Poirier, L and Costelloe-Kuehn, B. 2019. Data Sharing at Scale: A Heuristic for Affirming Data Cultures. *Data Science Journal,* 18: 48, pp. 1–7. DOI: https://doi.org/10.5334/dsj-2019-048

ESSAY

Data Sharing at Scale: A Heuristic for Affirming Data Cultures

Lindsay Poirier[1] and Brandon Costelloe-Kuehn[2]

[1] University of California Davis, US

[2] Rensselaer Polytechnic Institute, US

Corresponding author: Lindsay Poirier (lnpoirier@ucdavis.edu)

Addressing the most pressing contemporary social, environmental, and technological challenges will require integrating insights and sharing data across disciplines, geographies, and cultures. Strengthening international data sharing networks will not only demand advancing technical, legal, and logistical infrastructure for publishing data in open, accessible formats; it will also require recognizing, respecting, and learning to work across diverse data cultures. This essay introduces a heuristic for pursuing richer characterizations of the "data cultures" at play in international, interdisciplinary data sharing. The heuristic prompts cultural analysts to query the contexts of data sharing for a particular discipline, institution, geography, or project at seven scales – the meta, macro, meso, micro, techno, data, and nano. The essay articulates examples of the diverse cultural forces acting upon and interacting with researchers in different communities at each scale. The heuristic we introduce in this essay aims to elicit from researchers the beliefs, values, practices, incentives, and restrictions that impact how they think about and approach data sharing – not in an effort to iron out differences between disciplines, but instead to showcase and affirm the diversity of traditions and modes of analysis that have shaped how data gets collected, organized, and interpreted in diverse settings.

Keywords: data sharing; data culture; ethnography; data friction; metadata

Introduction

In the 1980s, the European Organization for Nuclear Research (CERN) was the most prominent particle physics laboratory in the world and at the cutting edge of coordinating international scientific research. Herwig Schopper (2014), Director-General for CERN, 1981–1988, describes the time as provoking "a new 'sociology' for international scientific collaboration;" with over 30 countries participating in experiments, the challenges for keeping track of researchers, workflows, and scientific data were enormous.

CERN hired Tim Berners-Lee as a contract programmer in 1980. To help keep track of projects, he toyed with designing Enquire – a knowledge organization system that enabled users to organize their data by creating links between documents stored in separate locations. Berners-Lee landed a fellowship in the Data Acquisition and Control division in 1983 at a time when CERN was upgrading its computing infrastructure to better network globally distributed researchers in laboratories that each followed their own methods, used their own operating systems, and often spoke different languages. In describing the systems that were proposed for addressing these challenges, Berners-Lee (1999: 15) writes:

> I had seen numerous developers arrive at CERN to tout systems that "helped" people organize information. They'd say, "To use this system all you have to do is divide all your documents into four categories" or "You just have to save your data as a WordWonderful document" or whatever. I saw one protagonist after the next shot down in flames by indignant researchers because the developers were forcing them to reorganize their work to fit the system. I would have to create a system with common rules that would be acceptable to everyone. This meant as close as possible to no rules at all.

The challenge was not to compel researchers to adopt a new standard; instead the challenge was learning to recognize and respect the different data cultures that guided how diverse researchers approached their work.[1] Berners-Lee's Enquire eventually evolved into a proposal for what became the World Wide Web – perhaps the most widely adopted information infrastructure in the world, in large part because the system has very few rules prescribing how users should organize their knowledge within it.

Today, we are contending with a sociology for international scientific collaboration on a much larger scale. Addressing the most pressing contemporary social, environmental, and technological challenges will require integrating insights and sharing data across disciplines, geographies, and cultures. Research into the socio-technical challenges of data sharing has begun to characterize complications that arise as researchers in different communities work to align their data cultures (Borgman, 2012). The process of integrating complex and heterogeneous data generated in different geographies, according to different disciplinary standards, and motivated by different epistemic commitments and incentive structures, can produce "friction," demanding that researchers make compromises to find common ground (Edwards et al. 2011). Different disciplines may speak different "languages" making it difficult to devise shared schemas and ontologies. Perhaps most notably, researchers in different settings often have diverse rationales for valuing data preservation, contextualization, integration, and dissemination. Strengthening international data sharing networks will not only demand advancing technical, legal, and logistical infrastructure for publishing data in open, accessible ways; it will also require recognizing, respecting, and learning to work across diverse data cultures. As Berners-Lee observed of collaborative research practice at CERN in the 1980s, prescriptively forcing researchers to reorganize their work to fit a standard limits adoption and collaboration. This essay, informed by our work exploring diverse data sharing communities at the Research Data Alliance (RDA), will introduce a heuristic we've developed in order to pursue richer characterizations of the "data cultures" at play.

