

Understanding data culture/s: Influences, activities, and initiatives

Gillian Oliver¹  | Jocelyn Cranefield²  | Spencer Lilley²  |
Matthew J. Lewellen² 

¹Department of Human Centered Computing, Monash University, Clayton, Victoria, Australia

²School of Information Management, Victoria University of Wellington, Wellington, New Zealand

Correspondence

Gillian Oliver, Department of Human Centered Computing, Monash University, Wellington Road, Clayton, VIC 3800, Australia.

Email: gillian.oliver@monash.edu

Abstract

Data culture/s as a research topic has begun to attract attention from a wide range of disciplines, albeit with inconsistent application of definitions, dimensions, and applications. This work builds on a call to investigate data culture/s within the information studies domain as a topic related to, but distinct from, information culture. The purpose of this study is to explore what is known about data culture/s in greater depth. We apply a retroductive approach to select and consider likely dimensions, inputs, and aspects of data culture/s in order to further map this construct to the literature, and thereby highlight gaps and opportunities to add to this body of knowledge. The initial candidate dimensions explored below include data-related skills and attitudes, data sharing, data use/reuse, data ethics and governance, and a specific focus on Indigenous perspectives to provide insights on why and how a group may contest the emergent dominant discourse of data culture/s. Our conclusion highlights areas needing further research to fully define and examine the dimensions, inputs, and aspects of data culture/s, and calls for greater understanding and engagement with data culture/s from the information studies community.

1 | INTRODUCTION

The emerging topic of data culture/s has been identified and discussed across a broad spectrum of academic disciplines, ranging from the humanities to the applied and pure sciences albeit without any evidence of a shared understanding of the meaning of the concept (Oliver et al., 2023). For the purpose of this article, we adopt Oliver et al.'s (2023) definition of data culture, based on a synthesis of definitions used in published research:

Data culture/s are the social, technical, and cultural characteristics, values and practices

that influence/determine the nature of data production, generation, acquisition, cultivation, use, curation, preservation, sharing, and reuse by individuals, organisations, governments, and societies. They may co-exist and compete at multiple levels and are dynamic and normative in nature (Oliver et al., 2023).

The purpose of this article is to explore contributing dimensions, inputs, and aspects of data culture/s in greater depth by identifying and analyzing peer-reviewed literature which has been thematically linked to data

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Journal of the Association for Information Science and Technology* published by Wiley Periodicals LLC on behalf of Association for Information Science and Technology.

culture/s. This investigation includes the role of data-related skills and attitudes, data sharing, data use/reuse, data ethics and governance, and a specific focus on Indigenous perspectives in explaining data culture/s. Our emphasis is on the literature published by the information studies and information systems disciplines, but we also include any significant publications which are likely to be influential in developing research agendas in those two information-related disciplines.

We begin with a summary of findings from a foundational multi-disciplinary literature review which provided us with insight into the extent of issues and concerns relating to data culture/s, and thus served to establish the parameters for the current study. We then explain the methodology followed to identify the thematic candidates covered in this article, which given the number of peer-reviewed publications available in targeted areas was selective rather than aiming for comprehensiveness. Subsequent sections consider data-related skills and attitudes, including data literacy and fluency, data sharing, data use/reuse, data ethics and governance, and last but by no means least, Indigenous perspectives about data sovereignty. Our conclusion highlights areas needing further research, and calls for greater understanding and engagement with data culture/s from the information studies community.

2 | PRELIMINARY REVIEW

We undertook a preliminary scoping review of the literature to ascertain the extent of the research already undertaken which specifically addressed the concept of data culture/s (Oliver et al., 2023). This initial search was restricted to English-language peer-reviewed publications that specifically included the phrase *data culture/s* in the title, abstract, or body of the article, which resulted in 80 papers published between 2004 and 2021. The range of disciplines concerned with data culture/s encompassed the humanities, arts, and social sciences, as well as science, technology, engineering, and mathematics (STEM), but with surprisingly little representation from information studies researchers. We reviewed the full text of each paper to identify how this diverse cohort of researchers understood and portrayed the concept of data culture/s, and we also applied a heuristic developed with the Research Data Alliance (Poirier & Costelloe-Kuehn, 2019) as an analytical tool to identify the level at which research was focused. We found that there was no unified understanding of what the construct *data culture* or *data cultures* represented, and that most publications focused on organizational settings. There was minimal research investigating data-related values and attitudes, and no

discussion or acknowledgment of Indigenous perspectives. The literature review demonstrated that most problems were perceived at the meso (organizational) and micro (practices and customs) levels. Notably, much emphasis was given to the need for people such as employees, students, or researchers to have digital and data expertise and understanding, which can be understood as data literacy or data fluency. This is an area central to information studies, and one where there is indeed a lot of research (further discussed in the data-related skills and attitudes sections below), but it is doubtful whether this body of research would be apparent to researchers coming from other disciplines.

