Building a global knowledge commons through open access and open science

Kathleen Shearer
Executive Director, COAR
kathleen.shearer@coar-repositories.org
@Kathlee.shearer
COAR?

• An international association founded in 2009
• Office is based in Göttingen, Germany
• Members & Partners: over 120 institutions from 35 countries in Australia, Africa, Asia, Europe, North and South America
• Institutional membership fees: EUR 500
• CARL was a founding member
• ? Members from Asian region
Vision

A global knowledge commons based on a network of open access repositories
Who is COAR?

- Over 100 members and partners from 35 countries in 5 continents
- Universities, libraries, government agencies, open access organizations, not-for-profit organizations, and platform developers
- Diverse perspectives that share a common vision

Contacts Us

http://www.coar-repositories.org
Email: office@coar-repositories.org
Phone: + 49 551 39 22215
Fax: + 49 551 39 5222
Facebook: COAR_eV
Twitter: @COAR_eV

How to participate?

- Organizations can join COAR for €500 Euros per year (about $600 US)
- Join as a single, consortial, or special member or partner
- Download the membership application (https://www.coar-repositories.org/about/join/become-a-member)

Major Activities

- International voice
  Raising the visibility of repository networks as key infrastructure for open science
- Alignment and interoperability
  Building a global knowledge commons through harmonization of standards and practices
- Cultivating relationships
  Supporting an international community of practice for repositories and open access
- Building capacity
  Advancing skills and competencies for repository and research data management
- Adopting value-added services
  Promoting the use of web-friendly technologies and new functionalities for repositories

Positioning Asia in the Global Movement of Open Science
November 14 - 15, 2016  Kuala Lumpur, Malaysia
Asia OA is a special forum hosted by COAR in which members of the Asian open access community can share information, meet each other and build relationships. It has a mailing list and organizes meetings to facilitate greater exchange beyond national boundaries.

This community is dedicated to people working in the academic environment based in the Asian region. It celebrates Asian cultural diversity and unique way of doing things.

To join the mailing list, send an email to: office@coar-repositories.org
What is open science?

• Began with open access to publications, but moving to data and other types of research outputs
• Parallels to Open Government/Open Data movement

Open science involves:

• Sharing and access to all types of research outputs
• Transparency of research findings
• Open peer review & open citations
• Equitable flow of knowledge
What is driving this trend?

1. Verification, reproducibility and transparency of scientific results
2. New scientific discoveries through re-use and integration of datasets
3. Greater social and economic benefits through application of research outputs
4. And because we can...

Eg. Canada’s Open Government Action Plan

Positioning Asia in the Global Movement of Open Science
November 14 - 15, 2016  Kuala Lumpur, Malaysia
Open science is a global trend

Open science: a hot issue for OECD and non-OECD countries

Number of countries reporting that the situation has recently substantially changed in the policy area, compared with other STI policy areas or instruments

Note: Simple counts do not account for the magnitude and impact of policy changes.
Source: Country responses to the STI Outlook policy questionnaire 2014.
Current priorities for open science:

- Publications (open access)
- Research data (open data)
The international publishing system is broken!
The access problem

### TABLE 1: AVERAGE 2015 PRICE FOR SCIENTIFIC DISCIPLINES

<table>
<thead>
<tr>
<th>DISCIPLINE</th>
<th>AVERAGE PRICE PER TITLE</th>
<th>DISCIPLINE</th>
<th>AVERAGE PRICE PER TITLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemistry</td>
<td>$4,871</td>
<td>Zoology</td>
<td>$2,073</td>
</tr>
<tr>
<td>Physics</td>
<td>4,341</td>
<td>Technology</td>
<td>2,058</td>
</tr>
<tr>
<td>Engineering</td>
<td>3,039</td>
<td>Math &amp; Computer Science</td>
<td>1,866</td>
</tr>
<tr>
<td>Biology</td>
<td>2,977</td>
<td>Health Sciences</td>
<td>1,694</td>
</tr>
<tr>
<td>Astronomy</td>
<td>2,730</td>
<td>General Science</td>
<td>1,643</td>
</tr>
<tr>
<td>Food Science</td>
<td>2,496</td>
<td>Agriculture</td>
<td>1,589</td>
</tr>
<tr>
<td>Botany</td>
<td>2,277</td>
<td>Geography</td>
<td>1,571</td>
</tr>
<tr>
<td>Geology</td>
<td>2,195</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SOURCE: LJ PERIODICALS PRICE SURVEY 2015
The participation problem

World scaled by number of documents in Web of Science by Authors Living There

Created by @juancommander using d3.js and cartogram.js
The pressure to publish in "luxury" journals encouraged researchers to cut corners and pursue trendy fields of science instead of doing more important work.
We are all complicit!

