OPEN SCIENCE, OPEN DATA, OPEN ACCESS ...
A UKeiG WHITE PAPER
OPEN SCIENCE, OPEN DATA, OPEN ACCESS …
A UKeiG WHITE PAPER
DAVID BALL, DAVID BALL CONSULTING

Published by UKeiG, the UK eInformation Group of CILIP, the Chartered Institute of Library and Information Professionals
www.ukeig.org.uk
Published in October 2015
Series: UKeiG White Paper: 2015/01
© David Ball and UKeiG, 2015
Abstract

This White Paper gives an overview of developments in openness in Open Science, Open Data, Open Access and Open Monographs.

Open Science is shown to be moving centre-stage, with a rationale of improving efficiency in science; increasing transparency and quality in the research validation process; speeding the transfer of knowledge; increasing knowledge spill-overs to the economy; addressing global challenges more effectively; and promoting citizens’ engagement in science and research.

Open Data is shown to have undergone a surge in practical development, mirroring the well established repositories for research outputs. The development and application of model policies and of principles is also discussed.

The current major developments in Open Access are discussed in detail, including the identification and mirroring of success factors in funders’ and institutions’ policies and mandates for driving Open Access deposits and the growth in Gold Open Access.

The appearance of publishers and supporting infrastructure for the publication of Open Access scholarly monographs and textbooks is outlined.

Finally there is a prediction of the growth and influence of Open Science.
1. Open Science

1.1. Why Open Science?

The concept of Open Access (OA) to research outputs such as journal articles has been common currency for many years. The seminal Budapest Open Access Initiative (BOAI) statement of February 2002 (http://www.budapestopenaccessinitiative.org/read), for instance, reads:

By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

There have also been developments in the field of Open Data (OD).

More recent thinking, however, for instance by the European Commission, has expanded the concept of openness even further, to Open Science (OS), which aims to transform science by making research more open, global, collaborative, creative and closer to society (https://ec.europa.eu/digital-agenda/en/open-science). This shift is potentially very important for the development and exploitation of research.

OS is about the way research is carried out, disseminated, deployed and transformed by digital tools, networks and media. It relies on the combined effects of technological development and of cultural change in the direction of collaboration and openness in research.

To elaborate, a recent OECD report identifies the following six rationales for policies that seek to implement and support OS, including Open Data:

- **Improving efficiency in science** – OS can increase the effectiveness and productivity of the research system, by: reducing duplication and the costs of creating, transferring and re-using data; enabling more research on the same data; multiplying opportunities for domestic and global participation in the research process.
- **Increasing transparency and quality** in the research validation process, by allowing greater replication and validation of scientific results.
- **Speeding the transfer of knowledge** – OS can reduce delays in the re-use of the results of scientific research, including articles and data sets, and promote swifter development from research to innovation.

---

1. An alternative term sometimes used is Science 2.0. However a recent consultation by the European Commission published in February 2015 showed that Open Science is the preferred term (Validation of the results of the public consultation on Science 2.0: Science in Transition, available at: https://ec.europa.eu/research/consultations/science-2.0/consultation_en.htm).

• *Increasing knowledge spill-overs to the economy* – Increased access to the results of publicly funded research can foster spill-overs and boost innovation across the economy as well as increase awareness and conscious choices among consumers.

• *Addressing global challenges more effectively* – Global challenges require co-ordinated international actions. OS and Open Data can promote collaborative efforts and faster knowledge transfer for a better understanding of challenges such as climate change, and could help identify solutions.

• *Promoting citizens’ engagement* in science and research – OS and Open Data initiatives may promote awareness and trust in science among citizens. In some cases, greater citizen engagement may lead to active participation in scientific experiments and data collection.