The diversity of disciplines concerned with data culture/s provided insight into growing awareness of the strategic importance of this topic; it was not in fact possible to identify any major discipline that had not undertaken research addressing data culture/s to a greater or lesser extent. However, it was of concern that information studies and information systems researchers did not appear to have a strong or obvious presence in research teams, suggesting that the unique perspectives and expertise we bring is not known about outside our own disciplinary bubble. Our awareness of the contributions from information studies and information systems research concerned with specific areas that influence and shape data culture/s motivated this further literature review. So, in this article, we embark on an examination of the thematically linked topics representing the dimensions, inputs, and aspects of data culture/s, in an attempt to bring together the rich body of relevant research that has been underway since the early 2000s, under the banner of data culture/s. In so doing, we highlight the specialist contributions from the information-related disciplines to an area of concern that is attracting attention from the humanities, arts, social sciences, science, technology, engineering, and mathematics disciplines, and which will become more and more important over time as data becomes inextricably interwoven into all aspects of everyday life.

3 | METHODOLOGY

In this section, we relate the methodological approach used to derive our initial thematic candidate topics, which will be used to more fully explore and explain the dimensions, inputs, and aspects of the data culture/s construct.

This work builds upon a scoping literature review completed in March 2022 that builds a case for investigating data culture/s within the information studies domain as a topic related to, but distinct from, information

culture (Oliver et al., 2023). In the literature, there is a clear differentiation between the concepts of data, information, and knowledge (Wilson, 2002). Furthermore, acknowledging the demonstrated links between information culture and management's openness to change and innovation (Ginman, 1993), there appears to be some gaps in the literature when considering organizational change and innovation without taking data (and data culture/s) into account.

Our goal then is to explore additional dimensions, inputs, and aspects of data culture/s in order to further map this construct to the literature, and thereby highlight gaps and opportunities to add to this body of knowledge. To that end, we have applied a retroductive approach to select and consider likely aspects of data culture/s for initial study, and we have sought to recursively explore further into the literature in an attempt to better frame and define this study area (Muganda, 2013). Use of a retroductive approach “does not entail an ‘anything goes’ approach to the generation and evaluation of empirical evidence made in its name. [...] [T]he ultimate tribunal of experience is the degree to which its accounts provide plausible and convincing explanations of carefully problematized phenomena for the community of social scientists” (Howarth & Stavrakakis, 2000, p. 7).

To arrive at the selection of component topics, an analysis of the initial scoping literature review (Oliver et al., 2023) was considered through a heuristic developed by Poirier and Costelloe-Kuehn (2019, pp. 4–5). This heuristic was inspired by a cultural anthropology discourse that leverages a number of strata or levels to guide analysis (Fortun, 2016). From these strata, we were able to extract a number of thematic candidates within the data culture/s domain for further exploration. These included:

- Data-related skills and attitudes (micro)
- Data sharing (meso)
- Data use/reuse (techno/data)
- Data ethics and governance (meso)
- Indigenous perspectives (macro)

These thematic areas represent an initial investigation into aspects of data culture/s, and are not meant to constrain this investigation; nor in naming these do we claim a complete understanding of data culture/s.

4 | CULTURES OF DATA-RELATED SKILLS AND ATTITUDES

In this section, we consider the literature discussing the need to build data-related competencies and capabilities

as well as other factors influencing attitudes and behaviors in relation to data.

4.1 | Data literacy and data fluency

The possession of data-related skills and attitudes that contribute to data culture/s require an underlying familiarity with data, that is, foundational data literacy. The earliest peer-reviewed article concerned with data literacy we identified was published in 2002, and was concerned with developing the appropriate skills in high school students (Yan et al., 2002). The importance of data literacy for global innovation is now recognized by the Organization for Economic Development (OECD) and included as a core competency in their curriculum for 2030 (Organisation for Economic Co-operation and Development (OECD), 2022).

It is not surprising that there has been considerable attention paid to the concept of data literacy in information studies, notably linking it to the existing body of research concerned with information literacy (see, e.g., Condon & Pothier, 2022; Koltay, 2015, 2017). Definitions of data literacy demonstrate this genealogy, for instance:

[...] a specific skill set and knowledge base, which empowers individuals to transform data into information and into actionable knowledge by enabling them to access, interpret, critically assess, manage, and ethically use data (Koltay, 2017, p. 10).

