Universities and knowledge sharing

Evaluating progress to openness at the institutional level

Lucy Montgomery, Curtin University, Australialucy.montgomery@curtin.edu.au, Cameron Neylon, Richard Hosking, Chun-Kai (Karl) Huang, Alkim Ozaygen and Katie Wilson

Introduction

1   Following Open Access (OA) declarations in the twenty-first century such as the Budapest Open Access Initiative, Bethesda Statement on Open Access Publishing and Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, the goals of OA and the sharing of knowledge have advanced through the establishment of institutional policies and repositories for research output, funder and government mandates and the growth of OA publishing platforms. However, OA publishing and sharing practices are being adopted unevenly: within institutions, as well as across countries and regions. Distribution of OA research outputs is dominated by transatlantic, European flows of knowledge with much lower movements towards and among African, Asian, Middle Eastern, Latin and South American countries. Further, declarations and mandates are only a first step. Worldwide rankings and evaluations that currently dominate the higher education landscape provide little insight into the nature of the open knowledge challenge now facing the sector. Established rankings also provide few clues as to how institutions might most effectively navigate the challenges and opportunities of open knowledge. Over the past decade, the rhetoric of open has shifted from the fringes of science, education and communication discourses, to the centre. To a greater or lesser extent, it is now to be found in the language and policies of research and research funding globally. Openness has become a hallmark of Good Science. New tools are needed to help governments, research funders, universities and researchers to understand their place within open knowledge landscapes, and to identify their challenges and successes in becoming Open Knowledge Institutions.

2   To support institutions to make real change we need to measure not just words and policies (intentions) but effort (investment and resource allocation) and outcomes

(evaluation). This involves extended analysis of institutional OA publishing output performance, research collaboration, and diversity in research output and staffing. Communication to the public including actions such as OA output and research data sharing contribute to the whole institutional profile. Diversity in research production and output contributes to knowledge diversity. Staffing diversity in terms of gender, ethnicity, disability and age enables open, diverse perspectives in teaching and research. Coordination of policy, communication and evaluation actions, and of interaction between diverse groups is critical to achieving the objectives of the Curtin Open Knowledge Initiative.

Key objectives of the project

3   The Curtin Open Knowledge Initiative is a research project within the Centre for Culture and Technology at Curtin University. With funding from Curtin University, we are exploring the extent to which universities are functioning as effective open knowledge institutions; as well as the types of information that universities, funders, and communities might need to understand an institution's open knowledge performance and how it might be improved.

4   We define an Open Knowledge Institution as an institution that does more than simply support or mandate specific practices. An effective Open Knowledge Institution embodies core values that deliver the benefits of open science culture and practice. It achieves this through providing an environment, platforms and culture that deliver and also hold in tension three key areas: **communication, diversity** and **coordination**. Cultural change at institutional level and in response to national or regional initiatives and mandates is fundamental to achieving openness. Where policies and practices manage and determine the sharing of research knowledge, collaboration and diversity, we can explore and analyse the extent of, and possible ways to enable openness. Through a theory of change we aim to identify institutional progress from aspirations (storytelling and narrative as well as policies) through investment and resource allocation to specific initiatives and on to delivering integrated practice and performance. Table 1 represents this process schematically.

Table 1: A schematic of the Curtin Open Knowledge Initiative project. Institutions can assess their positioning and progress through evaluating outcomes and measures.

	Aspiration (Policy/Narrative)	Action (Investment)	Outcomes (Evaluation)
Communication	• OA Policy • Data management policy • Public engagement/ communications policy • Communication in core documents	• Investment in repositories • Staffing for data management and/or OA • Funded support or resourcing for engagement and communications	• % OA outputs • Data shared and archived • Public access of outputs • Public engagement

Diversity and Equity	• Diversity and inclusion policy • Policy on communications and evaluation (output diversity) • Public engagement/ communications policy	• Engagement with diversity programs • Staffing and support • Training programs • Interdisciplinary programs	• Staff/student diversity • URM retention statistics • Diversity of revenue sources • Disciplinary diversity of research outputs
Coordination	• Library access policies • Campus planning and public access • Policy support for coordination and/or community building	• Investment in public transport integration/civil provisioning • Support for public events	• Attendance at public events • Collaboration measures • Public engagement/ citizen science measures

5   The Curtin Open Knowledge Initiative project is examining a range of dimensions and indicators of openness with the ultimate goal of developing new strategies to help institutions to identify, understand, extend and diversify their own open knowledge achievements. A key first project output is a collaborative OA book by thirteen international scholars - *Open Knowledge Institutions: Reinventing Universities* (Montgomery et al, 2018). The book is a manifesto for Open Knowledge Institutions, outlining the philosophy and principles for institutional openness and exploring the links between ideas of open science and open scholarship, diversity (of people, ideas and topics), and the coordination required to bring these ideas effectively together. The idea of an Open Knowledge Institution is a way of understanding the tensions between "open" and "closed". The internet has shifted the balance between control and disorder in knowledge production and the book discusses the opportunities the web brings directly to many of the challenges to be addressed in this contested space.