Support for Open Science is widespread, not only among scientists but also among politicians and the public. Fecher and Friesike⁴ following an extensive literature survey, identify five schools of thought in promoting and supporting Open Science:

<table>
<thead>
<tr>
<th>School of thought</th>
<th>Central assumption</th>
<th>Involved groups</th>
<th>Central Aim</th>
<th>Tools &amp; Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democratic</td>
<td>The access to knowledge is unequally distributed.</td>
<td>Scientists, politicians, citizens</td>
<td>Making knowledge freely available for everyone.</td>
<td>Open access, intellectual property rights, Open data, Open code</td>
</tr>
<tr>
<td>Pragmatic</td>
<td>Knowledge-creation could be more efficient if scientists worked together.</td>
<td>Scientists</td>
<td>Opening up the process of knowledge creation.</td>
<td>Wisdom of the crowds, network effects, Open Data, Open Code</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Efficient research depends on the available tools and applications.</td>
<td>Scientists &amp; platform providers</td>
<td>Creating openly available platforms, tools and services for scientists.</td>
<td>Collaboration platforms and tools</td>
</tr>
<tr>
<td>Public</td>
<td>Science needs to be made accessible to the public.</td>
<td>Scientists &amp; citizens</td>
<td>Making science accessible for citizens.</td>
<td>Citizen Science, Science PR, Science Blogging</td>
</tr>
<tr>
<td>Measurement</td>
<td>Scientific contributions today need alternative impact measurements.</td>
<td>Scientists &amp; politicians</td>
<td>Developing an alternative metric system for scientific impact.</td>
<td>Altmetrics, peer review, citation, impact factors</td>
</tr>
</tbody>
</table>

---

It is evident from the table that the central aims, and of course the tools and methods to achieve them, are quite diverse. This diversity may well be a strength: many different objectives supporting different interests can be achieved through a single, albeit complex, mechanism.

1.2 What it is
Each step of the research lifecycle is becoming more open\(^4\), for instance through:

- **Open Notebooks** - an emerging practice, documenting and sharing the experimental process of trial and error;
- **Open Data** - managing research data in a way that optimises access, discoverability and sharing for use and re-use;
- **Open Research Software** - documenting research code and routines, and making them freely accessible and available for collaboration;
- **Open Access** - making all published outputs freely accessible for maximum use and impact.

In order to achieve this openness in science, each element of the research process should:

- **Be publicly available** - it is difficult to benefit and use knowledge hidden behind username and password barriers, or if it does not contain the right metadata to make it discoverable.
- **Be re-usable** - research outputs must be licensed appropriately so that prospective users know clearly any limitations on re-use.
- **Induce collaboration** between researchers through better access and better online tools;
- **Be transparent and have appropriate metadata** to provide clear statements of how research output was produced, and can be re-used.

A more concrete exposition of Open Science and its many branches is provided by the taxonomy developed by the European FP7 FOSTER project in support of its aim of putting in place “sustainable mechanisms for EU researchers to FOSTER OPEN SCIENCE in their daily workflow, thus supporting researchers optimizing their research visibility and impact, the adoption of EU open access policies” (https://www.fosteropenscience.eu/about). The taxonomy covers not only the constituent elements of OS but also supporting tools, measurements and mechanisms.

\(^4\) See the FOSTER course Open Science to Scientific Research (in preparation, to be available through https://www.fosteropenscience.eu/project/).
1.3 How open is our research?

Further evidence of this widening from Open Access and Open Data to Open Science is provided by the development by SPARC Europe of a tool for visualising, discussing and monitoring how open an institution’s research is (http://sparceurope.org/how-open-is-our-research-a-checklist-for-institutions/) It takes the form of a radar diagram generated by confirming status or actions, or estimating percentages, in 11 main topic areas:

The 11 topic areas are exhaustive, but worth quoting in full because the checklist demonstrates the potential extent and complexity of Open Science, and the challenges faced by institutions and funders in bringing it about:

- **Policies and strategies – external**: signatory of/adherence to global/international statements (e.g. BOAI); compliance with supra-national policies (e.g. Horizon 2020); compliance with national policies (e.g. funders’) and domain policies.

