Becoming Open Knowledge Institutions: Divergence, Dialogue and Diversity

Katie Wilson[1]([✉]) [iD], Lucy Montgomery[1] [iD], Cameron Neylon[1] [iD],
Rebecca N. Handcock[2] [iD], Richard Hosking[2] [iD], Chun-Kai (Karl) Huang[1] [iD],
Alkim Ozaygen[1] [iD], and Aniek Roelofs[2] [iD]

[1] Centre for Culture and Technology, Curtin University, Perth, Australia
katie.wilson@curtin.edu.au
[2] Curtin Institute for Computation, Curtin University, Perth, Australia

Abstract. The Curtin Open Knowledge Initiative (COKI) is an innovative research project that collects and analyses publicly available research output data to assist and encourage researchers, academics, administrators and executives to understand the actual and potential reach of openness in research, and to assess their progress on the path towards open knowledge institutions. By taking a broad global approach and using multiple data sources, the project diverges from existing approaches, methods and bibliometric measures in the scholarly research environment. It combines analysis of research output, citations, publication sources and publishers, funders, social media events, open and not open access to provide overviews of research output and performance at institutional, funder, consortial and country levels. The project collects and analyses personnel diversity data such as gender, focusing on widening the reach of data analysis to emphasise the importance and value of diversity in research and knowledge production. Interactive visual tools present research output and performance to encourage understanding and dialogue among researchers and management. The path towards becoming open knowledge institutions involves a process of cultural change, moving beyond dominant publishing and evaluation practices. This paper discusses how through divergence, diversity and dialogue the COKI project can contribute to this change, with examples of applications in understanding and embracing openness.

Keywords: Open knowledge institutions · Open access · Research performance · Higher education · Data visualization

1 Introduction

The Curtin Open Knowledge Initiative (COKI) is a strategic research project established at Curtin University, Australia through a critique of global rankings dominating the higher education environment. It challenges the dependence of rankings on impact and citation data from scholarly publications with embedded epistemologies that favour scientific scholarly disciplines and the English language, ignoring a large body of research and knowledge from non-dominant populations and countries. The project embodies a theory of cultural change, focusing on openness and diversity in research production, and new

ways of thinking about institutional research impact. COKI collects, aggregates and analyses data from multiple sources, including information on open access publications, research collaboration, social media events, diversity, policies and infrastructure. The project includes open source software architecture and a cloud platform that gathers, integrates, manages and reports on the data. We undertake research projects that provide critical insights into research performance across countries and institutions, access to knowledge, and the outcomes and interventions of institutional policies [1, 2].

Knowledge is social in terms of its production, its capacity to support public interest, communication, and the different kinds of value that it creates. As discussed by [3], what universities contribute as Open Knowledge Institutions is found in how they support and interact with diverse communities, as well as how they practice communication to encourage dialogue in order to create the diverse values that underlie different conceptions of knowledge. Diversity, inclusion and equity lead to more effective communication, dialogue and knowledge production. Inclusive coordination provides platforms for dialogue among different types of knowledge participants. Broadening traditional measurement of research output to include a larger set of formats such as digital scholarship, short monographs, translations, creative writing, art works, and performance scholarship empowers a wider range of voices and encourages new forms of research and dialogue with extended communities. Diverse models of production and access provide opportunities for broader dialogues in knowledge creation [3]. This involves a process of thinking that diverges from existing publishing and evaluation practices within research institutions.

2   Divergence

Universities are influenced by global market rankings that depend on impact and citation data extracted from dominant scholarly publications primarily in two commercial citation indexes, Scopus from Elsevier and Web of Science from Clarivate Analytics [4]. The ranking systems make assumptions about the usefulness of their proxy data measures that are based on ontologically inconclusive assumptions and reasons [5]. However, the presence of influential world rankings pressure universities and researchers in many countries to aspire to publish in the dominant sources in order for their institutions to compete on the ranking scale. This devalues diverse research from non-dominant countries in the global south, non-English languages, indigenous and minority populations and in disciplines such as the humanities and political activist scholarship where the primary output format is not journal publications [6, 7]. Even universities that explicitly refuse to participate in rankings continue to be included in their listings [8].