Design and methodology

6   Building on the foundations established in the book *Open Knowledge Institutions: Reinventing Universities,* the project is gathering and analysing a range of data and information from institutions and sources around the world to identify how different paths to openness can be achieved. Key to this is the identification of common indicators which can be used to evaluate how institutions are achieving Open Knowledge. The methodology aims to compare universities through analysis of a range of established datasets and institutional documents.

7   We are expanding our knowledge of OA output by harvesting data entries from bibliographic and media sources on a large scale to build a database of research outcomes, providing detailed data for universities since 2000. In parallel we analyse OA policies and practices that express intentions and aspirations in relation to the publishing and availability of research output and the performance of institutional repositories. For this study, following Piwowar et al. (2018), we define OA output as that which is '*free to*

read online, either via the publisher website or in an OA repository.' Searching bibliographic sources Microsoft Academic, Scopus, Web of Science and UnPaywall we are obtaining DOIs representing research outputs for a set of institutions and identifying the distribution of types of OA found in institutional repositories. We have identified OA format types and analyse institutional output by such format types as a further indication of research diversity. Using Crossref funder data we are also analysing indicators of OA mandate compliance at an institutional level. Analysis of CrossRef event data shows media attention to OA publications and research output including patents. Examining the datasets of research output, we can identify collaboration networks among institutions to provide indications of the geographic spread and impacts of research over a period of time.

8   In addition to the OA data we are examining institutional intentions in the context of regional and local environments through policies from university website documents, and compliance with funders' OA mandates. Using narrative text analysis we are developing ways of understanding institutional positions on and commitments to diversity, equity and inclusion as broader dimensions of open approaches to knowledge production. This includes exploring data and indicators to understand practices and performance of universities' collaboration and engagement with the communities they form part of, locally and regionally. Through analysis of word occurrence and frequency in institutional annual reports we can discover aspiration and intention in relation to diversity and equality. Analysis of higher education statistics to understand staff characteristics such as gender and ethnic distribution provides evidence of outcomes and enables comparison with the intentions expressed in annual reports and policies.

9   In order to explore the potential of library access policies to act as a proxy for openness we use qualitative document and data analysis methods to review academic library access and library use policies to understand how they can facilitate knowledge access and privileges for non-institutional or unaffiliated library users. We gather institutional information about open access and open data policies and practices such as repositories, correlating library access with open access. To facilitate these processes, we developed a Jupyter notebook instrument to automate the search, retrieval and downloading of library access and use policy and membership documents from the internet. Sources include university websites and databases of repositories such as the Directory of Open Access Repositories (OpenDOAR) and The Registry of Open Access Repository Mandates and Policies (ROARMap).

Open access performance

10   We have analysed institutional OA data in terms of a body of DOIs of research output, OA format types, CrossRef event data showing attention to OA publications, and research collaboration among institutions. The OA analysis shows a breakdown into percentages of Gold (found in journals registered with DOAJ), Green (freely accessible in repositories), Hybrid (articles accessible at publisher websites in non-DOAJ journals with a clear access license (of any kind), Bronze (articles accessible at publisher websites for which a clear access license cannot be found. Micro-analysis identifies further nuances within these OA types, such as percentages of Green submitted, Green published, Green in institutional repository, Green only (all objects found only in a repository including author accepted

manuscript (AAM), submitted and published versions), Green AAM, Gold DOAJ, Hybrid any licence and Hybrid CC licence.

11 Extracting details of bibliographic formats (eg, article, book chapter, book, conference paper, report, reference entry, monograph) we can identify institutional research diversity as an indicator of openness. For example, Figure 1 shows the extent of OA bibliographic output diversity for one institution over the period 2000-2017.

Figure 1: Numbers of OA research output by format for one university
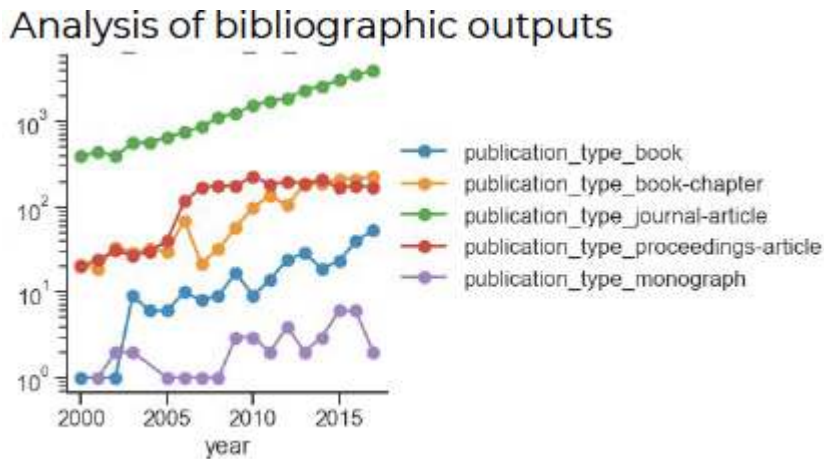
12 Analysis and cross-referencing of OA output across institutions enables the tracking of OA performance across regions. For example, Figures 2 and 3 show the variations in percentages of Green and Gold OA output in 13 geographic subregions and regions for the year 2017. The Americas and Asia have higher Gold OA while Green is predominant in the European regions, reflecting the different policy emphases and practices in these regions. In particular, the Gold OA in South America is driven by the collections of OA journals published and indexed in the Scientific Electronic Library Online (SciELO), established in 1998, representing research output from Latin America, the Caribbean and South Africa since 2009.