- **Policies and strategies – internal**: policies that support open data deposit; support Green OA; support Gold OA; support re-use/CC licensing; are published; are regularly reviewed and updated/revised as appropriate.

- **Text repository**: open repository, listed in ROAR (http://roar.eprints.org/); mandates for various outputs (ROARMAP - http://roarmap.eprints.org/); incentives in place to encourage self-archiving; import of SHERPA RoMEO (http://www.sherpa.ac.uk/romeo/) data; discoverability; metrics services, e.g. repository usage indicators.

- **Data repository**: open repository; data management plan and preservation policy; mandate for data deposit; incentives in place to encourage self-archiving; discoverability; usage metrics.

- **Open educational resources (OER)**: OER repository; mandate for deposit; incentives in place to encourage deposit; discoverability; percentage of courses that are OA.

- **Open research**: policy in place; open notebooks; open source software; open standards; involvement of public as appropriate.
Open Science, Open Data, Open Access … - David Ball Consulting

- **Publishing**: annual percentage of the range of outputs that are OA; provision of tools for scholar-led publishing initiatives (e.g. OJS software); university press or library campus-based publishing operation.

- **Licensing**: open licensing policy; institutional websites appropriately licensed; institutional research outputs, teaching materials etc. appropriately licensed.

- **Institutional programme on openness**: included in all library induction programmes for the range of research-connected staff; included in information literacy materials.

- **Archives**: digitisation programme; preservation policy in place; discoverability; metadata included in Archive Hub; appropriate viewing mechanisms for digital materials.

- **Culture, mission, practice**: public commitment to openness in mission statement; designated ‘open champion’ in a significant position of influence; community outreach; digest and presentation of research to media; promotion of resources via popular websites and social media.

Completing the checklist for an imaginary institution produces the following visualisation of the openness of its research:

Such diagrams can be used in a number of ways, for instance as an assessment tool, to generate discussion, to inform policy-making.

2. **Open Data**

As we noted at the start of this document, while Open Access to research outputs has a long history and development, Open Data have come into scope somewhat later. The OECD report already quoted makes the rationale specific:

5: reducing duplication and the costs of creating, transferring and re-using data; enabling more research on the same data; ... increasing transparency and quality in the research validation process, by allowing greater replication and validation of scientific results.

Research data can be defined simply as whatever is either produced in research or evidences research outputs.

---

The European Commission’s definition is: “information, in particular facts or numbers, collected to be examined and considered and as a basis for reasoning, discussion, or calculation”. Examples include: statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings, images.\(^6\)

The 2012 European Commission recommendation on access to and preservation of scientific information states that: “open access to scientific research data enhances data quality, reduces the need for duplication of research, speeds up scientific progress and helps to combat scientific fraud”.\(^7\) Unsurprisingly other funders also require open access to the data produced as a result of their funding. The Wellcome Trust for instance has been a leader in the field. Its *Policy on data management and sharing* (2010) states: “The Wellcome Trust expects all of its funded researchers to maximise the availability of research data with as few restrictions as possible” [http://www.wellcome.ac.uk/about-us/policy/policy-and-position-statements/wtx035043.htm](http://www.wellcome.ac.uk/about-us/policy/policy-and-position-statements/wtx035043.htm).

Over time policies have developed. Commonly they will now include the following elements\(^8\):

- **Timing**: when publication should take place;
- **Data plan**: requirements for a technical management plan;
- **Access and Sharing**: what exactly will need to be available for public use;
- **Long term curation**: data creation and sustainability;
- **Monitoring**: any monitoring that will be carried out by the funding body and guidance available;
- **Storage**: details of the appropriate repository or data centre to be used;
- **Costs**: where costs can be claimed from and when.

Making data open is in some, particularly technical, senses more complex than making research outputs open: data collected must be capable of being verified, processed and re-used. However there are many resources covering all aspects of Open Data policies and practice now made available for instance by the Digital Curation Centre ([http://www.dcc.ac.uk/resources/policy-and-legal/policy-tools-and-guidance/policy-tools-and-guidance](http://www.dcc.ac.uk/resources/policy-and-legal/policy-tools-and-guidance/policy-tools-and-guidance)).