10 simple strategies to increase the impact factor of your publication

by sven | Mar 5, 2015 | |

Impact factors are heavily criticized as measures of scientific quality. However, they still dominate every discussion about scientific excellence. They are still used to select candidates for positions as PhD student, postdoc and academic staff, to promote professors and to select grant proposals for funding. As a consequence, researchers tend to adapt their publication strategy to avoid negative impact on their careers. Until alternative methods to measure excellence are established, young researchers have to learn the “rules of the game”.

Positioning Asia in the Global Movement of Open Science
November 14 - 15, 2016  Kuala Lumpur, Malaysia
The Oligopoly of Academic Publishers in the Digital Era

Vincent Larivière, Stefanie Haustein, Philippe Mongeon

Published: June 10, 2015 • DOI: 10.1371/journal.pone.0127502

Abstract

The consolidation of the scientific publishing industry has been the topic of much debate within and outside the scientific community, especially in relation to major publishers’ high profit margins. However, the share of scientific output published in the journals of these major publishers, as well as its evolution over time and across various disciplines, has not yet been analyzed. This paper provides such analysis, based on 45 million documents indexed in the Web of Science over the period 1973-2013. It shows that in both natural and medical sciences (NMS) and social sciences and humanities (SSH), Reed-Elsevier, Wiley-Blackwell, Springer, and Taylor & Francis increased their share of the published output, especially since the advent of the digital era (mid-1990s). Combined, the top five most prolific publishers account for more than 50% of all papers published in 2013. Disciplines of the social sciences have the highest level of concentration (70% of papers from the top five publishers), while the humanities have remained relatively independent (20% from top five publishers). NMS disciplines are in
Open access has arrived!
The two roads to open access

Open access journals

- Journals without subscriptions
- A variety of business models
- The large publishers are using APCs

Open access repositories

- More than 3000 based at institutions around the world
- Repository networks (e.g. NII, OpenAIRE, LA Referencia, SHARE, etc.)
- Embargoes
Smooth transition?

Photo credit: Roy Gumple (www.allposters.com)
The solutions of the major publishers: APCs (Article Processing Charges)

ELSEVIER
Open access publication fee
A fee is payable by the author, or their institution or funder to cover the publication costs. Fees range from $500 to $5,000 US Dollars. Visit your journal's homepages for specific pricing information.


Open Choice allows you to publish open access in the majority of Springer's subscription-based journals.
The price to publish open access

The initial wide variety in APC prices and their general convergence shows that APC prices are not grounded in the actual cost of producing an article but rather are reflections of what the market can bear (Lawson, “APC Pricing”, 2014). The report’s estimate of £1,500-£2,000 may have encouraged cheaper journals to raise prices in order to be seen as high quality.

“Article processing charges (APCs) and subscriptions. Monitoring open access costs”, Jisc. United Kingdom. Junio 2016 https://www.jisc.ac.uk/reports/apcs-and-subscriptions

£0.75 GBP = $1 US
Average price = $2334.58 US
In dramatic statement, European leaders call for ‘immediate’ open access to all scientific papers by 2020
Another option...

Strengthen and add value to our local journals and repository networks
#9: Local infrastructure that is sustainable and inclusive

(Open, distributed systems, like the Internet, are more flexible, sustainable and less likely to failure or being bought out by commercial industry)
The vast majority of open access policies are green

OA Policy Requirements - Pasteur4OA Project (European Commission)

<table>
<thead>
<tr>
<th>Criterion (Green OA)</th>
<th>Number of policies</th>
<th>Criterion (Gold OA)</th>
<th>Number of policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deposit in repository required (Green OA)</td>
<td>381</td>
<td>OA publishing required</td>
<td>2</td>
</tr>
<tr>
<td>Deposit in repository requested</td>
<td>140</td>
<td>Recommended alternative to Green OA</td>
<td>97</td>
</tr>
<tr>
<td>Deposit in repository not specified</td>
<td>141</td>
<td>Permitted alternative to Green OA</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not specified/other</td>
<td>463</td>
</tr>
<tr>
<td>Total</td>
<td>663</td>
<td></td>
<td>663</td>
</tr>
</tbody>
</table>