The COKI project diverges from dominant scholarly publishing thinking and practices in universities through a critique of the limited bibliometric measures used by global university rankings. It regards information as a community asset and highlights an opportunity to move beyond established, commercial data resources towards the construction of data resources that are governed by the higher education and research community. Universities possess the technical skills and expertise to capture data relating to research and scholarly communication at very large scales, and to build the tools and frameworks needed to interpret and contextualize such data once captured. However, a lack

of common approaches and shared datasets has hampered measurement and evaluation, particularly on an international scale. Development of our project has generated intense interest from a wide range of researchers, organisations and institutions, revealing an unfulfilled information need and leading to sharing and dialogue among different communities who come together to exchange knowledge and experiences. These include institutional executives, researchers, librarians, funding bodies, consortia, educational networks and associations from Africa, Arab nations, Europe, Latin America, North America, Oceania and the United Kingdom.

3   Dialogue

COKI collects, stores and analyses publicly available data to enable researchers and universities to track and understand the potential for increasing the reach of their research through open publication formats and options that extend beyond citation chasing. The project has developed a dataset of more than 12 trillion items related to scholarly communication, open access, equity, diversity and inclusion. The dataset draws affiliation, publication, funding and social media events data from Microsoft Academic, Unpaywall, Crossref, Open Citations and ORCID to understand open research output and performance at institutional, funder, publisher and country levels. It includes over 100 million research outputs, 20,000 organisations and 20,000 funders covering a breadth of research output to reach an understanding of who creates research knowledge. This approach aims to encourage dialogue among academics, researchers, library and research staff, administrators and senior executives in universities, research and funding organisations. In order to engage with a wide audience, including those who are not necessarily experts in research measurement, COKI's visual interactive dashboard tools facilitate sharing of data and dialogue among diverse groups and individuals within institutions and communities. The project's data resources and dashboards present open research publication and performance data for research institutions and consortia, funders and publishers.

Existing and new data analysis combine to provide visual, interactive research performance overviews, including open research, funder compliance, collaboration, social media events and diversity. Evidence of the strength and uniqueness of the project emerged from responses to the COKI dashboard developed for the Council of Australian University Librarians (CAUL) and the Council of New Zealand University Librarians (CONZUL). We undertook a research survey of librarians and research staff from the two consortia (CAUL and CONZUL) who used the dashboard in 2020. Respondents commented on the accessibility of the visual format, the value of data such as citation advantages of open access publications (see Fig. 1) and compliance with funder open access mandates, that will enable them to initiate dialogue with researchers, directors and senior executives:

- *Citation data OA/non-OA, averages collated visually in this way very useful for showing OA reluctant researchers the benefit of OA in terms of citations.*
- *Very useful to have all the data together in one place and presented visually.*
- *Reports like these are difficult to get from the IR* [institutional repository]*, and cannot get visualisations like these.*

years.publish...	Open Access			Not Open Access			Grand total		
	Articles	Citations	Average Citations	Articles	Citations	Average Citations	Articles	Citations	Average Cit...
2007	250	3,188	12.75	557	6,941	12.46	807	10,129	12.61
2008	373	5,040	13.51	672	7,779	11.58	1,045	12,819	12.54
2009	476	6,551	13.76	680	7,137	10.5	1,156	13,688	12.13
2010	574	8,050	14.02	890	9,288	10.44	1,464	17,338	12.23
2011	605	7,392	12.22	1,013	9,682	9.56	1,618	17,074	10.89
2012	686	7,265	10.59	1,074	8,900	8.29	1,760	16,165	9.44
2013	1,098	12,173	11.09	1,156	7,711	6.67	2,254	19,884	8.88
2014	1,263	10,889	8.62	1,213	6,935	5.72	2,476	17,824	7.17
2015	1,419	9,396	6.62	1,431	6,773	4.73	2,850	16,169	5.68
2016	1,635	8,497	5.2	1,749	4,809	2.75	3,384	13,306	3.97
2017	1,692	3,904	2.31	1,929	2,562	1.33	3,621	6,466	1.82
2018	1,473	416	0.28	2,069	411	0.2	3,542	827	0.24
Grand total	12,323	96,964	12.67	16,755	109,078	9.12	29,078	206,042	10.89

Fig. 1. Graph from the CAUL/CONZUL institutional dashboard. Analysis of open access and non-open access article publication citation data for Curtin University 2000–2018 shows overall citation averages for open access article publications are higher than not open. Note that only 2007–2018 data display in the image but the Grand total refers to the extended graph data from 2000–2018. Analysis and graph: COKI.