Research Councils UK (RCUK), on behalf of the UK Open Research Data Forum, has also recently (July 2015) published a draft *Concordat On Open Research Data*, which is meant to:

[help] ensure that the research data gathered and generated by members of the UK research community is made openly available for use by others wherever possible in a manner consistent with relevant legal, ethical and regulatory frameworks and norms. ... The intention [of the Concordat] is to establish sound principles which respect the needs of all parties. It is not the intention to mandate, codify or require specific activities, but to establish a set of expectations of good practice with the intention of establishing open research data as the desired position for publicly-funded research over the long-term.

---


The development and promulgation of such principles is welcome. However there have been some criticisms (which may be held generally applicable to the current stage of development of Open Data), to the effect that the Concordat concentrates more on the sciences than on the arts and humanities (https://unlockingresearch.blog.lib.cam.ac.uk/?p=285).

3. **Open Access**

3.1 *Some definitions*

One of the early and seminal definitions of Open Access (OA) occurs in the statement of the Budapest Open Access Initiative (BOAI) quoted at the start of this document, which was first published in 2002, then re-affirmed ten years later in 2012. The statement is long and radical. A more concise definition is provided by Peter Suber: Open Access literature is “digital, online, free of charge, and free of most copyright and licensing restrictions”.

These definitions are in contrast to the traditional method of obtaining access to research outputs: *Toll Access* (TA), by means of institutional or personal subscription to journals, or aggregations of content, or by means of paying publishers for access to individual articles.

Intellectual property laws in the UK and elsewhere offer limited “fair dealing” or “fair use” exemptions. With OA a distinction is made between *Gratis* and *Libre*. *Gratis* OA is free of charge to access but subject to the limits of fair dealing; it removes toll barriers but not permission barriers. *Libre* OA is both free of charge and free of at least some legal and licensing restrictions; it removes toll barriers and at least some permission barriers; the BOAI definition is libre.

3.2 *Green OA*

A further distinction is made between *Green* and *Gold* OA, Green being long established, Gold being more recent.

*Green OA* is delivered through self-archiving: authors deposit manuscripts in institutional or disciplinary repositories. It relies on the relatively recent but well established infrastructure of institutional and subject repositories. The repositories can be found through the Registry if Open Access Repositories (ROAR - http://roar.eprints.org/), which aims “to promote the development of open access by providing timely information about the growth and status of repositories throughout the world”. In October 2015 it listed 4053 repositories, the majority being institutional or departmental. It is fully searchable and browsable.

Green OA is easy and cheap: each article incurs only a very small portion of the overhead costs of setting up and running repositories (estimated by Swan at between £6 and £15 per article). It does not incur the overheads of peer-review, yet deposited articles may be, most often have been, peer-reviewed for publication in TA journals. It is compatible with subscription journal publishing: scholars can publish in TA journals and, through self-archiving, still make their articles OA. Generally an embargo period, of 6 or 12 months, is imposed by publishers, and authors generally need to obtain rights from publishers to deposit and make articles available. Finally, it is hospitable to many other types of document, notably theses and multiple-media offerings.

As we shall see later, current thinking on maximising the deposit of OA papers is that they should be deposited on acceptance by a publisher. This practice does not contravene publishers’ embargo

---


requirements. The deposit step is a separate action from making an article openly available and the publisher has no sanction over it. The aim is to get authors to deposit their articles as they are accepted for publication, which is the moment they are dealing with the paper for the last time in practical terms. So long as a paper is deposited, the author, and support staff, need not worry about it any longer: if it is under a publisher’s embargo the repository software automatically opens the article and makes it public at the end of the embargo period.\textsuperscript{11}

