact that go beyond counting citations and made the case for strongly influencing movement toward open and transparent systems. In the USA development is very vendor-driven so a lock-in strategy prevails; there is limited discussion on cross-institutional initiatives. In stark contrast Jisc is leading a more inclusive approach.

On hand to outline Jisc’s approach, Phil Richards, Jisc’s Chief Innovation Officer talked about the work being done to keep Jisc at the forefront of research and teaching from a digital perspective. Continuing Lynch’s theme it was noted that citation metrics is a closed world and yet the influencing power of those metrics cannot be understated: citation-based metrics contribute 20% to a university’s position in the Times Higher Education world rankings for example.

Picking up the thread of the first day’s breakout session on research metrics and indicators, the REF evaluation process is likely to continue to use expert panels at a cost that some believe to be too great. It is hoped that the proposed National Research Information Infrastructure unveiled on the first day of the conference will help moderate the cost to universities of participation in the process. There is some dissatisfaction about the expert panels’ detailed analyses not being available for further analysis on the grounds of confidentiality but Jisc is looking to license similar information directly from ten universities that have a process of paying experts to score the papers they plan to submit to the REF. The plan is to see what value can be extracted from detailed analysis of the data.
Working with the Open University, Jisc has been funding the work of CORE as it creates an aggregation of open access research papers. Already CORE offers over 36 million records with four million full text items and work is under way to get open access papers funded by the gold route into the database. CORE is being positioned in the same space as Elsevier's SciVal and will become a useful resource for data mining. The team is developing a dashboard to enable users to see citation counts. Moving beyond the traditional approaches to using citations as a proxy measurement for quality, Jisc has been funding an experiment called Semantometrics to use the full text of journal articles to assess quality.

Turning from research to learning analytics, Jisc has been working with stakeholders to achieve consensus to build a national learning analytics architecture. There is recognition that a learning records warehouse, which sits in the cloud, will contain sensitive information about students and their activities so great care has been taken to ensure data security. It will very much be a “walled garden”. The plugin architecture is openly defined and allows different vendors to offer services. It is conceived that metrics derived from the system will permit accurate benchmarking and may also produce metrics for use by the forthcoming Teaching Excellence Framework (TEF). There are no plans to generate metrics for use in league tables since this would probably discourage universities from participating. In brief, the expected outcomes of Jisc's learning analytics system - making use of big data such as demographics data sets - will be a deep understanding of e-learning, metrics for engagement and learning gain and personalized, next generation e-learning content and delivery. In terms of additional value, using the Blockchain distributed database provides the opportunity to design currency for something of value such as badges for all forms of accreditation. Work on this is under way at the Open University.

Discussion

The ensuing discussion tackled the key areas covered in the presentation. There is a good level of interest in creating a public utility of bibliographic data upon which commercial organisations would be able to build services and which people would be able to mind and derive new datasets. There are already various places such as LOCKSS where there are large aggregations of bibliographic data but license restrictions against secondary use exist. One of the problems with such aggregations is knowing whether they are receiving or harvesting all the content they should have; experience at LOCKSS indicates this is not the case due to incomplete feeds from publishers. In addition metadata can be low quality or incomplete so it would be useful to be able to cross reference with other bibliographic databases. It was suggested that the community could approach publishers

---

25 https://core.ac.uk
26 https://en.wikipedia.org/wiki/Open_access_journal
27 https://www.elsevier.com/solutions/scival
28 http://semantometrics.org
31 https://www.lockss.org
to request greater access to reference data and that the development of a standard for citation data would be useful. There was debate about whether CORE has sufficient content yet to reliably produce bibliographic metrics. Currently CORE has 20% of the total number of records so is not yet at the point where it could be used as a substitute for commercial services. There is also a judgment to be made about how much effort should go into replicating traditional metrics rather than the development of new metrics that go beyond citation data.

Turning to learning analytics it was pointed out that in the USA universities are in tremendous competition with one another and that these competing interests are not necessarily in the best interests of research or learning. One outcome of this competition is that it is very difficult to fail poorly performing students because that affects a university's ranking. It was reported that the Gates Foundation runs a big grants program for personalised learning. There are concerns that specific funded initiatives are tied to commercial providers of learning resources; these providers are often start-ups so there are issues around their longevity. The program places no emphasis on Open Educational Resources (OER). The grants seem to be aiming to create a market. The concern with packaged learning systems is that it might make students think they don't have to or need not use products outside a narrowly prescribed set. Many believe higher education is about a broader outlook.
Key question 4

What are the major challenges in developing an integrated infrastructure to enable those subject to OA and research data policies to comply, and demonstrate compliance, with those policies? What steps might be taken next in addressing these challenges?

In the final session, MacKenzie Smith, University Librarian at the University of California, Davis, explained that the policy compliance situation in the USA is different to that in the UK and is much messier. There is a real need to focus on standards: some need to be built on; some are missing. The stakeholders in Open Access policy all have different perspectives on what needs to be done. Compliance is highly delegated to the organisation that owns the policy. The challenge is how to coordinate the chaos that is emerging. The chaos cannot necessarily be controlled but a better job can be done of setting the landscape, building the basic infrastructure and coming up with appropriate standards.