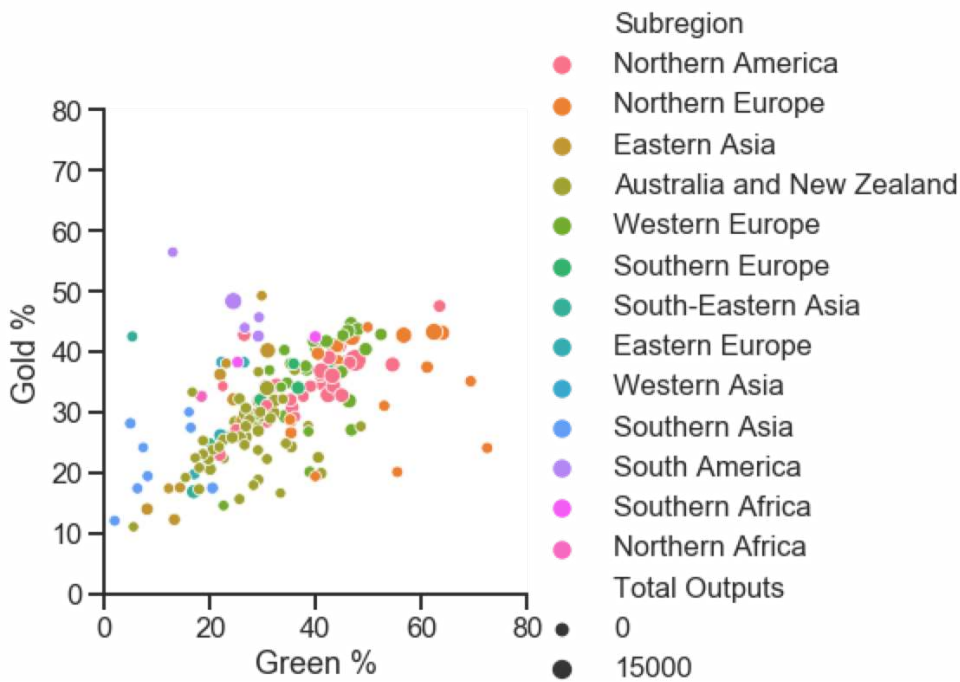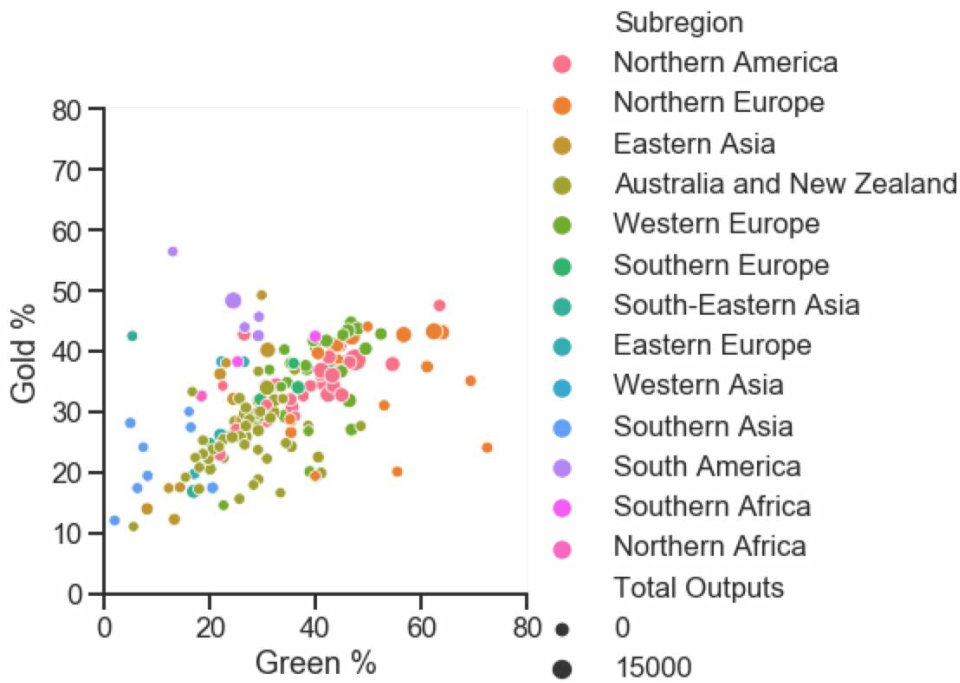**Figure 2: Gold vs Green OA output by sub-region**

Figure 3: Percentage of Gold vs Green OA output by region 2017. Europe shows a steady, large growth in Green output.