3.3 Gold OA

Gold OA has been delivered mainly through journals: these may be completely OA or hybrid, where some articles are OA and others toll access. Articles, in both OA and hybrid journals, are paid for by the authors or their institutions or funders. Articles are peer-reviewed for publication; thus Gold OA may incur much the same costs for the editorial and peer review process as TA journal publishing. Gold OA is always immediate, while Green OA is often subject to time embargoes imposed by subscription journal publishers. It provides access to the published version of an article, while Green OA generally provides access only to the author’s final peer-reviewed manuscript, without the formatting or pagination of the published version. By its nature Gold OA is confined to post-prints and generally obtains rights and permissions direct from the rights-holder (usually the author). Both Green and Gold OA are gratis. Green OA generally is only gratis; Gold OA may be libre.

We noted above that Green OA offers a cheap and easy route, based on the peer review practice of traditional TA journals. Gold OA is more problematic. Like TA publishing, it has to bear the costs of editorial boards, peer review, production and marketing. However its costs are met not through TA’s well established market route of subscriptions, which are now generally paid at the wholesale level of the Big Deals, but generally through individual article processing costs (APCs). Payment and administration by institutions are therefore at a very granular level, especially when compared with the Big Deals.

Despite these complications and difficulties, and perhaps a general reluctance to adopt new means of publishing, there is evidence of a continuing rise in Gold OA. In 2011 a study by Mikael Laakso et al.\textsuperscript{12} (p.8) identified three earlier cycles in the development of Open Access publishing: the “Pioneering Years” (1993 to 1999), the “Innovation Years” (2000 to 2004), and the “Consolidation Years” (2005 to 2009).

The Pioneering Years were characterised by innovation by individuals or small groups of scholars, using simple technologies. There was rapid growth from, obviously, a small base: in 1993, it is estimated that 20 open access journals published 247 articles; by 2000, 741 journals are estimated to have published 35,519 articles. Many of these early journals did not survive.

The Innovation Years coincided with the wholesale movement of journal content to electronic delivery. In terms of OA they were characterised by burgeoning advocacy of OA and the development of economic models for Gold OA, notably article processing costs (APCs). BioMed Central and PLoS demonstrated the viability and high quality of Gold OA. There was significant


growth of both titles and articles: by 2005, 2,837 journals published 90,720 articles, an increase of 155% on 2000.

The *Consolidation Years* saw the growth of infrastructure to support OA, such as open source publishing software, the DOAJ, Creative Commons licences. Discovery was enhanced and enabled by Google and Google Scholar. Growth was not as spectacular, but still very strong: in 2009 4,767 journals published 191,851 articles, an increase of 111% on 2005. One might add that the Consolidation Years also saw the adoption by funders of policies on deposit and public availability of the results of the research they fund. The first was the Wellcome Trust, followed by the National Institutes of Health and many others.

In October 2015 the Directory of Open Access Journals (DOAJ - [https://doaj.org/](https://doaj.org/)), “an online directory that indexes and provides access to high quality, open access, peer-reviewed journals” recorded 10,624 Gold OA journals from 135 different countries, comprising 2,104,841 articles. There has therefore been a doubling in the number of OA journals in the six years since 2009.

The DOAJ has recently been used by Walt Crawford\(^\text{13}\) to analyse the development of Gold OA. He offers the following conclusions and observations:

- Gold OA constitutes a significant portion of scholarly publishing – at least 20%, possibly more;
- Gold OA is growing, but not (generally) very rapidly;
- Subject areas differ hugely, so it is very difficult to generalise;
- Every subject area demonstrates a significant number of free OA journals publishing a significant number of articles;
- There are subject areas in which there is no significant amount of APC-charging OA activity (library science, for example), but not many;
- No-fee publishing is declining in some subjects but by no means all.

We may therefore conclude that there is significant and healthy growth in Gold OA publishing, and that it is increasingly generating income.