The RDA (rd-alliance.org) is an international community of researchers aiming to design and sustain the socio-technical infrastructure needed to enable open research data sharing across geographies and disciplines. We became involved in the RDA as cultural anthropologists looking to advance frameworks for data sharing in our own field. However, we found that RDA's bi-annual plenaries are rich ethnographic fieldsites for examining how commitments guiding scientific practice inform how data gets produced, structured, and semantically-enriched, and how researchers in different communities think about, value, and practice data sharing.[2] We have been developing the heuristic presented in this paper since 2017, when we were asked, as ethnographers of data practice, to collaborate in a session at the RDA Plenary 10 in Montreal on addressing barriers to adoption of RDA outputs. We devised a series of questions that researchers might ask themselves in considering the aims, assumptions, and commitments brought to their work, their congruence with the RDA output, and the infrastructure and incentives available for enabling adoption. We revised the questions in 2019 for a workshop addressing the socio-technical challenges of international and interdisciplinary data sharing.

Data Sharing at Scale

Global, interdisciplinary data sharing is a cultural system – one that assembles many actors, institutions, technologies, and frameworks. It is a system animated by a diverse set of forces operating at many different locations and across many different scales. Understanding this system will demand that we learn to simultaneously observe the multiple forces acting upon and interacting with researchers and the data they produce.

Scale has historically referred to many things. For instance, geographers may refer to spatial scales that designate different geo-political boundaries; computer scientists may refer to nested IT infrastructures (i.e. data, computers, networks, the internet, etc.). Scholars studying the history and social dynamics of information infrastructures have shown how examining data systems across macro, meso, and micro scales of society can reveal the complexity of socio-technical forces in shaping knowledge, modernity, and computing history (Edwards 2004; Acker 2015). When we refer to scale in the context of cultural systems, we aim to evoke the frames of reference – diverse in their breadth and modes of cultural ordering – that anthropologists have crafted to examine how culture is enacted and mediated. When we talk of frames of reference, we are not referring to particular places, contexts, or phenomena, but rather, focusing devices that order

[1] We refer to "researchers" here quite expansively to denote any individual involved in the collection, designation, analysis, stewardship, and/or use of empirical data. This may refer to scientists, humanists, industrial analysts, and government actors in a variety of locations.

[2] In positioning the RDA as a fieldsite, we have methodologically employed what Fassin and Rechtman (2009: 11) refer to as "observant participation" in the study of data cultures. Ethnographic observation has come second to and has been inflected by our own participation in the organization. This research was carried out under the approval of Institutional Review Boards at Rensselaer Polytechnic Institute and the University of California Davis.

what the ethnographer pays attention to (i.e. customs, politics, discourse, etc.). Kim Fortun (2009: 75–6), a cultural anthropologist that has written extensively on theories of ethnographic practice, argues that "scale is a heuristic, which, like all heuristics, provides a way of seeing that frames and orients perspective. At its best, scale provides a way to see many types of action in motion at once, evoking a sense of the system at hand." Fortun proposes seven "strata" (including meta, macro, meso, micro, techno, nano, and natural) for guiding cultural analysis and argues that examining cultural systems across these scales can help constitute the "meta-data" needed to make sense of the systems.