- *Funder OA/non-OA data useful…The [funder] compliance data is very interesting, especially to some Senior Executives* [9].

COKI offers two publicly available dashboards. The COKI Open Access Dashboard provides country level publication research output by open/non-open access and open/non-open article citation analysis [10]. Figure 2 from this dashboard presents Australia's research output from 2000–2020: analysis of research output by total OA, Gold OA (publisher-mediated), Green OA (repository-mediated), Hybrid OA (published in a journal not listed in the Directory of Open Access Journals (DOAJ), but free to read from publisher with any license), and Bronze OA (free to read online via a publisher but without a license).

The public COKI Research Funding Dashboard provides country level analysis of national and international research funding sources acknowledged in publications from 185 countries: the ratio of domestic to international funding; country of funder; major funders grouped by publication year; and funder name with number of acknowledgements per funder [11]. Figure 3 shows major funders for Australian institutional research output.

The strength of COKI's large dataset is its comprehensiveness, and while the detailed data analysis may be specialized, visual presentations such as those in the figures above provide clear messages about research performance and funding sources. Such messages can initiate further investigation, dialogue and analysis, and assist institutions in charting their progress and publishing options. COKI is working with a range of North American consortia, specifically building dashboards to address consortium level tracking of research publications. Alongside publication and funder data, we analyze public staff institutional demographic data such as gender, indigeneity, ethnicity or nationality and disability, where available, to highlight understanding of diversity within research production and creation. This is discussed in the next section.
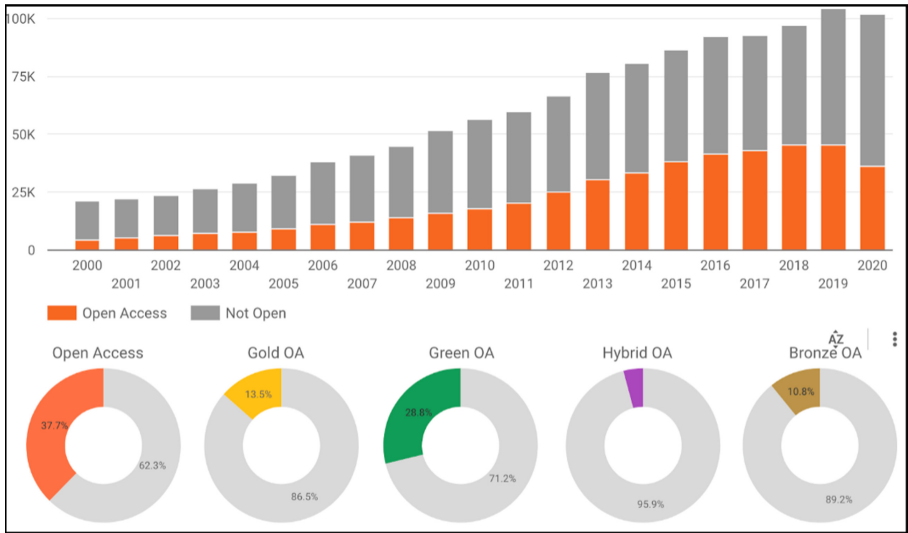
Fig. 2. Australia's research output in terms of open access and not open access, plus percentages of Gold, Green, Hybrid and Bronze OA, 2000–2020, from the public COKI Open Access dashboard. Analysis and image: COKI. (Color figure online)
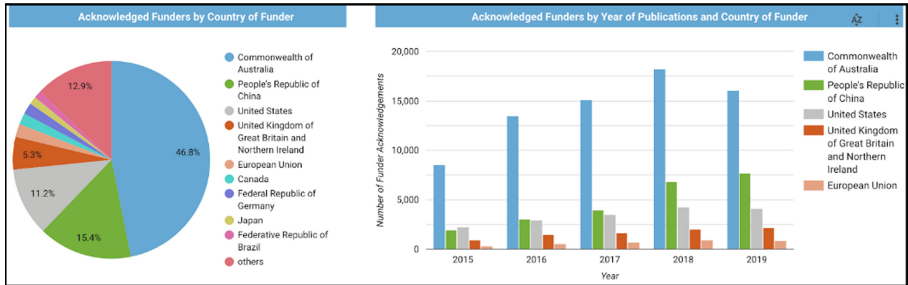

Fig. 3. Australia: acknowledged funders by country and major funders by publication year, 2015–2019, from the public COKI Research Funding dashboard. Analysis and image: COKI.