13 Through analysis of CrossRef Event data (CrossRef, 2017) using its API, we are able to identify institutions' research performance, impact and engagement in mainstream and social media. Crossref tracks research output with DOIs in sources such as Twitter, Reddit, Wordpress, blogs; citations from patents in Cambia Lens, an open patent and scholarly search service; annotations in Hypothes.is; newsfeeds, scholarly literature, research

recommendations in F1000, question and answer websites and Wikipedia. Comments, links, shares, bookmarks and references to output with DOIs are tracked. The event data are available from 2015 and include output with earlier publications.

14   Analysis of author affiliations in open access publications provides information on the distributions of collaboration among universities through geographical research networks as a further indicator of research openness and diversity. Figure 4 shows national collaboration among Australian universities.

Figure 4. A chord diagram showing the intensity of collaborations between Australian institutions in 2017. The width of the connecting lines shows the number of collaborative publications between universities.
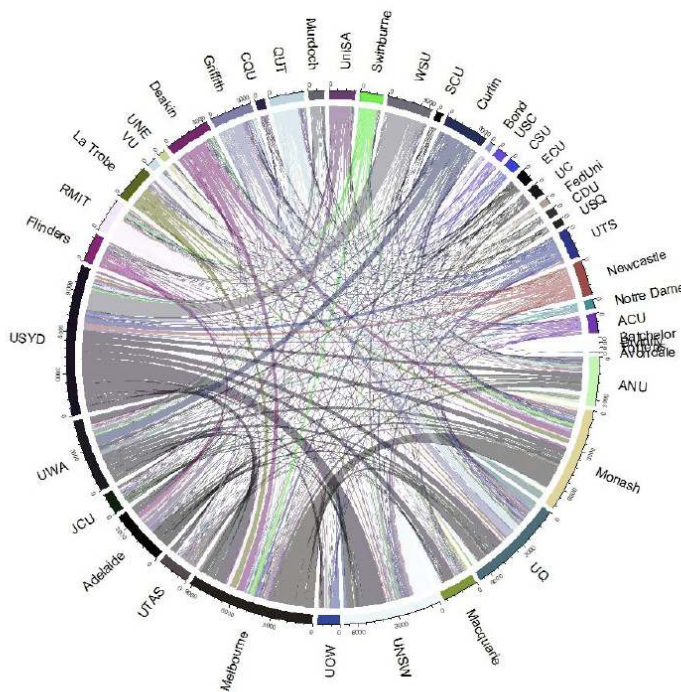
Staffing diversity

15   To understand institutional openness in terms of staff diversity (gender, disability, underrepresented minorities and ethnicity) we analyse statistics and relate to policy and legislation. Staff gender, indigenous, nationality, citizenship and ethnic percentages calculated from higher education statistics demonstrate real outcomes in response to institutional policies and programs. Disparities between intention and aspiration expressed in public documents and performance in statistical outcomes emerge from this analysis. For example, sex discrimination legislation was enacted in Australia in 1984, and gender equal opportunity legislation in 1999. All universities have policies in place in compliance with such legislation, and 33 of the 43 Australian universities receiving public funding have signed up to the Athena Swan accreditation framework as part of the Science in Australia Gender Equity project (SAGE, 2019). Gradual progress in achieving gender balance since 2000 is apparent, with Australian universities in 2017 showing the overall percentage of full-time equivalent (FTE) women staff as 57%. However, individual

university figures range between 47% and 65%. Analysing further by classification levels, the overall percentage of women in non-academic positions is 66%, but below parity at 46% for academic positions. The percentages decrease further for senior academic positions of Associate Professor and Professor: 32% overall for all Australian universities (Department of Education and Training, 2018). European universities showed similar results in 2015, with a majority of countries and institutions achieving an academic staff gender balance of between 40 and 60%, but a minority of women (15 to 30%) in full professor positions (Hovdhaugen and Gunnes, n.d.).

16 A sample of 12 universities from Australia, New Zealand, Southern Africa, Europe and East Asia in Table 2 shows that although academic staff gender balance is close to 50% in some institutions, fewer women hold professor and senior academic positions than men. These figures suggest that despite intentions and policies to diversify the communities carrying out research in some regions, progress in achieving such diversity at the highest academic levels is slow. There are some variations in the Professor/Senior academic classifications in the categories in this table but the overall percentages are indicative.

Table 2: Gender balance in senior academic positions from four regions, 2015.