In this paper, we adapt Fortun's approach in order to outline a heuristic for examining the data cultures of different research communities. We do so bearing in mind that all heuristics for framing perspective delimit insight, and that phenomena are constantly crossing these scales that are only analytically (and uneasily) separable. Further, the issues that emerge within each strata, themselves involve cultural forces playing out across diverse spatial, temporal, and infrastructural scales. Data cultures can be characterized as a cumulation of phenomena oscillating between various sites at different times in the context of the different cultural frames of reference we outline below. While the heuristic we introduce provides but one way of unpacking complex cultural systems, it does serve as a starting point, enabling comparative perspective between fieldsites, informing examinations of how cultural systems evolve, and signposting the various forces that can impact data sharing. The heuristic we outline below (**Figure 1**) can serve as a template for querying researchers and examining data cultures within the context of a particular discipline, institution, geography, or project.[3]

META	In what ways have researchers in this community developed discourse around the promise of open science – particularly at a time when global support for research itself is under attack in many settings? To what extent have researchers in this community developed discourse that encourages investment in the development and maintenance of research infrastructure?
MACRO	What guidelines or frameworks are in place to support researchers in this community in discerning how to share data within the parameters of existing laws – such as copyright and requirements for human subjects research? To what extent are there financial structures in place to support and sustain the development and implementation of data infrastructure in this community?
MESO	To what extent do promotional incentives encourage researchers in this community to participate in data sharing and infrastructure development? Do the governance mechanisms that researchers in this community work within support collaborative research projects? Which external organizations are available to support researchers in this community in adopting and implementing data sharing infrastructure?
MICRO	In what ways does data sharing change the research workflow in this community? Are researchers in this community prepared to work collaboratively? Are there guidelines and frameworks available to researchers in this community that prepare them to integrate recommended data practices into their day-to-day work?
TECHNO	To what extent are the technologies and infrastructures needed to support data sharing accessible and affordable to researchers in this community? Do they have the time to learn, build, and implement them? How accessible are the guidelines regarding the implementation of tools and practices for data sharing to the researchers in this community?
DATA	To what extent do the data sharing architectures available to researchers in this community respect and preserve the unique epistemologies, commitments, and modes of analysis they bring to their research? How are the vocabularies researchers use in this community distinct from those used in other communities? What translational work must researchers perform when adopting shared taxonomies?
NANO	What beliefs and values motivate researchers in this community to engage in data sharing and collaborative work? What do they hope to gain from collaborative perspective? How do educational programs for researchers in this community teach the value and practice of research data sharing?

Figure 1: A Multi-Scaled Heuristic for Examining and Affirming Data Cultures.

[3] As RDA working groups plan for the design of new data sharing infrastructure, this heuristic may be used as a template for interviews or surveys designed to elicit from diverse communities the data cultures that shape their thinking, their practice, and the resources available to them. For communities seeking to adopt RDA outputs, the heuristic may be a tool to help analyze and make sense of the diverse cultural forces that shape possibilities for infrastructural implementation.

Meta: Discursive

Meta-level analysis, or the way forces characterized by the scales below get talked about, queries the dominant discourse and counter-narratives guiding how a community values data sharing. Different communities have differently prioritized investments in open science and the data infrastructure needed to support it. While advocates in all disciplines may struggle to communicate the value of open science to their administrators, funders, and peers, the conversation has advanced much further in certain disciplines. Early policies regarding data sharing in the Human Genome Project propelled discourse around open science to the forefront in genomics research (Kaye et al. 2009), setting global expectations that sequencing data would be publicly available. Meta-level analysis may also consider how geopolitics shape different discourses around open science. In Low- and Middle-Income Countries, some researchers have voiced concerns that opening data can exacerbate existing global inequalities by heralding in new opportunities for extraction (Serwadda et al. 2018).

Macro: Legal, Political-economic, and Financial

Macro-level analysis attends to the financial and legal structures that support the work of data sharing communities and organizations. While globally funders are increasingly requiring researchers to deposit data in public access repositories, their willingness to fund data infrastructure development and maintenance differs drastically. For instance, the European Union has consistently prioritized Open Science in their research and innovation funding programmes (including Horizon 2020 and Horizon Europe), supporting the propagation of consulting bodies such as GoFAIR and the European Research Infrastructures Initiative, as well as domain-specific bodies such as DARIAH and the Europrean Marine Biological Resource Center. In other parts of the world, financial resources supporting research infrastructure are not as readily available.

Macro-level analysis also considers differences in legal structures that disciplines operate within. For instance, social scientists must comply with country-specific research ethics laws, and public health researchers must comply with country-specific laws for safeguarding a patient's privacy with respect to medical information (Panhui et al. 2014). Such legislation restricts the degree to which researchers can engage in interdisciplinary data sharing.