Initially, at least, data literacy appeared to be exclusively considered in the context of research (Koltay, 2015, 2017), with data training and education initiatives considered most appropriately targeted at would-be researchers and data specialists (Koltay, 2015, p. 411), but subsequently, attention has extended to diverse workplaces (Pothier & Condon, 2020).

Consideration in the much broader context of the *datafication* of society, however, results in more expansive conceptualization. For instance, “Data literacy [...] involves both critical understandings of the technological infrastructure and the political economy of digital platforms, as well as strategies and tactics to manage and protect privacy and resist being profiled and tracked” (Pangrazio & Sefton-Green, 2020, p. 214), where data literacy is seen as having the potential to protect individuals from the negative consequences of data in society. Similarly, Gray and colleagues formulate the idea of data *infrastructure* literacy as a means to promote critical consideration of datafication (Gray et al., 2018). The

potential for positive benefits to society from datafication were prominently promoted with the formation of the United Nations' Data Revolution initiative in 2014, which resulted in a call for global data literacy (The United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG), 2014).

The concept of data fluency has also emerged, accompanied by a similar lack of definitional precision. Data fluency has been described as the ability and confidence to select appropriate software tools for data analysis (Kirkwood, 2016), and has been differentiated from data literacy on the basis that literacy implies a novice status whereas fluency implies having more expertise—for example, the nuanced understanding necessary in order to “...ask informed questions to make highly articulated data-driven decisions” (Kennedy-Clark & Reimann, 2022, p. 44). Perhaps the reality is that the terms data literacy and data fluency are used interchangeably as reported by Capdarest-Arest and Navarro (2021), particularly in situations where it is important to avoid the negative connotations of illiteracy. Regardless of the terminology used, it is very clear that during the last decade, in particular, considerable attention has been directed toward emphasizing the importance of people having the requisite knowledge and skills to effectively interact with data, both in the context of scientific research or in the context of navigating everyday life.

Given the recognition of the critical importance of data literacy and data fluency, it is not surprising that attention has focused on how these concepts can be assessed and measured. Canada's National Statistics Agency has considered the meaning of data literacy in the context of the public sector, providing an overview of existing objective and self-assessment tools (Bonikowska et al., 2019). The authors distinguish between the skills required by data specialists and those required by nonexpert users in the public sector, and warn against overreliance on self-assessment tools because of the risks of generating a distorted picture of skills (Bonikowska et al., 2019, p. 15).

A maturity model developed to evaluate the data literacy of employees in a nongovernmental organization provides the opportunity to look at one of these tools in more detail. The model distinguishes between four levels: from *uncertainty*, through to *enlightenment*, *certainty*, and *data fluency* at the highest level (Sternkopf & Mueller, 2018). A matrix is provided to evaluate what the authors describe as competencies according to these four levels. The so-called competencies include relatively simple skill-based tasks such as how to ask a question, find, get, and verify an answer; to much more diffuse concepts—namely, data culture, data ethics, and security (Sternkopf & Mueller, 2018,

p. 5053). The authors make the connection between data fluency and a data culture where “Psychological barriers of data have been brought down (e.g., insecurities, fear, resignation), and comfort around data is promoted. Higher-level management and project managers understand and support the importance of dedicated resources (time, budget, human resources) for data handling and conversion” (Sternkopf & Mueller, 2018, p. 5051).

5 | CULTURES OF DATA SHARING

Within information studies, the emergence of concerns about the sustainability of digital information and associated data curation objectives have motivated considerable research effort toward understanding attitudes and behaviors relating to data sharing and fostering data-sharing cultures. Patterns of data-sharing attitudes and behaviors form another aspect to be explored under the banner of data culture/s. As explained by Oliver and Harvey (2016, p. 96), “Collaboration is, in fact, firmly embedded in digital curation practice. Active management of data for current and future use relies on effective sharing of data, which in turn relies on agreement about and adoption of standards.” The extent of activity is clearly indicated by the existence of the Research Data Alliance (RDA). This is a global organization—in May 2022 consisting of 12,600 members from 145 countries (The Research Data Alliance (RDA), 2022a)—with the following vision and mission statements emphasizing the importance of data sharing culture/s:

The RDA Vision: Researchers and innovators openly share and re-use data across technologies, disciplines, and countries to address the grand challenges of society.