4 Diversity

Diversity is a key element of open knowledge institutions, and to achieve openness, universities also need to address challenges of diversity in input and output. This means understanding *who* is involved in knowledge creation and research, as well as *how* knowledge is shared within disciplines and scholarly communities, across disciplines, between universities and wider communities. A diversified staff that reflects the diversity of institutional student bodies and communities contributes to openness and the inclusion of diverse ideas, knowledges and languages. Systemic biases within academic research practices and institutions, including colonisation, genderism, sexism and racism perpetuate and legitimize marginalisation of groups of people and thought [12–15]. Diversity

in open knowledge is enacted through the transmission of knowledge to students in teaching and the creation of diverse models and examples. Underpinning research and teaching is internal diversity among personnel and practices in institutional administration, management and decision-making. External diversity encompasses openness in organisational relationships and partnerships, funding arrangements and bodies, crossing geopolitical boundaries [16].

COKI collects and analyses data to enable understanding of how universities operate as effective open knowledge institutions and how to improve their open knowledge performance. In turn this enables universities to work more effectively with diverse local and global communities in the production of knowledge. Staff demographic statistical data provide quantitative evidence of institutional diversity. While this is only part of the story, it assists understanding of institutional and national performances in relation to policy intentions and outcomes. We identify publicly available staff diversity data globally in order to develop potential indicators to understand progress in response to diversity, inclusion and equity policies and programs at institutional levels.

The path towards institutional openness involves moving from closed or traditional ways. At the beginning of this change model, many policies appear to address deficits and institutional preservation and reputation. The extent to which an institution translates its intentions into reality is not always visible or transparent in terms of outcomes. Narratives and policy around equity, diversity and inclusion may be associated with good intentions but do not necessarily achieve transformation. Discriminatory practices within universities are structural and systemic, often reflecting wider cultural customs within countries and regions as well as academic traditions. Legislation, policies and programs that aim to address discrimination frequently are introduced by those who are part of the dominant system [16]. Tools may be regarded as indicators of progress but outcomes can indicate a different picture [17]. COKI analyses staff demographic diversity data to understand the results of institutional policies and practices and how these can contribute to openness.

We identify public data sources from ministries and departments of education, higher education funding bodies and research associations that collect and collate institutional staff demographic statistics. These are at country level and wider regional levels. For example, in Europe, the European Tertiary Education Register (ETER) provides detailed data for multiple countries, but statistics are also available from individual country sources. Differences at country levels present challenges in collecting global data. Some countries do not provide detailed statistical collections for public access, for example, those undertaking higher education structural and financial reform. Nations experiencing war and political turmoil may have lower priorities and limited funding directed towards higher education and research [16].

The diversity dimensions collected and reported on most consistently are gender, origin (race, ethnicity or nationality) and age, but detail and availability of the data vary widely by country. Gender or sex of university staff members is the most commonly available dimension across the statistical sources analysed. Limitations exist within this collection however. Data are sometimes only provided for academic staff or faculty. Non-binary sexual preferences are not frequently available, and are only recently recorded

in a few countries where numbers are small and are excluded from statistics for privacy and confidentiality.

Where possible, we incorporate staff diversity data into the project's interactive visual data dashboards. For example, the data analysis in Fig. 4 shows the breakdown of women and men in academic and non-academic positions at Curtin University, from 2001 to 2018. The number of women reached 57% in 2018, but only 46% of academic staff were women in the same year.
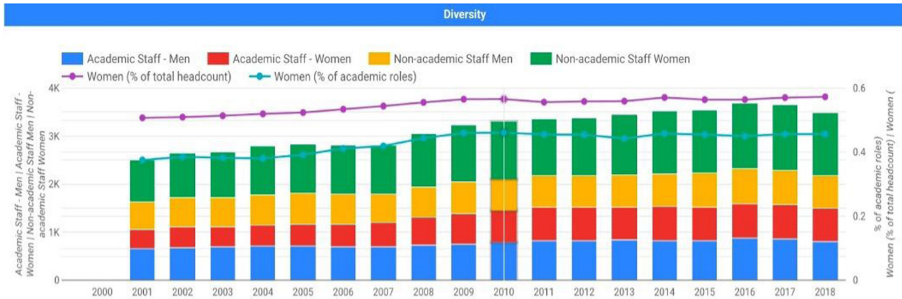
Fig. 4. Gender (men, women) share of academic and non-academic staff, Curtin University, 2001–2018. Source data: Australian Department of Education, Skills and Employment 2019. Analysis and image: COKI.