Meso: Organizational

Meso-level analysis focuses attention on organizations and networks. Ethnographic research examining barriers to data sharing and collaborative practice has often focused here, examining how a lack of incentives for sharing data, engaging in collaborative work, and investing time in the design and maintenance of research infrastructure has inhibited participation amongst some communities in data sharing networks (Edwards et al. 2011; Borgman 2012). Notably, this dearth of impetus is felt disproportionately in certain disciplines (such as those where collaborative publications are discouraged), in certain institutions (such as those that do not count the design and maintenance of research infrastructure as "service" in tenure cases), and at different career stages. Early career researchers, for instance, are often concerned that publishing their data in open repositories may lead to their research findings being "scooped up" before they have an opportunity to establish their credibility as scholars (Bahlai et al. 2019).

Micro: Research Practices and Customs

Micro-level analysis focuses attention on customs and practice – both data practice and research practice. For instance, Broom, Cheshire, and Emmison (2009) argue that data sharing imperatives can undermine qualitative researchers' understanding of what it means to "do" qualitative research. Since many qualitative researchers understand their data to be so contextually situated and containing indeterminacies that require an "expert" eye to be perceived, they have voiced concern that adopting data publishing into their workflows erodes the integrity of doing qualitative research.[4] On the other hand, in some research communities, data sharing procedures have already fundamentally reshaped and advanced research practice. Leonelli and Ankeny (2012) have documented how the development of community databases for model organisms has provoked a shift in biological research practice away from single species analysis towards more comparative, cross-species research. Micro-level analysis also considers the time researchers can devote to implementing data sharing protocols (Acord and Harley 2013).

[4] Ann Zimmerman (2008) similarly demonstrates how the locally-situated knowledge ecologists acquire in fieldwork can be difficult to translate into public data through available standards and thus often get left behind.

Techno: Technological

Techno-level analysis attends to the availability, accessibility, and fitness of technologies and data standards for supporting data sharing practices. In certain research communities, suites of data sharing technologies and standards have already been networked into data repositories that can support domain-specific data publishing and management without requiring advanced data infrastructure expertise on the part of data depositors. While certain communities can submit their data to domain-specific data repositories (such as the earth data repositories networked through DataONE) or region-specific data repositories (such as Research Data Australia) researchers in other data domains and regions do not necessarily have access to infrastructure designed specifically to meet their data management needs (Tenopir et al. 2011). Further, at some institutions, librarians and information scientists are more readily available for helping researchers address the technical challenges of implementing data infrastructure to support sharing and management. The California Digital Library, for instance, supports the entire University of California system in addressing challenges of data curation and preservation.

Data: Data Architecture

Data-level analysis focuses attention on data architecture and configuration and the extent to which the logics of data sharing standards embody the assumptions that different communities bring to their research practice. Some disciplines, such as evolutionary biology and chemistry, have a long tradition of grouping and sorting entities according to various taxonomic systems, while in other research domains the very act of discretizing knowledge runs counter to researchers' epistemological commitments. For instance, Pulsifer et al. (2011) have shown how designing formal data management processes for documenting Inuit knowledge posed a unique set of challenges because the complexity and dynamism of relationships characterized in indigenous narratives are not amenable to the standardized data practices Western science promotes. At the data level, we consider the ways in which the complexity of a community's context-specific knowledge inevitably gets transformed to fit into data sharing infrastructures (because all infrastructures structure and delimit entities and their flows), and how new data infrastructures can be designed to better represent diverse knowledge forms.

Nano: Individual Beliefs and Values

Nano-level analysis focuses attention on the embodied beliefs researchers bring to data sharing practice – why they value data sharing and what they hope to get out of collaboration.[5] For instance, in many of the natural sciences, an oft-cited motivator for advancing robust data sharing infrastructure is to confront a crisis of the scientific method – that a great deal of published research has not been documented in such a way that its results can be reproduced (Jasny et al. 2011). The humanities, however, are typically guided by a different set of motivations; in cultural anthropology, for instance, sharing data and encouraging its reuse can help ensure that interpretations of cultural objects are not univocal but instead represent an array of perspectives (Poirier et al. 2019). Unpacking the ideological underpinnings of different data cultures highlights the fallacy of designing one-size-fits-all solutions; to advance global and interdisciplinary data sharing in an inclusive way, we need infrastructures and policies that affirm the wide array of stakes that animate different research communities' investment in data sharing and open science.