The RDA Mission: RDA builds the social and technical bridges that enable open sharing and re-use of data. (The Research Data Alliance (RDA), 2022b)

An earlier ARIST review considering data sharing in the academic research environment provides the state-of-the-art view from the first decade of the new millennium (Kowalczyk & Shankar, 2011). The authors defined data sharing as consisting of a number of complex challenges, which they distinguish into two categories: the practical *how to do it* issues faced by information professionals, and the broader *societal level concerns* about the nature of research itself, including freedom of access to the outcomes (Kowalczyk & Shankar, 2011, p. 249).

The challenges relating to data sharing have a long history in the context of international scientific endeavors

as described, for example, in an account of US and Soviet initiatives during the Cold War (Jacobsen et al., 2021). More recently, establishing a culture of widespread sharing of data within the international scientific research community has become a strategic priority, both between and within domains, and widespread sharing of research data has even been positioned as fundamental to human progress (Organisation for Economic Co-operation and Development (OECD), 2007, p. 3). There are various drivers for the elevated societal significance of data sharing. These include calls for interdisciplinary research teams to address grand challenges (Faniel & Zimmerman, 2011), the increasingly collaborative nature of research, the generation by research projects of large datasets whose value exceeds individual studies or programs, a shift towards valuing data-driven discovery as a mode of research (Thessen & Patterson, 2011), and a drive for greater efficiency, accountability, and return from public investment in research (Organisation for Economic Co-operation and Development (OECD), 2007). The vision and benefits of sharing research data were articulated in the Organisation for Economic Co-operation and Development's (OECD) (2007) Principles and Guidelines for Access to Research Data from Public Funding, by then OECD Secretary General Angel Gurría, who stated that, "access to research data increases the returns from public investment in this area; reinforces open scientific inquiry; encourages diversity of studies and opinion; promotes new areas of work and enables the exploration of topics not envisioned by the initial investigators" (Organisation for Economic Co-operation and Development (OECD), 2007, p. 3). In combination, these drivers and the perceived benefits of data sharing have led to the emergence of mandates for data sharing at the level of governments, organizations, and institutions, as well as the increased status of raw datasets in the research community (MacMillan, 2014). A data-sharing research culture is therefore seen as critical to scientific progress, accountability, and credibility. (The rationale for data sharing is strongly predicated on the assumption that data will be reused, an aspect of data culture/s that we consider below.)

Despite the fact that researchers recognize the high-level benefits of data sharing, there is a documented gap between demand and practice in sharing research data (Pampel & Dallmeier-Tiessen, 2014), with many scientists being reluctant to share data sets (MacMillan, 2014). Data sharing enables other scientists to reproduce and validate research results, diagnose methodological errors, and reduce unnecessary duplication (Borgman, 2012; MacMillan, 2014), but this can also be perceived as a threat. Diverse barriers to a data-sharing culture have been identified. They include data-related practices that are associated with different research disciplines and cultures, ethical issues, concerns about security and loss of

control of data leading to possible misuse, perceived effort and costs (such as loss of reputation if errors are exposed), and infrastructural barriers such as lack of suitable repositories and standards (MacMillan, 2014). Notable among these barriers are the different data cultures associated with specific academic disciplines. Data culture in this context has been defined as "the explicit and implicit data practices and expectations that determine the destiny of data. It relates to the social conventions of acquisition, curation, preservation, sharing and reuse of data" (Thessen & Patterson, 2011, p. 19). Scientific disciplines have very different data cultures relating to how data is collected, stored, and shared, and who it is shared with (MacMillan, 2014; Thessen & Patterson, 2011). Fields such as biomedicine and earth sciences (Pampel & Dallmeier-Tiessen, 2014) and molecular biology (MacMillan, 2014) have strong, established intra-disciplinary data-sharing cultures, while for others such a culture is emergent. A 2009 study of data preservation practices among 1,200 European researchers, data managers, and publishers found that only 20% of respondents deposited their research data into a digital archive (Kuipers and van der Hoeven (2009), as cited by MacMillan, 2014). While this percentage may have increased over the intervening years, the above cultural barriers seem likely to persist.