2015	Academic men	Academic women	Total academic	% Academic women	Professor/ senior academic men	Professor/ senior academic women	Total professor/ senor academic	% Women professor/ senior academic
AU*	1069	639	1708	37.41	425	137	562	24.38
AU*	828	690	1518	45.45	271	117	388	30.15
N.Europe	803	278	1080	25.73	118	6	124	4.84
UK	3895	3320	7220	45.98	760	270	1030	26.21
NZ**	1,790	1,590	3,385	46.97	260	80	340	23.53
SA	634	545	1179	46.23	164	56	220	25.45
W.Europe	1409	897	2306	38.90	161	39	200	19.50
W.Europe	4736	3654	8390	43.55	477	136	613	22.19
W.Europe	1566	1300	2865	45.35	225	60	285	20.94
E. Asia***	1166	669	1836	36.44	215	35	250	14.00

* These sources combine Associate Professor with Professor
** Professors and Deans
***Category is Senior Academic
Sources: National and regional statistical agencies.

Access to academic libraries

17 Library access policies have the potential to serve as a useful proxy for the openness of an institution. By their nature, such policies provide insight into the ways in which a university views its role within the knowledge landscape of a wider community. It is arguable that library access policies reflect the extent to which a university views its knowledge resources as assets to be managed on behalf of an exclusive group of staff or students; or as resources most likely to benefit both the community and the institution if they are shared beyond the university. To explore the correlation of library access policies with open access and open science policies, we analysed policy documents and library practices to identify types and levels of institutional and external library membership for a sample of 12 academic libraries from Australia, China, Hong Kong, Singapore, South Africa, the United Kingdom and the United States. Users who are unaffiliated with institutions may be granted membership or access to university libraries. However, multiple categories of membership reflect differing levels of eligibility for privileges and services, fees charged, and restrictions on physical access. Overall categories of library access can be grouped in three concentric positions indicating their relationship to the core business of the university: the academic community; individuals and organisations who have prior, established relationships adjacent to the university; and the general, unaffiliated public who have specific research or other information needs:

18 **Core:** faculty, staff, students of an institution

19 **Adjacent:** retired, former, ex-staff; spouses; alumni; visiting researchers, scholars; reciprocal scheme borrowers; business and/or industry; societies, non-profit organisations; government departments and agencies.

20 **General public:** community or public members, independent or private researchers, commercial researchers, other university students, school students, one-time or day visitors. (See Figure 5).

Figure 5: Model of library user categories
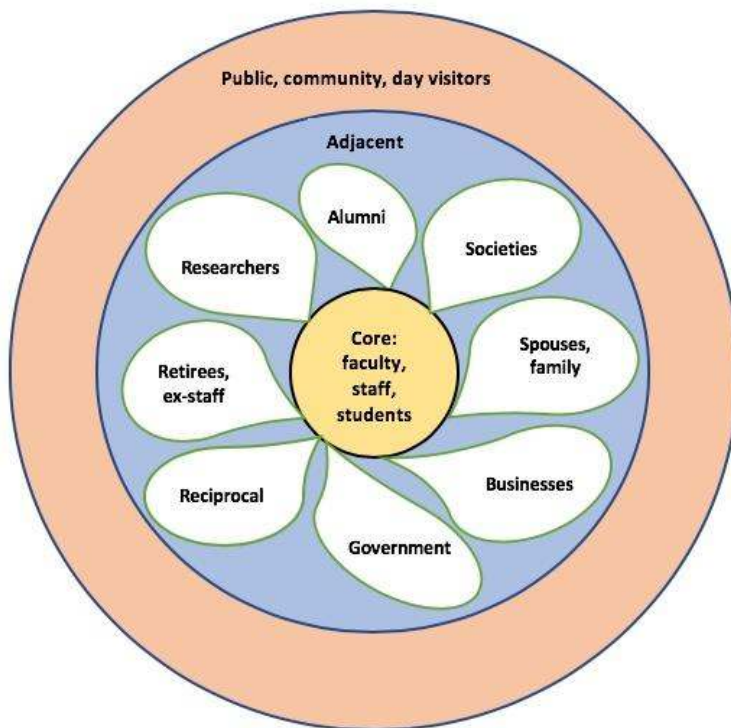
21 We find that in most libraries in the sample, external users in the adjacent group have physical and collection access through special agreements with the institution, such as alumni, retired staff, reciprocal borrowing programs, learned societies, government departments and businesses. This access includes borrowing rights and sometimes, but not always, limited access to electronic resources such as databases and online journals. However, access is most restrictive for the general public who in some cases are required to pay a daily, weekly, monthly or annual membership fee in order to enter an academic library. This varies by country and within countries and can depend on the location and research status of institutions. Limitations to access can develop in response to particular problems or situations such as security, campus unrest and high demand for information and access from the unaffiliated public. Library privileges and access for the general public with no institutional affiliation emerges as a key differential factor and indicator of openness to knowledge within academic libraries.