Diversity also extends to equal inclusion of the knowledges, research and output from researchers and institutions in non-dominant nations. As Fig. 5 illustrates, the highest open access output, both Gold (published in the Directory of Open Access Journals (DOAJ) or free to read via publisher with license) and Green (free to read via an open access repository), emerges from Africa, Asia, Latin America and Europe.

Through a focus on research diversity, the project raises questions about the processes of data collection and data use relating to research output. What are the underlying intentions and epistemologies of the collection process; who determines and legitimizes data collection methods; how are datasets analysed and presented; how do these factors affect the outcomes of analysis, and how the data are used? Concerns about research inequality and data sovereignty include the assumptions and values at many levels, from institutional to governmental, that shape approaches to data collection [18, 19]. For example, research output data from many African-based publications are excluded from the citation sources on which so-called global rankings depend. To extend the project's dataset beyond the global north and English language publications COKI is actively collaborating with partners and organisations on the African continent to incorporate and analyze their research output data.
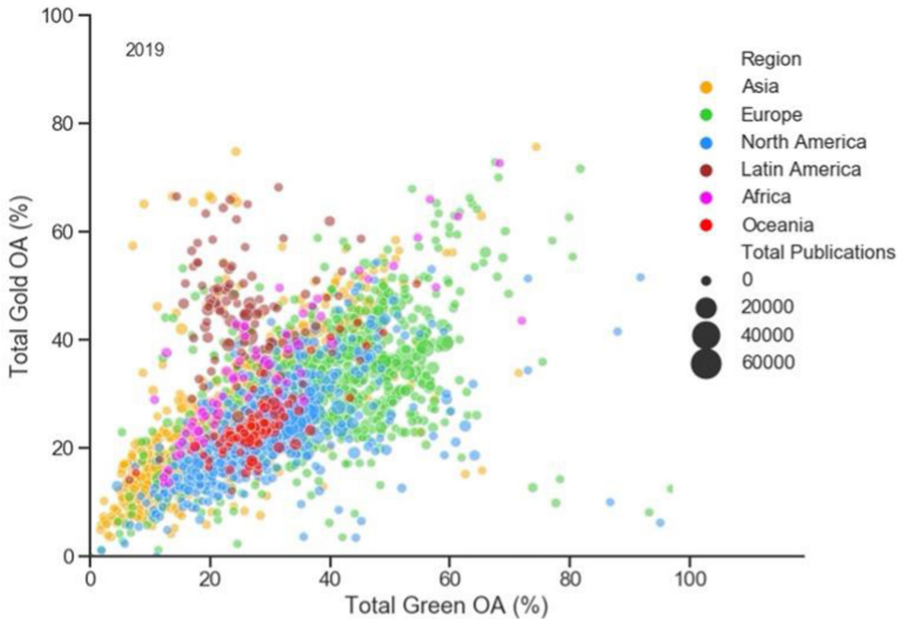
Fig. 5. Regional open access performance shows strong output in Gold (DOAJ and hybrid journals) Latin America and Asia; Green (repository-mediated) OA output for Africa and Europe, 2019. Each dot represents an institution and the size of the dots indicate the number of publications. Data sources: Microsoft Academic, Scopus, Web of Science, CrossRef, UnPaywall. Analysis and image: COKI. (Color figure online)

5 Conclusion

Through data analysis and open practices COKI provides opportunities for researchers, institutions, funding organisations and publishers to understand the reach and potential of open research, and to explore their individual and collective progress in becoming open knowledge institutions. While the project identifies a need for institutional cultural change in this process, it recognizes that this change is complex. The process involves divergence from existing assessment and evaluation practices and associated attitudes within research institutions which reward prestige publishing, follow world university rankings, and ignore open knowledge. It requires multidirectional change from individuals at grassroots levels and management levels within institutions, in the scholarly community and the marketplace. COKI user-centered dashboards are designed to encourage dialogue and assist researchers, managers and students to understand and critique the tools, datasets and methodologies used to measure research performance, as well as the agency that institutions and individuals have to make change within this process.

COKI embraces diversity in bringing together evidence of open research practices from different sources and challenging institutions to review their diversity and inclusion practices and outcomes. We encourage sharing of project data, software and visual dashboard code among researchers, library, research and technical staff and senior executives to enable critical data literacy upskilling [9, 20] and to facilitate dialogue. Options for

alternative research publishing practices are well established, but are overshadowed by processes constrained by third party organisations with commercial interests in maintaining dominance in the scholarly communication market. Expanding dialogue among interested members of the scholarly community can include collaboration with commercial sources where relevant and appropriate. With a divergent, dialogic and diverse approach COKI hopes to enhance the possibilities for change within institutions and progress towards openness in research and knowledge sharing.

References

1. Huang, C.-K., et al.: Meta-Research: evaluating the impact of open access policies on research institutions. ELife **9**, e57067 (2020). https://doi.org/10.7554/eLife.57067
2. Wilson, K., et al.: 'Is the library open?': Correlating unaffiliated access to academic libraries with open access support. LIBER Q. **29**(1), 1–33 (2019). https://doi.org/10.18352/lq.10298
3. Montgomery, L., et al.: Open Knowledge Institutions: Reinventing Universities. MIT Press, Cambridge (Forthcoming)
4. Tennant, J.P.: Web of Science and Scopus are not global databases of knowledge. Eur. Sci. Edit. **46**, e51987 (2020). https://doi.org/10.3897/ese.2020.e51987
5. McCormick, A.C.: The intersection of rankings with university quality, public accountability and institutional improvement. In: Hazelkorn, E. (ed.) Global Rankings and the Geopolitics of Higher Education, pp. 205–215. Routledge, London (2017)
6. Shahjahan, R.A., Wagner, A.E.: Unpacking ontological security: a decolonial reading of scholarly impact. Educ. Philos. Theory **51**(8), 779–791 (2019). https://doi.org/10.1080/00131857.2018.1454308
7. Vera-Baceta, M.-A., Thelwall, M., Kousha, K.: Web of Science and Scopus language coverage. Scientometrics **121**, 1803–1813 (2019). https://doi.org/10.1007/s11192-019-03264-z
8. McKenna, S.: University rankings don't measure what matters. The Conversation, 15 September 2020 (2020). http://theconversation.com/university-rankings-dont-measure-what-matters-145425
9. Wilson, K., et al.: Extending researcher knowledge of open scholarship performance (Forthcoming)
10. COKI Open Access Dashboard. http://openknowledge.community/dashboards/coki-open-access-dashboard/. Accessed 06 Jan 2021
11. COKI Research Funding Dashboard. http://openknowledge.community/dashboards/funder-countries/. Accessed 06 Jan 2021
12. Ajil, A., Blount-Hill, K.-L.: "Writing the Other as Other": exploring the othered lens in academia using collaborative autoethnography. Decolonization Criminol. Justice **2**(1), 83–108 (2020). https://doi.org/10.24135/dcj.v2i1.19
13. Daza, S.L., Tuck, E.: De/colonizing, (post)(anti)colonial, and indigenous education, studies, and theories. Educ. Stud. **50**(4), 307–312 (2014). https://doi.org/10.1080/00131946.2014.929918
14. Farmer, L.B., Robbins, C.K., Keith, J.L., Mabry, C.J.: Transgender and gender-expansive students' experiences of genderism at women's colleges and universities. J. Divers. High. Educ. **13**(2), 146–157 (2020). https://doi.org/10.1037/dhe0000129
15. Smith, L.T.: *Decolonizing methodologies: Research and indigenous peoples*, 2nd edn. Zed Books, London (2012)
16. Wilson, K., Neylon, C., Montgomery, L., Hosking, R., Huang, C-K., Ozaygen, A.: Global diversity in higher education staffing: towards openness. (Forthcoming)
17. Ahmed, S.: Living a Feminist Life. Duke University Press, Durham (2017)

18. Chan, L.: Open Insights: An interview with Leslie Chan. Open Library of Humanities (2018). https://www.openlibhums.org/news/314/
19. Ríos, C.D., Dion, M.L., Leonard, K.: Institutional logics and indigenous research sovereignty in Canada, the United States, Australia, and New Zealand. Stud. High. Educ. **45**(2), 403–415 (2020). https://doi.org/10.1080/03075079.2018.1534228
20. Carmi, E., Yates, S.J., Lockley, E., Pawluczuk, A.: Data citizenship: rethinking data literacy in the age of disinformation, misinformation, and malinformation. Internet Policy Rev. **9**(2), 1–22 (2020). https://doi.org/10.14763/2020.2.1481