Table 3: Open Access policies: Green and Gold OA criteria
Current state of repositories internationally

Proportion of Repositories by Continent - Worldwide

- Europe: 45.6%
- Asia: 19.9%
- North America: 18.3%
- South America: 8.6%
- Africa: 4.7%
- Australasia: 2.7%
- Caribbean: 1.6%
- Central America: 1.5%
- Other: 0.8%

Total = 3189 repositories

OpenDOAR - 30-Aug-2016
Dual mission of repositories

1. Provide access to intellectual outputs of the institution, to the local community and to the world
2. Contribute as a node in a global knowledge commons
An institutional repository is/are…

“a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members.” (Cliff Lynch 2002)

“digital collections capturing and preserving the intellectual output of a single or multi-university community” (Raym Crow 2002)
Science is increasingly global!
The current role of repository networks

• Harvest from repositories
• Define community policies, practices, and standards
• Facilitate a community of practice
• Curate and transform metadata
• Track open access publications
• Create a brand for repositories
Figure 26.7: Scientific publication trends in Malaysia, 2005–2014

Malaysian publications have grown rapidly since 2005, overtaking those of similarly populated Romania

<table>
<thead>
<tr>
<th>Year</th>
<th>Malaysia</th>
<th>Romania</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>1,559</td>
<td>2,543</td>
</tr>
<tr>
<td>2006</td>
<td>1,813</td>
<td>2,934</td>
</tr>
<tr>
<td>2007</td>
<td>2,225</td>
<td>3,983</td>
</tr>
<tr>
<td>2008</td>
<td>2,852</td>
<td>5,165</td>
</tr>
<tr>
<td>2009</td>
<td></td>
<td>4,266</td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td>5,777</td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td>6,485</td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td>6,657</td>
</tr>
<tr>
<td>2013</td>
<td>8,925</td>
<td>7,550</td>
</tr>
<tr>
<td>2014</td>
<td>9,998</td>
<td>6,651</td>
</tr>
</tbody>
</table>

- **0.83** Average citation rate for Malaysian publications, 2008–2012; the OECD average is 1.08; the G20 average is 1.02
- **8.4%** Share of Malaysian papers among 10% most-cited, 2008–2012; the OECD average is 11.1%; the G20 average is 10.2%
- **46.4%** Share of Malaysian papers with foreign co-authors 2008–2014; the OECD average is 29.4%; the G20 average is 24.6%
Aligning repository networks

1. Strategic coordination
To have a shared vision and a stronger voice for the repository community internationally

2. Data exchange
To demonstrate that we are building truly global services!

3. Harmonization and standardization
To support the development of value added services across region
Latin America (LA Referencia)

North America (SHARE)

Europe (OpenAIRE)

Japan (NII)

China (CAS)

Asia

Africa

Australia/Pacific
COAR Controlled Vocabularies: #1 resource type (currently in English, Chinese, Dutch, French, Italian Portuguese, Russian, Spanish, Turkish)  
Version 1.0 available on the COAR website  
https://www.coar-repositories.org/activities/repository-interoperability/ig-controlled-vocabularies-for-repository-assets/deliverables/  
Version 1.1 will be available soon
In their current form, repositories only perpetuate the flawed system

“What if we don’t change at all ... and something magical just happens?”
COAR Working Group, Next Generation Repositories

Eloy Rodrigues, chair (COAR, Portugal)
Andrea Bollini (CINECA, Italy)
Alberto Cabezas (LA Referencia, Chile)
Donatella Castelli (OpenAIRE/CNR, Italy)
Les Carr (Southampton University, UK)
Leslie Chan (University of Toronto, Canada)
Rick Johnson (SHARE/University of Notre Dame)
Paolo Manghi (CNR, Italy)
Lazarus Matizirofa (NRF, South Africa)
Pandelis Perakakis (Open Scholar, Spain)
Oya Rieger (Cornell University, US)
Jochen Schirrwagen (University of Bielefeld, Germany)
Daisy Selematsela (NRF, South Africa)
Kathleen Shearer (COAR, Canada)
Tim Smith (CERN, Switzerland)
Herbert Van de Sompel (Los Alamos National Laboratory, US)
Paul Wain (EDINA, UK)
David Wilcox (Duraspace/Fedora, Canada)
Kazu Yamaji (National Institute of Informatics, Japan)
Next generation repositories