22 Correlating library access with institutional open access policies, repositories and percentages of open access publishing showed limited correlation, in particular lower scoring for library access policies. This suggests the intentions expressed by the two policy actions have not been applied in similar directions, in other words to reach the same ends. As open access publishing options and mandates expand, public physical entry and access to print and electronic resources in academic libraries is contracting (Wilson et al., 2019). This position varies across countries and regions, but it conflicts with global library and information commitments to open access to knowledge, and illustrates a lack of coordination between policies and practices at institutional levels. We do not judge libraries' access and membership policies, but point out the impact, intentional or not, policy restrictions can have on institutional positions on openness.

Challenges

23 The Curtin Open Knowledge Initiative project aims to achieve global and language diversity in collecting and analysing data. However, obtaining objective and equivalent data at institutional levels across countries and regions presents a number of challenges.

Affiliation and date

24 Affiliation identifiers assigned to each object in bibliographic data sources can be aligned at multiple and varying levels of the institutional hierarchy, ranging from research centres, campuses, universities to university systems. Authors may use different language, terminology and abbreviations when providing affiliation in a publication byline analysed by a data source. ORCID aims to disambiguate author names and standardise institutional affiliations through the allocation of unique identifiers to registered members. However, the adoption of ORCID at country and institutional levels, and registration by individual authors, is a work in progress. Linguistic and translated versions of institutions, name variants including formal renames, abbreviations, common names, name changes and merging of institutions at points in time create further challenges for identifying and aggregating institutional data.

25 In relation to publication dates, each data source may have a different rule or policy in terms of what is to be recorded. For publications, a source may record the date of first online appearance, date of publication online, date of print, date of volume or issue, sometimes with only minor differences between each case. The granularity in metadata also varies, ranging from day-month-year date to just the year.
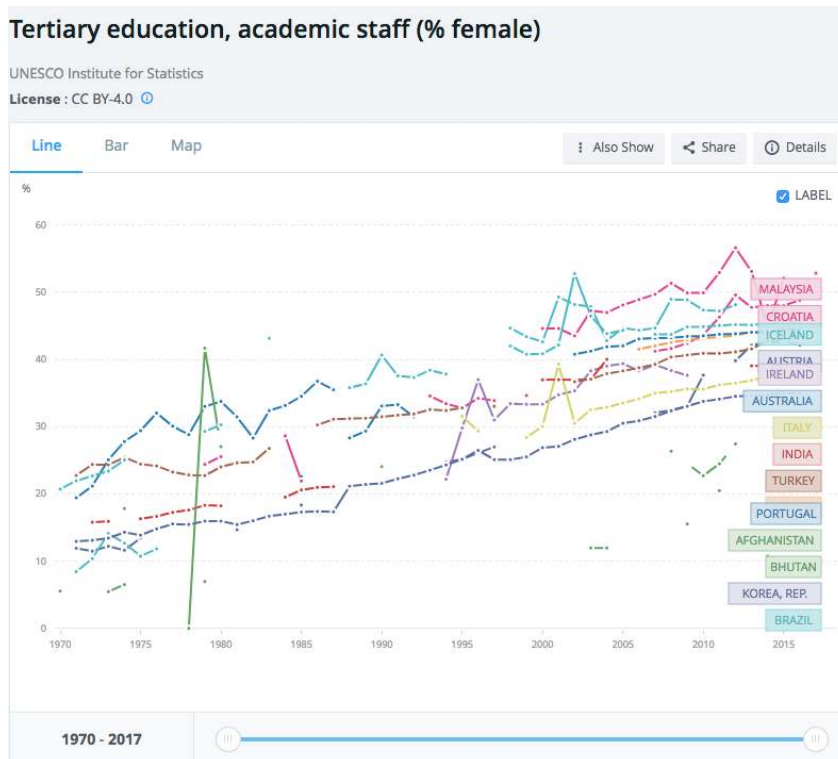
Collection and cross-classification of statistical data

26 Additional measures and initiatives extend and expand our understanding of institutional openness to knowledge. Policy documents and statements may express intentions, but evidence is required to assess progress in achieving outcomes.

27 However, when collecting, normalising, analysing and integrating data across regions, countries and continents there is limited standardisation of measures. At the first level, diversity in terms of gender and ethnicity in populations may be collected and collated differently, making integrations of data problematic and even irrelevant. The motivation or requirement to gather statistics is often driven by policy, affirmative action programs and legislation that may not be present in all countries. Thus data may not be collected for analysis. Cultural influences and embedded assumptions can affect the collection and analysis of statistical diversity data such as staff percentages of gender, ethnicity, age, disability and salary. Questions of appropriation of data include who collects and who owns the data; what cultural assumptions are inherent in data collection and how they affect the results. In relation to higher education populations, indigeneity, ethnicity and gender dimensions within the statistics gathered may not always be politically neutral or culturally appropriate. For example, individuals or population groups may not want to declare a status for reasons such as systematic, embedded racism. Questions or levels of data gathered may not contain sufficient granularity or options to reflect reality, and broad categories of ethnicity may exclude some population groups. Data sovereignty and

decolonisation are key considerations in relation to the relevance and quality of data affecting indigenous and ethnic minorities. (Wilks et al. 2018).