### 3.4 Fostering growth in OA: Policies

The most important factor in driving the growth of OA is the policies adopted by funders and individual institutions. Recognising this, the European Commission’s FP7 Project PASTEUR4OA ([http://www.pasteur4oa.eu/](http://www.pasteur4oa.eu/)) “aims to support the European Commission’s Recommendation to Member States of July 2012 that they develop and implement policies to ensure Open Access to all outputs from publicly-funded research”. As part of the work of PASTEUR4OA, the database of Open Access policies, ROARMAP ([http://roarmap.eprints.org/](http://roarmap.eprints.org/)), was extended and elaborated. It now records, and links to, every known policy’s conditions under an exhaustive set of categories, and is fully searchable. This database as a whole provides a rich source of data to analyse when studying policy effectiveness.

The project also looked at the mandatory policies in place at over 120 universities around the world and assessed the effectiveness of each policy. This was measured in terms of the percentage of

---

Open Access material available from each institution compared to the total number of articles published from those institutions each year.

Using regression analysis, the project determined that the important elements of a policy, whether of a funder or an institution, are as follows:

- The policy states that research articles must be deposited in a repository (that is, the policy is mandatory);
- The policy states that this action cannot be waived: that is, whatever the conditions of embargo, the article must be deposited at the point specified by the policy;
- If the policy states that an author should retain certain rights over the published work, this action is mandatory and cannot be waived;
- The policy states that deposited items must be made Open Access, and if there is an embargo then they must be made Open Access immediately the embargo comes to an end;
- The policy links the deposit of articles with research assessment/performance evaluation procedures within the institution: that is, the policy states that articles that are not deposited in line with policy requirements will not count towards performance reviews or research assessment exercises.

The critical elements of a policy are:

- The policy requires that research articles be deposited in an Open Access repository;
- In addition, the policy must state that this deposit step cannot be waived;
- The policy links deposit with research assessment (performance evaluation).

The first two elements in the list above are significantly correlated with resulting high levels of Open Access and, of course, they make the policy a mandatory one. All three of these policy elements are significantly associated with success.

<table>
<thead>
<tr>
<th>Policy element</th>
<th>Positive correlation</th>
<th>Significant correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articles must be deposited</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Deposit cannot be waived</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Deposit of articles is linked to research evaluation (performance assessment)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Articles must be made Open Access</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Where the policy stipulates that authors retain certain rights, this cannot be waived</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

A policy that includes the elements above and is implemented properly by funders and institutions will succeed in gathering a large volume of Open Access content. The requirement to deposit, and

---

the insistence that this step cannot be waived for any reason, ensure that authors deposit their work.

There is evidence from the PASTEUR4OA project to show that the adoption of strong policies by funders drive the adoption of policies, particularly aligned policies, in institutions. Most importantly the policy for the European Commission’s Horizon 2020 research funding programme is also of this type, meaning that institutions making this type of policy are aligning their own policy with that of the European funding programme. This is important, as researchers within the institution may be funded under this programme and will therefore have the agreeable experience of their funder’s and institution’s policies have matching requirements, making it simple to comply with both through one set of actions.

An excellent example of a policy linking deposit of articles to research evaluation is provided by HEFCE (Higher Education Funding Council for England). The policy insists that “to be eligible for submission to the post-2014 REF [the next research assessment exercise], authors’ outputs must have been deposited in an institutional or subject repository” (http://www.hefce.ac.uk/pubs/Year/2014/201407/). Deposit must also take place on acceptance by a publisher. There is evidence from individual institutions that this policy is already having the effect of increasing the number and proportion of OA deposits. At UCL (University College Lindon) for instance the repository contained 10,000 OA outputs in 2011 and 14,000 OA papers in 2013; OA content then sharply increased to 22,500 papers by September 2015.

There is as yet little evidence of funders preferring Gold OA over Green OA. At most their policies are neutral, but they do state that Gold OA publishing costs may be included in bids for funding. Obviously there will be a considerable delay in such Gold OA articles being published – generally only after completion of the research project. There is however also a current European Commission initiative: a pilot, run by OpenAIRE, to provide retrospective Gold OA funding for FP7 projects. In summary according to the website (https://www.openaire.eu/postgrantopenaccess):

- This post-grant Open Access Pilot provides an additional instrument to make FP7 project research results Open Access. It does not affect authors' freedom to choose which way their project publications should be made Open Access.
- The post-grant Open Access Pilot will cover OA Article Processing Charges (APCs) for FP7 projects up to two years after they end.

