Universities and knowledge sharing: The Curtin Open Knowledge Initiative

Associate Professor Lucy Montgomery
Centre for Culture and Technology, Curtin University
ELPub conference, 4 June 2019
Introductions
Introductions – COKI

• COKI: The Curtin Open Knowledge Initiative
• Housed within the Centre for Culture and Technology
• In collaboration with the Curtin Institute for Computation
• Strategic funding from Curtin University
• Multidisciplinary
What does it mean for a university to be an ‘Open Knowledge Institution’?

What kind of information might universities need to understand how they are tracking?

Can publicly available data help?
Introductions – The COKI Team

Cameron Neylon, Professor of Research Communications
Lucy Montgomery, Associate Professor, Internet Studies
John Hartley, Professor of Cultural Science
Chun-Kai (Karl) Huang, Data Scientist
Katie Wilson, Research Fellow
Richard Hosking, Data Scientist
Alkim Ozaygen, Data Scientist
Chloe Brookes-Kenworthy, Data Wrangler
Emma de Francisco, Project Coordinator
‘This book advocates for universities to become Open Knowledge Institutions which institutionalise our world’s creative diversity to contribute to the stock of human knowledge’

Lucy Montgomery, John Hartley, Cameron Neylon, Malcolm Gillies, Eve Gray, Carsten Herrman-Pillath, Chun-Kai (Karl) Huang, Joan Leach, Jason Potts, Xiang Ren, Katherine Skinner, Cassidy Sugimoto, Katie Wilson
Introductions - COKI

• Open Access and open data are important elements of vibrant knowledge systems
• But they are not enough

Successful OKIs also involve

• Diversity – who gets to make knowledge?
• Cross boundary collaboration: disciplines + university/wider community
Can publicly available data help us to begin mapping:

- Intentions – Institutional policy documents, stated aspirations
- Effort – Investment and resource allocation
- Outcomes - Evaluation
<table>
<thead>
<tr>
<th>Aspiration (Policy/Narrative)</th>
<th>Action (Investment)</th>
<th>Outcomes (Evaluation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>What aren’t we doing?</td>
<td>Where do we choose to invest?</td>
<td>What do we do well?</td>
</tr>
<tr>
<td>What are our comparators doing?</td>
<td>Can we find synergies?</td>
<td>How are we different?</td>
</tr>
</tbody>
</table>

- **Diversity**
- **Communication**
- **Coordination**

**Deficit Models**

**Investment/Priority Models**

**Capacity Models**

**Impacts**
Some numbers:
~1.5 Billion records
~2 terabytes
~1000’s of primary sources
(not including snapshot, backups & temporary workflow steps, which is significant)

Person hours of effort:
• Core team @ Curtin
• Collaborators in UK, USA, Netherlands, Australia and NZ
• Data Scientists, Software engineers and Qualitative researchers working together

Computational usage:
• 500,826 function invocations
• 5,578,408.86 GHz-second
• 4,578,284.74 gibibyte second
• 36.96 gibibyte of log volume
• 867,936 Firestore DB writes
• 9,824,690 Firestore DB reads
• Multiple tebibyte’s of data processed in BigQuery analytical queries
How ‘open’ are campuses?
Library access policies and practices

- General public least likely to have access
- Subscription costs a factor in capacity of libraries to accommodate non-core communities. But...
- Limited co-ordination between OA and physical access policies
Spearman rank correlation:
%OA vs OA policy $\rho = 0.5$
%OA vs access policy $\rho = -0.33$
Access to academic libraries: an indicator of openness?

Katie Wilson, Cameron Neylon, Lucy Montgomery, and Chun-Kai (Karl) Huang

Introduction. Open access to digital research output is increasing, but academic library policies can place restrictions on public access to libraries. This paper reports on a preliminary study to investigate the correlation between academic library access policies and institutional positions of openness to knowledge.

Method. This primarily qualitative study used document and data analysis to examine the content of library access or use policies of twelve academic institutions in eight countries. The outcomes were statistically correlated with institutional open access publication policies and practices.

Analysis. We used an automated search tool together with manual searching to retrieve Web-based library access policies, then categorised and counted the levels and conditions of public access. We compared scores for institutional library access feature with open access features and percentages of open access publications.

Results. Academic library policies may suggest open public access but multi-layered user categories, privileges and fees charged can inhibit such access, with disparities in openness emerging between library policies and institutional open access policies.

Conclusion. As open access publishing options and mandates expand, physical entry and access to print and electronic resources in academic libraries is contracting. This conflicts with global library and information commitments to open access to knowledge.

Introduction

In recent years global library and information communities have reaffirmed commitments to providing access to information and knowledge. The International Federation of Library Associations and Institutions (IFLA) adopted the theme Libraries Driving Access to Knowledge from 2009-2011 (Lau, Tammaro and Bothma, 2012), confirming IFLA support for the open access movement. IFLA advocated for the United Nations Sustainable Development programme to incorporate access to information, and information and communication technologies, in achieving the UN sustainable goals. In 2015 the United Nations member states made access to information a specific target under Goal 16 of the 2030 Agenda for the UN Sustainable Development Goals. Libraries provide and promote access for people to ‘the information they need to live, learn, create, and innovate’ (International Federation of Library Associations and Institutions, 2017, p.
Who is Making Knowledge?
Also -

• Gathering data about linguistic and cultural diversity
• Working with international collaborators to identify locally relevant measures of diversity and inclusivity
• Analyzing institutional policies (or lack of)
How are we tracking with OA?
Methodology

- Scopus
- Web of Science
- Microsoft Academic

DOIs

- Unpaywall
- Crossref

- Open Citations
- DOAB
- Worldcat
- ORCID
Challenges

Incomplete and inconsistent databases

Limited to DOIs

DOIs

Grid ID

OA status

Year

DOI

Events

Incomplete Coverage

Unpaywall

Crossref

Open Citations

DOAB

ORCID

Worldcat
Outline of results: Australia & New Zealand

Graph showing trends over years for various categories related to open access.

- Black line: output_openaccess_coki_percent_total_oa
- Orange line: output_openaccess_coki_percent_gold_doaj
- Red line: output_openaccess_coki_percent_hybrid_any_license
- Green line: output_openaccess_coki_percent_green
- Filled green: output_openaccess_coki_percent_green_home_repo

Graphs for Australia and New Zealand separately, showing trends from 2000 to 2015.

Bar chart showing percentages for AU university, Australia, New Zealand, and Global leader.

- Bronze %
- Hybrid Any License %
- Gold Doaj %
- Green Only %

Curtin University logo present at the bottom right.
Outline of results: Australia & New Zealand
Data Framework – Exploration and Communication

Dashboards - DataStudio

Jupyter Notebooks and Python

BigQuery – SQL + ML
**Future Directions – NLP, Graphs, Viz**

**The HIVE @ Curtin**
Prototyping new means of communicating the ‘Big Picture’
A timeline of research and research outputs in Australia

* [www.frankenplace.com](http://www.frankenplace.com)*
* Topic Modelling, Named Entity Recognition
  And Explorative search

* [https://www.canterbury.ac.nz/science/contact-us/people/ben-adams.html](https://www.canterbury.ac.nz/science/contact-us/people/ben-adams.html)*
Next Steps

- Quality assurance and coalition building
- Dashboard
- Libraries in Australia and New Zealand invited to review their own data
- August validation workshop:
  - Key players in open science and scientometrics from Europe, South America, Africa, the US and Australia
  - Stress-test our approach and methodology
Thank you

Questions?