Conclusion: Affirming Culture

> *"We all tend to treat diversity as a problem. It's here to stay and it's beautiful."*
> *–Dr. Devika P. Madalli, RDA Technical Advisory Board and Consultant, RDA Plenary 13, Philadelphia, PA (April 2019)*

Over the past five years, we have found ourselves in numerous data sharing workshops, meetings, and plenaries where "culture" gets cast as a problem to be fixed. We have heard folks say "if only we could get everyone to speak the same language" or "if only we can align different data sharing cultures." However, in working to tame culture in data sharing practices, there is great risk that the nuances that make interdisciplinary research so robust and appropriate for tackling complex, multi-scaled, and multi-dimensional problems

[5] Of all the scales presented here, the distinction between the first (meta/discursive) and the last (nano/individual beliefs and values) is perhaps the most difficult to stabilize. To understand this particular crossing of scale, we find Althusser's (1971) concept of interpellation useful, which in its simplest terms involves the processes by which ideology (embodied in the various scales presented here) conditions and even constitutes individual subjects' identities, beliefs, and values.

will be eclipsed. The heuristic we introduce in this essay aims to elicit from researchers the beliefs, values, practices, incentives, and restrictions that impact how they think about and approach data sharing – not in an effort to iron out differences between disciplines, but instead to showcase and affirm the diversity of traditions and modes of analysis that have shaped how data gets collected, organized, and interpreted in diverse settings.

This is a key and often overlooked component within efforts to design and implement data sharing policy and infrastructure. While designers of data sharing infrastructure often attempt to gather feedback from diverse domain communities when developing new standards, tools, or frameworks, we have found that they attempt to structure the feedback in ways that control for difference. For instance, designers may distribute use case templates that ask representatives in different data domains to outline scenarios, triggers, motivations, goals, costs, and risks involved in a particular data practice. While the structure of the document allows the designers to compare and contrast key factors of a data practice across communities, it also presets the conditions for comparative perspective in ways that can eclipse more fundamental differences – such as why the researchers value data in the first place, how they leverage theory, what they hope to gain through collaboration, the assumptions they have about language and representation, and the unique historical and institutional conditions that have shaped their communities. These considerations can have a profound effect on how data sharing practices get taken up in different settings. Studying data cultures at scale can help to foreground these often neglected considerations, animating capacity to design data sharing infrastructure and policies that are not only acceptable to everyone, but also affirm and respect the diversity of cultures that guide global and interdisciplinary research practice.

Acknowledgements

The ideas developed in this essay were supported through the RDA/US Data Share fellowship sponsored through a grant from the Alfred P. Sloan Foundation. We also thank Kim Fortun and Mike Fortun for helping us scope our involvement in RDA and reviewing multiple iterations of this heuristic.

Competing Interests

The authors have no competing interests to declare.

References

Acker, A. 2015. "Toward a Hermeneutics of Data." *IEEE Annals of the History of Computing*, 37(3): 70–75. DOI: https://doi.org/10.1109/MAHC.2015.68

Acord, SK and **Harley, D.** 2013. Credit, time, and personality: The human challenges to sharing scholarly work using Web 2.0. *New Media Society*, 15: 379–397. DOI: https://doi.org/10.1177/1461444812465140

Althusser, L. 1971. Ideology and Ideological State Apparatuses (Notes towards an Investigation). In: *Lenin and Philosophy and Other Essays. Monthly Review Press.*

Bahlai, C, Bartlett, LJ, Burgio, KR, Fournier, AMV, Keiser, CN, Poisot, T and **Whitney, KS.** 2019. Open Science Isn't Always Open to All Scientists. *American Scientist*. URL https://www.americanscientist.org/article/open-science-isnt-always-open-to-all-scientists (accessed 6.5.19). DOI: https://doi.org/10.1511/2019.107.2.78

Berners-Lee, T. 1999. Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor. San Francisco: Harper Business.