3.5 Fostering growth in OA: Practice

Library staff and repository managers have a vital role to play, both in supporting or prompting the creation and adoption of policies, and in ensuring the deposit of items. A range of pointers to good and effective practice was produced by a recent survey of the managers of 24 of the most successful repositories in a range of countries.

Of the respondents 83% have a formalised institutional strategic plan or co-ordinated approach for open access; a plan is in preparation for a further 8%. 58% have a formal preference for Green OA; none has a preference for Gold OA; the rest express no formal preference.

---

Three quarters of institutions have key stakeholders as champions; for 56% this is a formal arrangement. Most used are researchers (89%), followed by librarians (78%) and senior management (56%). In one institution this is an integral part of the Associate Dean role.

Repositories are served by between 3 and 6 FTE staff in 80% of institutions. Most have a full-time repository manager, with varying types and levels of support. Two institutions have integrated repository work with other responsibilities: one might see this as a continuing trend as the novelty of IRs declines.

Turning to training, 40% of institutions include OA in induction for staff and students; in all responding institutions it is given by library staff. The three most common events provided (by 90% of respondents) are: formal presentations, individual one-to-one training, website based instructions; 50-60% offer briefing sessions and interactive face-to-face presentations; 2 respondents also mention help desks. The most successful offerings are interactive face-to-face presentations and individual one-to-one training (both at 50%). Least successful offerings are: formal presentations and website based instruction.

One institution commented:
The approach has tended to be more integrated than one off events - we might not have an "open access" day but we would be part of a Faculty awayday where open access would be discussed alongside research strategy and impact... We will participate in other disciplinary events that are researcher led e.g., publishers coming to speak like RSC or funder training days e.g. AHRC. We look to embed into the research environment where possible.

3.6 Indicators of the Transition to OA

Responding to a recommendation of the Finch Report, the Research Information Network has just published (in August 2015) a report for the Universities UK Open Access Co-ordination Group entitled Monitoring the Transition to Open Access. The findings represent a first attempt at generating such indicators covering five sets of issues:

- **OA options available to authors**: the numbers of fully-OA and hybrid journals, along with issues such as the level of article processing charges (APCs), the availability of CC-BY and other licences, and the length of embargo periods;
- **Accessibility: authors’ take-up of OA options**: the numbers - and the proportions of the overall population – of articles accessible on OA terms via different routes;
- **Usage**: the levels of usage of OA articles as compared to those that are not accessible on OA terms;
- **Financial sustainability for universities**: the amounts paid by UK universities in subscriptions and in APCs; and

---


• Financial sustainability for learned societies: the overall income and expenditure – as well as the volumes of journal-related income and expenditure – of UK learned societies which have some publishing income.

This is a preliminary but very worthwhile investigation of how we might monitor the transition to OA.

4. OA Monographs

Scholarly monographs, the staple of subjects in the arts and humanities, have lagged behind journals. However, with a typical print-run of 200-250 and low financial returns for all involved in traditional publication of scholarly monographs, OA is an obvious avenue to be tested and developed.

Gold OA monographs are starting to appear, through publishers such as Ubiquity (http://www.ubiquitypress.com/) and Open Book Publishers (http://www.openbookpublishers.com/), especially in the humanities and social sciences.