One of the biggest funders in the USA, the NIH, has a Public Access Compliance Monitor, a web based tool that institutions can use to track compliance of publications that fall under the NIH Public Access Policy. If you have an NIH grant they track it and they enforce it: they do not permit follow-on grants if articles derived from a previous grant are not publicly accessible. Other agencies are taking different approaches. Universities have their own infrastructure including institutional repositories. The unfortunate truth is, however, no university in the USA has a clue what their faculties have been doing. These data are locked up by faculties and will not be released. Some universities are experimenting with working with publishers to get copies of articles their faculty members have published but there is no serious attempt to coordinate this activity.

In terms of research data agencies are doing their own thing; NOAA and NASA have data repositories that they are trying to market to other agencies. The US Government has told NTIS\(^\text{32}\) that they will be the national data agency but since government does not fund them a cost recovery model must be used. In addition, there are disciplinary repositories and universities’ own repositories. On top of that, Elsevier has a new research data strategy department where integrated tools will be brought to bear on the whole data life cycle. There is so much going on that it feels paralyzing. Because researchers can choose to use any of a number of data storage options the problem becomes one of coordination. There is a need to accept and adapt to researchers’ behaviour. Better identifiers and standards would give libraries and their institutions a better chance to track the research outputs of their faculty members as well as monitor funder compliance.

Building on this theme Neil Jacobs of Jisc called for greater alignment of policies not least because researchers may have to comply with more than one policy at any one time, which can be complicated. Greater alignment will probably be difficult to achieve in the short term but an important first step would be to aim to have policies that adopt consistent vocabularies expressed in a standard way. This would enable the people charged with building infrastructure to encode policies in logic and software. There is a clear need to agree sets of persistent identifiers for funders, universities and other types of organisation. There are promising signs but contenders

\(^{32}\) http://www.ntis.gov/about/
such as ISNI\textsuperscript{33} and the Open Funder Registry\textsuperscript{34} provided by Crossref are getting insufficient traction. Jisc is trying to understand what the missing components are and is talking to existing players including Crossref to move things forward. Many actors have a stake in improving the information landscape: this is the moment to start collaborating.

Publishers use Crossref and deposit lots of information so Jisc and other organisations are keen to work with Crossref. Jisc will look to use its influence on behalf of the community not only to work towards improving the quality of metadata but also to promote the use of ORCID consistently. Agreeing best practice at the most basic, “plumbing” level is a small but important step.

**Discussion**

There was little optimism that agencies in the USA will agree on common areas with respect to policy or compliance; the key is getting conversations started around the lowest common denominators, echoing Jisc’s call for small but important steps to begin with. Pragmatism in this case means modularity and accepting that the community will need to be flexible in its approach. One suggestion was to look to the experience of the open source software arena. The Open Source Initiative\textsuperscript{35} catalogued the open source licenses and then selected to create a new license of their own. Perhaps the scholarly communications community could do the same in pushing for a smaller number of broadly acceptable licenses, perhaps breaking them down into components to make them more politically acceptable. It was noted, for interest, that few open access policies are explicitly openly licensed.

In the absence of watertight compliance systems it was suggested that a reasonable survival technique for the next few years would be to think less about a transactional deposit model and more about a messier model where content is copied between different repositories every few weeks to deflect part of the burden from the researcher. Whether funders would agree to this is another matter, particularly those that have invested in their own compliance systems. As a basic principle, however, most agree that asking researchers to enter metadata and deposit digital objects once should be sufficient.

In terms of moving the collaboration agenda forward, it was suggested that the SHARE team work with groups like CORE and OpenAIRE\textsuperscript{36}. There needs to be a focus on collecting better quality, more accurate metadata. While Jisc has convening powers to express the sector’s desire for certain things to publishers, the power of requests made to publishers would be significantly bolstered if there was concerted action with US-based actors such as ARL. Also, funders have more power that they might think to negotiate with publishers. Jisc has drafted a list of things the sector would like them to do, but working with counterparts from the USA to come up with a set of joint principles could be a very useful point of collaboration.

\textsuperscript{33} https://en.wikipedia.org/wiki/International_Standard_Name_Identifier
\textsuperscript{34} http://www.crossref.org/fundingdata/index.html
\textsuperscript{35} https://opensource.org/licenses
\textsuperscript{36} https://www.openaire.eu
Closing remarks

Rachel Bruce and Clifford Lynch rounded off the meeting, collaborating to provide some comments on the outcomes of the meeting. A lot of ground was covered during the day, some high level, some in great detail. There can be no tidy synthesis of the day but the outcomes point in several directions and these need to be followed up. The community needs to come together to reach out to publishers and government agencies to make the system incrementally run more smoothly. There are no overnight wins but work on identifiers and standards must be pushed through the pipeline. There was a lot of discussion around monographs and developmental editing. Authors will do what is easy for them so providing the right tools and systems to support them at scale is an important aspiration. There is a lot of folklore that is widely accepted such as the role of grant funding versus other types of funding, separating innovation from maintenance and trying to figure out when to start thinking about sustainability in the project life cycle: a short white paper may be helpful food for thought for funders and other stakeholders.

As people read through this report it is hoped there will be some part of the record of the presentations and discussions that will pique their professional interest and sow the seeds of fruitful collaboration which might be between individuals, individual organisations or national level organisations. Being alive to the opportunities for UK/US collaboration is the key as we all play our part in creating a bright future for all those who rely on an efficient scholarly communications infrastructure.