To position repositories as the foundation for a distributed, globally networked infrastructure for scholarly communication _on top of which layers of value added services will be deployed,

_thereby transforming the system, making it more research-centric, open to and supportive of innovation,

_while also collectively managed by the scholarly community.
Two central ideas for actualize our vision:

1. Improve the functionality of repositorios
   - To be of, not just on the web
   - Global interoperability (exposing content in a standardized way)
   - Pro-active repositories
Two central ideas for actualize our vision:

2. Support and encourage the development of value added services
   - Comentaries and peer-review
   - Usage measures
Our vision involves more than just articles

All the valuable products of research should be shared!
EUROPEAN OPEN SCIENCE CLOUD
BRINGING TOGETHER CURRENT AND FUTURE DATA INFRASTRUCTURES

A trusted, open environment for sharing scientific data

Open and seamless services to analyse and reuse research data

Linking data

Connecting across borders and scientific disciplines

Connecting scientists globally

Improving science

Long term and sustainable

EUROPEAN DATA INFRASTRUCTURE
UNLOCKING THE VALUE OF BIG DATA; DIGITAL BY DEFAULT
Research Data Management is like a three-legged stool...

Infrastructure and services

Culture

Policies

Positioning Asia in the Global Movement of Open Science
November 14 - 15, 2016  Kuala Lumpur, Malaysia
Funders’ Data Policies

United Kingdom (all funders have data policies)
And United States, European Commission, Canada, Finland... and a few others.
## Typical Elements of a Data Policy

<table>
<thead>
<tr>
<th>Policy requirements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Data quality and standards</td>
<td>Investigators are required to adhere to international standards to enable access and reuse in the discipline.</td>
</tr>
<tr>
<td></td>
<td>Data documentation and metadata must accompany data so that the data is understandable by others.</td>
</tr>
<tr>
<td>Data access and sharing</td>
<td>Investigators are required to make data available to be shared (usually upon publication of results or shortly thereafter, although some agencies do allow embargo periods).</td>
</tr>
<tr>
<td></td>
<td>Requirements for deposit of metadata into a local or national catalogue</td>
</tr>
<tr>
<td>Data retention and preservation</td>
<td>Data should be retained for a certain time limit, where possible, investigators must deposit their data in a long-term archive to ensure the preservation of their data.</td>
</tr>
<tr>
<td>Data management plans</td>
<td>Research proposals must include a Data Management Plan in proposal.</td>
</tr>
</tbody>
</table>
### Exceptions

#### Common exceptions to policies

<table>
<thead>
<tr>
<th>Category</th>
<th>Exception</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy</td>
<td>The rights and privacy of individuals who participate in research must be protected at all times. Thus, data intended for broader use should be free of identifiers that would permit linkages to individual research participants and variables that could lead to deductive disclosure of the identity of individual subjects.</td>
</tr>
<tr>
<td>Traditional knowledge</td>
<td>Where local and traditional knowledge is concerned, rights of the knowledge holders shall not be compromised.</td>
</tr>
<tr>
<td>Data of a sensitive nature</td>
<td>Where data release may cause harm, specific aspects of the data may need to be kept protected (for example, locations of nests of endangered birds or locations of sacred sites).</td>
</tr>
<tr>
<td>Intellectual property/Data ownership</td>
<td>It may be necessary on occasion to delay publication for a short period to allow time for applications to be drafted.</td>
</tr>
</tbody>
</table>
A lot of this is about cultural change
Many researchers would rather share their toothbrush than their data...
Data sharing practices

From Wiley’s Research Data Insights Survey, 2014
2,250 responses from around the world
Both policy requirements and incentives are critical for cultural change.
Infrastructure and Services

Data sharing requires good management across the data lifecycle.
The data landscape

The 2011 survey by *Science*, found that 48.3% of respondents were working with datasets that were less than 1GB in size and over half of those polled store their data only in their laboratories. *Science* 11 February 2011: Vol. 331 no. 6018 pp. 692-693 DOI: 10.1126/science.331.6018.692
These services still only support a small portion of the research datasets produced by researchers around the world!
Concluding thoughts

• Repositories are a technology, and technologies change
• What we are really promoting is a vision in which institutions, universities, and their libraries are the foundational nodes in a global scholarly communication system
• We can do this by leveraging, expand and enhancing the already globally connected international repository network
• We need to start now with a shared vision
terima kasih!