Various business models are being trialled. One of the most interesting, pioneered by Knowledge Unlatched (http://www.knowledgeunlatched.org/), is a form of crowd funding:

• The Knowledge Unlatched model depends on many libraries from around the world sharing the payment of a single Title Fee to a publisher, in return for a book being made available on a Creative Commons licence.
• The Title Fee represents the basic cost of publishing a book. Because the Title Fee is a fixed amount, as more libraries participate in Knowledge Unlatched, the per-library cost of ‘unlatching’ each title declines.
• Access to the Title Fee allows publishers to feel confident that they will not make a loss on a title if it is made Open Access. Publishers offer libraries deductions if another format is also purchased to thank libraries for their support.
• Once KU has reached scale, this model is expected to be financially self-sustaining: the costs of operating Knowledge Unlatched will be covered by a small service charge added to the Title Fee to cover KU’s administrative costs.

Services to support OA publication are also starting to appear. One such is OAPEN (http://www.oapen.org/home):

OAPEN provides a platform for the full text dissemination of Open Access books and provides services to publishers and libraries. OAPEN builds a quality controlled collection of Open Access books, mainly in the area of Humanities and Social Sciences, and develops services for publishers, libraries and research funders in the areas of dissemination, quality assurance and digital preservation.

OAPEN also takes the lead in developing common standards and has over 30 publishers in membership. It is responsible for the Directory of Open Access Books (DOAB - http://doabooks.org/doab?uiLanguage=en), which lists 3380 Academic peer-reviewed books from 121 publishers.
In the US, the Library Publishing Coalition (http://educopia.org/lpc/index.php/Main_Page#Library_Publishing_Coalition_Project) is seeking to build capacity and experience for libraries and university presses to become or improve as publishers. Not all the outputs will necessarily be OA, nor all of them monographs. However this is a significant initiative, showing a clear direction of travel away from the commercial publishers.

Tackling a very problematic area of textbooks, another significant US initiative is State University New York’s (http://textbooks.opensuny.org/) publication of 15 OA e-textbooks, with more forthcoming.

In the UK Jisc Collections (https://www.jisc-collections.ac.uk/Institution-as-E-textbook-Publisher/) is funding a project to produce a range of e-textbooks. The fundamental question addressed is: “Will the institution as e-textbook creator help students by providing a more affordable higher education, and promote a better, more sustainable information environment for libraries, students and faculty?” Four project teams, The University of Liverpool, The University of Nottingham, The University of the Highlands & Islands with Edinburgh Napier University and University College London, are working to: a) create their own e-textbooks; b) apply a business, licensing and distribution model; and c) report back on the impact, value and viability of the models they chose. The textbooks and reports on business models are due in 2017.

5. The Future

This White Paper has given a necessarily incomplete overview of developments in openness in Open Science, Open Data, Open Access and Open Monographs.

Open Science is moving centre-stage, with a rationale of improving efficiency in science; increasing transparency and quality in the research validation process; speeding the transfer of knowledge; increasing knowledge spill-overs to the economy; addressing global challenges more effectively; and promoting citizens’ engagement in science and research.

We are already seeing the development of supporting infrastructure, such as the FOSTER taxonomy and the visualisation tool from SPARC Europe with its exhaustive eleven criteria.

Turning to Open Data, we have seen a surge in practical development, mirroring the well established repositories for research outputs. We are now also seeing the development and application of model policies and of principles.

Open Access is long established, but there are currently major developments, in terms for instance of the identification and mirroring of success factors in funders’ and institutions’ policies and mandates for driving Open Access deposits. We are also seeing steady growth in Gold OA at least in a number of subject fields, despite the historical difficulties associated with business models and general acceptance by the research community. Indicators to monitor the transition to OA are also in development.

Finally there is the appearance of publishers and supporting infrastructure for OA scholarly monographs.
It is very easy to become blinded by the interesting detail of these advances. However it is salutary to paraphrase Bill Clinton’s mantra on the economy: “it’s the research, stupid”. It is researchers themselves, funders, governments, supra-national bodies such as the European Commission and industry and commerce that will benefit directly from this openness. The benefits are potentially, huge, multiplying the return on investment in research, accelerating the research process and involving a full range of interested citizens.

We can already see new paradigms and structures arising. As information professionals we are already closely involved in their development. We must be seen to be leading the move to the new pervasive openness.