28   Second, for a range of different reasons statistical data are not shared publicly to the same degree in all regions and countries, or considered worth sharing. Gathering statistics on a regular, or even irregular basis is resource intensive, and governments and institutions may not always be resourced or wish to allocate resources to gather and share such data. Even comparatively well-resourced countries and governments may not regard diversity data as a priority to collect or share publicly. To recoup some resource costs, collection agencies and governments may charge for access to full, raw data, and only provide summary analysis freely. This can impact the process of obtaining objective data for research and analysis.

29   The collection, availability and extent of data on gender breakdown, disabilities and underrepresented minorities employed by universities is often mandated by legislation. In Australia, universities are required to gather and report to government the numbers of women and men employed by academic classification level, teaching or research function and indigenous staff numbers, but not disability or salary (Department of Education and Training 2018). In the United Kingdom, to meet legislation and government requirements, higher education providers report staff data by sex, age group, disability, ethnicity and salary range (Higher Education Statistics Agency HESA, n.d.). Within Europe, the Eurostat tertiary education statistics collection shares staff gender figures at country level only. Details of gender, citizenship of academic staff and professors, numbers of non-academic staff at institutional level are available in the ETER (European Tertiary Education Register, n.d.) but not all countries provide equivalent data for equivalent time periods. In the United States the Integrated Postsecondary Education Data System (IPEDS) provides a large amount of diversity data gathered by survey from over 7,000 institutions (Department of Education National Center for Education Statistics, n.d). Statistics indicating degrees of disability of university staff members are collected and shared in few countries. The variations in categories of statistical data across multiple locations can mean reduction to a smaller, common set of data.

30   Diversity data may also be collected and provided via survey methodology. While useful, such data are not always comprehensive. These problems compound at global levels. As an illustration of this, the World Bank databank visualisation tool provides analysis of UNESCO Institute for Statistics data to display percentages of women staff in tertiary education. One can select countries and graph available data over the period 1970 to 2017. Figure 6 demonstrates the variability in data for a sample of 12 countries. Overall the percentages of women academic staff have increased across the years, but data availability and completeness are intermittent or limited for many countries in the sample set. The highest levels (above 50%) are shown by Malaysia and Brazil and the longest period of unbroken data comes from South Korea (The World Bank Group, 2019). Other countries show breaks in the time period or very limited data. Although the reporting is at country level, this demonstrates the challenges in obtaining global data.

Figure 6: World Bank visualisation showing percentages of women in tertiary academic staff for selected countries 1970-2017. Colours of country names correspond to line colour on the graph. A single dot represents one year.

31   The collection and availability of staff diversity and equity statistical data is inconsistent and non-standard throughout the world. Thus, the challenge is how to cross-classify and integrate data from regions with disparate policies, political and cultural approaches in order to understand issues of diversity, gender equity and inclusion.

Linguistic diversity

32   Language and terminology in policy documents and statistics sources vary, making cross-comparison of textual and statistical data difficult. To assist in locating and retrieving policy documents and analysing data from institutions across countries and languages we have developed a multilingual lexicon of open knowledge and scholarly communication terminology, benefitting from the contributions of others in the scholarly communication field through the twittersphere (Lexicon Contributors, 2018). This is helpful in constructing web searches, addressing the institutional and linguistic variations along with translation services and sources, but challenges of terminology, particular for retrieval on a large scale.

Limitations

33   The key limitations in the collection of open access and event data relate to the timeframes over which data sources have collected data and the coverage of repositories by Unpaywall. Funder data, reporting funding sources for published scholarly research, only exists from the commencement of the Crossref Funder Registry in 2013 and does not

provide complete coverage of all funding bodies. Crossref Event Data is collected from 2015 onwards. The bibliographic data sources have substantial biases and limitations with respect to discipline coverage, with Scopus and Web of Science providing less coverage of Humanities and Social Sciences. We believe these are mitigated through our use of three independent data sources. As discussed above, gathering statistical data and institutional intentions through policy documents and annual reports is limited by the availability, accessibility, equivalence and comprehensiveness of sources.

Conclusion

34   This paper has discussed key areas identified by the Curtin Open Knowledge Initiative project as potential indicators of institutional progress towards openness. There is potential for such data to contribute towards a university's knowledge of its own progress towards positions of openness, open access to research and data, inclusion and diversity in staffing and research. The project is ongoing and continues to extend and enhance data collection and analysis to confirm a model for understanding the current and potential reach of openness to knowledge.

35   Currently we have detailed data for a set of higher education institutions for individual years between 2000 and 2018. We are expanding the methods and capacity to gather and analyse such data globally and longitudinally in the areas discussed here. At the same time, we recognise and acknowledge the challenges in relation to obtaining data and presenting analysed data objectively, but this forms part of the process of understanding the barriers to, and ways of achieving openness. The project builds on the principles, paths and proxies identified in this paper and welcomes the expertise and contributions of others in the scholarly open knowledge community to continue to assemble and analyse data for institutions to assist in their future decision-making on the journey to openness.