Borgman, CL. 2012. The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63: 1059–1078. DOI: https://doi.org/10.1002/asi.22634

Broom, A, Cheshire, L and **Emmison, M.** 2009. Qualitative Researchers' Understandings of Their Practice and the Implications for Data Archiving and Sharing. *Sociology*, 43: 1163–1180. DOI: https://doi.org/10.1177/0038038509345704

Edwards, P, Mayernik, MS, Batcheller, A, Bowker, G and **Borgman, C.** 2011. Science Friction: Data, Metadata, and Collaboration. *Social Studies of Science*, 41: 667–690. DOI: https://doi.org/10.1177/0306312711413314

Edwards, PN. 2004. Infrastructure and Modernity: Force, Time, and Social Organization in the History of Sociotechnical Systems. In: Misa, TJ, Brey, P and Feenberg, A (eds.), *Modernity and Technology*, 185–266. MIT Press.

Fassin, D and **Rechtman, R.** 2009. *The Empire of Trauma: An Inquiry Into the Condition of Victimhood.* Princeton University Press.

Fortun, K. 2009. Scaling and Visualizing Multi-sited Ethnography. In: Falzon, M-A (ed.), *Multi-Sited Ethnography: Theory, Praxis and Locality in Contemporary Research*, 73–86. Surrey: Ashgate.

Jasny, BR, Chin, G, Chong, L and **Vignieri, S.** 2011. Again, and Again, and Again …. *Science*, 334: 1225–1225. DOI: https://doi.org/10.1126/science.334.6060.1225

Kaye, J, Heeney, C, Hawkins, N, de Vries, J and **Boddington, P.** 2009. Data sharing in genomics — re-shaping scientific practice. *Nature Reviews Genetics*, 10: 331–335. DOI: https://doi.org/10.1038/nrg2573

Leonelli, S and **Ankeny, RA.** 2012. Re-thinking organisms: The impact of databases on model organism biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, Data-Driven Research in the Biological and Biomedical Sciences*, 43: 29–36. DOI: https://doi.org/10.1016/j.shpsc.2011.10.003

Poirier, L, Fortun, K, Costelloe-Kuehn, B and **Fortun, M.** 2019. Metadata, Digital Infrastructure, and the Data Ideologies of Cultural Anthropology. In: Crowder, J, Fortun, M, Besara, R and Poirier, L (eds.), *Anthropological Data in the Digital Age: New Possibilities-New Challenges*. Palgrave Macmillan.

Pulsifer, PL, Laidler, GJ, Taylor, DRF and **Hayes, A.** 2011. Towards an Indigenist data management program: Reflections on experiences developing an atlas of sea ice knowledge and use. *The Canadian Geographer/Le Géographe canadien*, 55: 108–124. DOI: https://doi.org/10.1111/j.1541-0064.2010.00348.x

Schopper, H. 2014. Viewpoint: The 1980s: Spurring collaboration – CERN Courier. URL http://cerncourier.com/cws/article/cern/56613.

Serwadda, D, Ndebele, P, Grabowski, MK, Bajunirwe, F and **Wanyenze, RK.** 2018. Open data sharing and the Global South—Who benefits? *Science*, 359: 642–643. DOI: https://doi.org/10.1126/science.aap8395

Tenopir, C, Allard, S, Douglass, K, Aydinoglu, AU, Wu, L, Read, E, Manoff, M and **Frame, M.** 2011. Data Sharing by Scientists: Practices and Perceptions. *PLOS ONE*, 6. DOI: https://doi.org/10.1371/journal.pone.0021101

van Panhuis, WG, Paul, P, Emerson, C, Grefenstette, J, Wilder, R, Herbst, AJ, Heymann, D and **Burke, DS.** 2014. A systematic review of barriers to data sharing in public health. *BMC Public Health*, 14: 1144. DOI: https://doi.org/10.1186/1471-2458-14-1144

Zimmerman, AS. 2008. New Knowledge from Old Data The Role of Standards in the Sharing and Reuse of Ecological Data. *Science Technology Human Values*, 33: 631–652. DOI: https://doi.org/10.1177/0162243907306704

How to cite this article: Poirier, L and Costelloe-Kuehn, B. 2019. Data Sharing at Scale: A Heuristic for Affirming Data Cultures. *Data Science Journal,* 18: 48, pp. 1–7. DOI: https://doi.org/10.5334/dsj-2019-048

Submitted: 01 July 2019 **Accepted:** 12 September 2019 **Published:** 30 September 2019

]u[*Data Science Journal* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS