



OPEN ACCESS

EDITED AND REVIEWED BY
Chris C. Funk,
College of Letters & Science (UC),
United States

*CORRESPONDENCE
Lauren A. Jackson
✉ lauren.jackson@noaa.gov

SPECIALTY SECTION
This article was submitted to
Climate Services,
a section of the journal
Frontiers in Climate

RECEIVED 26 October 2022
ACCEPTED 15 December 2022
PUBLISHED 17 January 2023

CITATION
Butler KA, Jackson LA, Kruk MC,
Merati N and Vance TC (2023)
Editorial: Democratizing data:
Environmental data access and its
future. *Front. Clim.* 4:1081021.
doi: 10.3389/fclim.2022.1081021

COPYRIGHT
© 2023 Butler, Jackson, Kruk, Merati
and Vance. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Editorial: Democratizing data: Environmental data access and its future

Kevin A. Butler¹, Lauren A. Jackson^{2*}, Michael C. Kruk²,
Nazila Merati³ and Tiffany C. Vance⁴

¹Environmental Systems Research Institute (ESRI), Redlands, CA, United States, ²National Centers for Environmental Information, National Oceanic and Atmospheric Administration (NOAA), Stennis Space Center, MS, United States, ³National Centers for Environmental Information, National Oceanic and Atmospheric Administration (NOAA), Boulder, CO, United States, ⁴United States (U.S.) Integrated Ocean Observing System, National Oceanic and Atmospheric Administration (NOAA), Silver Spring, MD, United States

KEYWORDS

data discoverability, data access, data and service equity, data usability, reproducibility, environmental data

Editorial on the Research Topic

Democratizing data: Environmental data access and its future

Brief introduction to the Research Topic

Data democratization is the equal and interdependent responsibility of data producers, consumers, and curators to make data discoverable, accessible, equitable, and usable (Figure 1). The genesis of democratization stems from critical changes in the supply and demand of data over the last decades. First, the supply of data, particularly environmental data, has exploded, partly due to improvements in the number and resolution of observational platforms and governmental entities and private organizations embracing the philosophy of open public data declaring data as a public good. Likewise, the increased demand for environmental data stems from the realization that planet-scale problems require planet-scale analyses and an increased emphasis on data-driven environmental decision-making (Tonn et al., 2000).

The global phenomenon of *datafication* (Mayer-Schönberger and Cukier, 2013) has resulted in ever-increasing availability, demand, and use of environmental data at unprecedented physical and social scales. Taylor (1997, p. 327) prophetically states, “in the twenty-first century, the community—not the federal government—will be the principal unit of solution to social and economic difficulties.” This shift has expanded the collection and reach of environmental data to previously unforeseen data consumers such as citizen scientists. However, even traditional consumers of environmental data, science practitioners, face data challenges in this era of big data. Meeting the goals of data as a public good and supporting science at local scales requires a new perspective on the production, management, and curation of data. We refer to this new perspective as data democratization—introducing democratic principles into all aspects of data processes

(see Tilly, 2001 for a complete discussion of democratization). The submissions in this Research Topic raise important questions and provide practical examples for making data more discoverable, accessible, equitable, and usable.

Demographics of participation

The nine articles on the Research Topic were submitted by 43 authors representing 27 institutions or organizations. The majority (58%) were submitted by authors affiliated with educational institutions, with the remainder from commercial (7%), governmental (21%), or non-profit (14%) entities. The heterogeneous institutional affiliations suggest that democratizing data presents challenges and opportunities for all data producers.

Topics of the collection

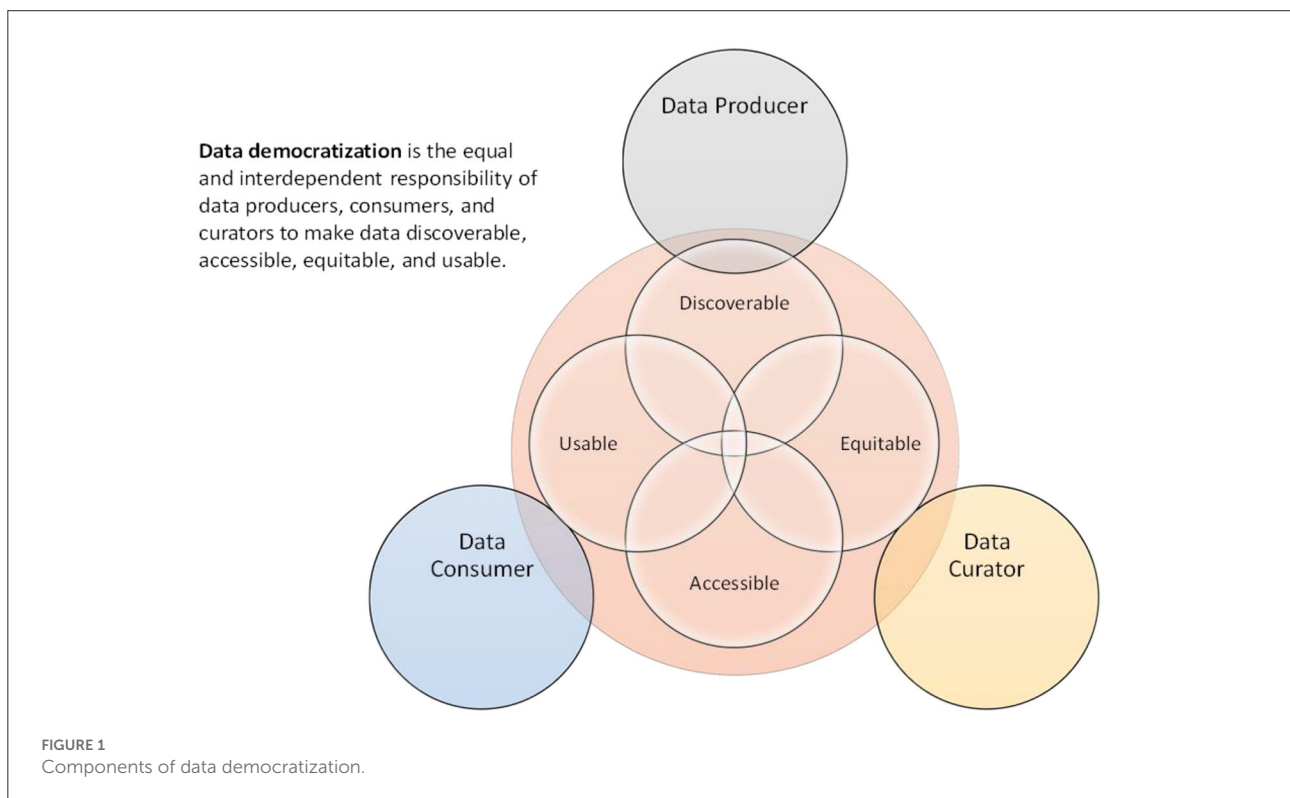
Three broad themes emerged from the collection:

- Maintaining a **user focus** in all aspects of the data lifecycle: Virapongse et al. challenge Earth scientists to be introspective about the methods and processes they use to produce more effective information products to help place-based communities build resilience. Cantor et al.

suggest that the decision-maker needs to be incorporated directly into data systems design. Finally, Gärtner-Roer et al. highlight the recent increased availability of glacier data and the role a centralized user-focused repository and standards organization can play.

- Making **data usability** a priority: Stern et al. describe an open-source platform for extracting archival data and creating analysis-ready, cloud-optimized data stores that empower a broader community of scientists. Contributors also highlighted the challenges of repurposing and making existing data repositories more usable (Rossi et al.) and unique usability challenges in 3D data (Paxton et al.).
- Ensuring **data veracity** and **equity**: In addition to the traditional data veracity challenges in big environmental data, ensuring veracity in time series (Sweeney) and simulation data (Mullendore et al.) presented unique challenges. Finally, Dosemagen and Williams state that prioritizing environmental data as a public good is a key to data usability, and usability is key to addressing environmental justice issues.

Increased pressure from funding agencies that promote or even mandate open data sharing has resulted in a new perspective on data—an explicit focus on the user. Historically, the data lifecycle supported the process of *knowledge production*, collecting, and analyzing data for peer-reviewed journal publications (Baker and Mayernik, 2020). A complementary



data lifecycle, *data production*, “creates data intended for release to a data repository that makes data accessible for reuse by others” (p. 4). This user-focused perspective dramatically impacts how data are documented, managed, and shared. Just as a democratic system of government requires an informed citizenry, democratized systems of data require informed data users.

While big environmental data presents tremendous opportunities, they also come with tremendous challenges. For example, data’s sheer volume, variety, and velocity can impede its usability. Contributors touched on many aspects of the findability, accessibility, interoperability, and reuse (FAIR) principles (Wilkinson et al., 2016), providing the following insights that help reach the goals of democratization:

- View data not as stand-alone datasets but as a system in and of itself (Stern et al.).
- Design data systems to meet the data needs of decision-makers (decision-driven data systems) rather than requiring decision-makers to adapt to existing systems (Cantor et al.).
- Include data users and producers in the design of data access systems (Dosemagen and Williams).
- Strive for *effective* use of data (Virapongse et al.).

Data veracity is, arguably, the riskiest aspect of data democratization. Issues of data quality and fitness for purpose become more critical as data sharing and data use networks grow beyond the data producer. Producers and curators are responsible for summarizing and communicating data quality issues and fitness for purpose in a form and tone approachable by data users outside the subject domain and the technical expertise of the data producers. This approach is consistent with the trend in several academic conferences and journals requesting a plain-language summary of scientific research.

Beyond FAIR

How is data democratization different from the FAIR principles? While the FAIR principles are an essential first step to promoting the democratization of data, they are, in our view, focused on the data provider. Boeckhout et al. (2018, p. 931) argue that “even though the principles create a powerful platform for furthering data sharing and improving data stewardship, they do not address

the normative issues and challenges associated with data sharing.” Strict check-listed adherence to the FAIR principles is a necessary but insufficient first step. Data democratization is a more holistic, comprehensive view of a process to make data discoverable, accessible, equitable, and usable.

Although not mentioned in our Research Topic, the CARE principles are a seminal example of shifting the focus from data providers to data consumers and moving beyond FAIR (Carroll et al., 2021). Developed by and for indigenous communities, these principles promote data ecosystems that provide **collective benefit**, where the **authority to control** the data resides with the data subjects and where there is a recognized **responsibility** to engage respectfully with data subjects. In addition, the **ethics** of the data subjects should inform data use.

Author contributions

KB and LJ drafted the content of this document. MK, TV, and NM provided edits. KB contributed figures. All authors contributed to the article and approved the submitted version.

Acknowledgments

The topic editors thank the authors that contributed papers to this topic. We also thank the reviewers for their dedication and time invested in providing invaluable feedback to the authors.

Conflict of interest

KB was employed by Environmental Systems Research Institute (ESRI).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Baker, K. S., and Mayernik, M. S. (2020). Disentangling knowledge production and data production. *Ecosphere* 11, e03191. doi: 10.1002/ecs2.3191
- Boeckhout, M., Zielhuis, G. A., and Bredenoord, A. L. (2018). The FAIR guiding principles for data stewardship: fair enough? *Euro. J. Hum. Genet.* 26, 931–936. doi: 10.1038/s41431-018-0160-0
- Carroll, S. R., Herczog, E., Hudson, M., Russell, K., and Stall, S. (2021). Operationalizing the CARE and FAIR principles for indigenous data futures. *Sci Data* 8, 108. doi: 10.1038/s41597-021-00892-0
- Mayer-Schönberger, V., and Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston, MA: Houghton Mifflin Harcourt.
- Taylor, H. L. Jr. (1997). No more ivory towers: connecting the research university to the community. *J. Plan. Liter.* 11, 327–332. doi: 10.1177/088541229701100304
- Tilly, C. (2001). Mechanisms in political processes. *Ann. Rev. Polit. Sci.* 4, 21–41. doi: 10.1146/annurev.polisci.4.1.21
- Tonn, B., English, M., and Travis, C. (2000). A framework for understanding and improving environmental decision making. *J. Environ. Plan. Manage.* 43, 163–183. doi: 10.1080/09640560010658
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* 3, 160018. doi: 10.1038/sdata.2016.18