References

36   Australia Department of Education and Training, *Higher education statistics: Staff data,* 2018. https://www.education.gov.au/staff-data

37   Crossref, *Event data,* 2017 https://www.crossref.org/services/event-data/

38   European Tertiary Education Register, *ETER.* n.d. https://www.eter-project.com

39   Higher Education Statistics Agency, *HESA.* n.d. https://www.hesa.ac.uk/

40   Hovdhaugen, Elisabeth, and Hebe Gunnes, *What ETER tells us about gender balance among academic staff in European HEIs.* European Tertiary Education Register, n.d., https://www.eter-project.com/uploads/assets/pdf/ETER_gender.pdf

41   Lexicon Contributors, *Scholarly communications multilingual lexicon v0.2.,* 2018, https://doi.org/10.5281/zenodo.1954813

42   Montgomery, Lucy, Hartley, John, Neylon, Cameron, Gillies, Malcolm, Gray, Eve, Herrmann-Pillath, Carsten, Huang, Chun-Kai, Leach, Joan, Potts, Jason, Ren, Xiang, Skinner, Katherine, Sugimoto, Cassidy R., and Katie Wilson *Open knowledge institutions: Reinventing universities.* (in review). MIT Press WIP https://wip.mitpress.mit.edu/pub/oki. doi: 10.21428/99f89a34

43 Montgomery, Lucy and Cameron Neylon,"In a globalised and networked world, what is the unique value a university can bring? Introducing Open Knowledge Institutions", *LSE Impact Blog, September 17, 2018,* https://blogs.lse.ac.uk/impactofsocialsciences/2018/09/17/in-a-globalised-and-networked-world-what-is-the-unique-value-a-university-can-bring-introducing-open-knowledge-institutions/

44 Piwowar, Heather, Priem Jason, Larivière Vincent, Alperin Juan Pablo, Matthias, Lisa, Norlander, Bree, Farley, Ashley, West, Jevin, and Stefanie Haustein, "The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles", 2018, *PeerJ* 6: e4375, https://doi.org/10.7717/peerj.4375

45 Science in Australia Gender Equity, *SAGE,* 2019, https://www.sciencegenderequity.org.au/

46 The World Bank Group, *World Development Indicators,* 2019, https://data.worldbank.org/indicator/SE.TER.TCHR.FE.ZS?view=chart

47 U.S. Department of Education National Center for Education Statistics, *IPEDS - integrated Postsecondary Education Data System*, n.d, https://nces.ed.gov/

48 Wilks, Judith, Kennedy, Gillian, Drew, Neil, and Katie Wilson. "Indigenous data sovereignty in higher education: Towards a decolonised data quality framework", *Australian Universities Review, 60*(2), (2018): 4-14 http://www.aur.org.au/archive

49 Wilson, Katie, Neylon, Cameron, Montgomery, Lucy, and Chun-Kai. Huang, "Library access policies: an indication of openness?", *Information Research, 24*(1), paper 809, 2019, http://InformationR.net/ir/24-1/paper809.html (Archived by WebCite® at http://www.webcitation.org/76tOSpfrn)

ABSTRACT

Universities are key sites of knowledge creation. Governments and research funders are increasingly interested in ensuring that their investments in the production of new knowledge deliver a quantifiable return on investment, including in the form of 'impact'. Ensuring that research outputs are not locked behind paywalls, and that research data can be interrogated and built upon are increasingly central to efforts to improve the effectiveness of global research landscapes. We argue that mandating and promoting open access (OA) for published research outputs, as well as the sharing of research data are important elements of building a vibrant open knowledge system, but they are not enough. Supporting diversity within knowledge-making institutions; enabling collaboration across boundaries between universities and wider communities; and addressing inequalities in access to knowledge resources and in opportunities to contribute to knowledge making processes are also important. New tools are needed to help universities, funders, and communities to understand the extent to which a university is operating as an effective open knowledge institution; as well as the steps that might be taken to improve open knowledge performance. This paper discusses our team's efforts to develop a model of Open Knowledge that is not confined to measures of OA and open data. The Curtin Open Knowledge Initiative is a project of the Centre for Culture and Technology at Curtin University. With funding from the university, we are exploring the extent to which universities are functioning as effective open knowledge institutions; as well as the types of information that

universities, funders, and communities might need to understand an institution's open knowledge performance and how it might be improved. The challenges of data collection on open knowledge practices at scale, and across national, cultural and linguistic boundaries are also discussed.

INDEX

Keywords: Open access, open knowledge, diversity, inclusion, openness

AUTHORS

CAMERON NEYLON

Curtin University, Australia

RICHARD HOSKING

Curtin University, Australia

CHUN-KAI (KARL) HUANG

Curtin University, Australia

ALKIM OZAYGEN

Curtin University, Australia

KATIE WILSON

Curtin University, Australia