The Open Science movement has ambitions for data sharing and access that go beyond establishing a data-sharing culture among the science community. It aims to make data accessible and usable to anyone, anywhere, at any time, for any purpose (Faniel & Zimmerman, 2011). Driven by societal demand and academic policies (Pampel & Dallmeier-Tiessen, 2014), Open Science is leading to increased participation in data sharing and reuse, notably by citizens who are nonscientists (Faniel & Zimmerman, 2011). New "contexts of reuse" arise when expert or nonexpert users come from outside the scientific community in which data was generated, shared and reused and this "raise[s] questions about how individuals from different cultures and with varied knowledge and expertise find, understand, and reuse data" (Faniel & Zimmerman, 2011, p. 59). This suggests that there is fertile ground for future research to examine the dominant discourse of data culture/s involved in the open data movement and how the different data cultures across disciplines interact. Pampel and Dallmeier-Tiessen (2014) argue that successful implementation of Open Science is reliant on establishing a dominant culture of sharing, which is a far-reaching challenge that will require changes to the scientific reputation system. For example, they suggest that scientific performances should be valued with a "sharing factor" that goes beyond considering citation frequency to rate the sharing of data that is undertaken for the good of society (Pampel & Dallmeier-Tiessen, 2014, p. 221).

Open Government is another notable context in which data sharing has been elevated to a strategic level. The goals of the open government data (OGD) movement can be seen as broadly similar to those of open science (International Open Data Charter, 2015). Three key goals of data sharing in this context are transparency of government to citizens, facilitating the release of social and commercial value inherent in government data, and participatory governance, based on the idea that stakeholders who are better informed can make better decisions (Attard et al., 2015). Janssen et al. (2012) point out that this rests on two assumptions about government that require considerable transformation of the public sector: (a) that public agencies are ready “for an opening process which considers influences, discourses, and exchanges as constructive and welcomes opposing views and inputs,” and (b) that government is [ready to] to give up control to some extent (Janssen et al., 2012, p. 258). This is counterpointed against the observation that [Dutch] public sector managers “often have the tendency to avoid opening their data” (Janssen et al., 2012, p. 258). Risk averse organizational cultures and subcultures have been identified as a barrier to data sharing and the success of OGD systems (Barry & Bannister, 2014; Janssen et al., 2012). In relationship to Open Government, Barry and Bannister (2014) found that members of the public service have a high level of concern about misuse of poor information by the media due to its so-called “gotcha” culture, and that this concern is aggravated by the limited ability to correct or counter stories after they have appeared in the press. To address this deeply embedded risk-aversion and establish a culture of broader data sharing, the creation and institutionalization of a culture of open government (Janssen et al., 2012) as well as change from the top down (Barry & Bannister, 2014) is seen as being required.

The above examples show that data sharing has been elevated in societal significance, highlighting the critical need to foster data culture/s that value data sharing. It has been argued that issues relating to sharing data need to be at the forefront of concerns for those responsible for developing policy, technologists, the public, and scientific researchers (Kowalczyk & Shankar, 2011, p. 283). It is clear that over time, societal awareness of the benefits of data sharing has dramatically increased, and it can be argued that this demand is extending beyond scientific research to virtually all areas of human endeavor.

6 | CULTURES OF DATA USE/REUSE

In the previous section, we discussed the emerging cultures of data sharing from their origins in scientific

research. Internationally, data (and the instruments used to record the data) have also played a key role in diplomacy (Jacobsen et al., 2021). More recently, the concept of data diplomacy has been explored and expanded to encompass the use of data at different levels and in diverse settings, resulting in a definition of data diplomacy as “the harnessing of diplomatic actions and skills by a diverse range of stakeholders to broker and drive forward access to data, as well as widespread use and understanding of data” (Boyd et al., 2019, p. 2). The high-level vision for data sharing embodied in the open science and open government movements are reliant on data reuse. It is therefore important to consider the literature on data use/reuse as it relates to data culture/s. Questions that arise are: how do data culture/s affect re-use of data? and how do such cultures develop?

Safarov et al. (2017) argue that the key challenge to the success of OGD initiatives is not the sharing of data but its limited use (reuse) in practice. Surprisingly, their review of research into OGD utilization identifies only one study that found that organizational culture was a precondition to OGD use (i.e., Barry & Bannister, 2014). The pre-conditions for data reuse most commonly identified in the OGD utilization literature by Safarov et al.’s (2017) review were data quality, legal/policy, skills, infrastructure, availability, and privacy. The absence of data culture/s in this literature review is surprising and suggests that further investigation is needed into the role of data culture/s in open data reuse.

In contrast to this [apparent] paucity of research into data-use cultures in the OGD literature, literature on science data reuse offers insights into cultures of data use and how they develop. Cultures of data reuse in the sciences have been found to be field-specific and perpetuated through acculturation based around a system of apprenticeship: Kriesberg et al. (2013) used a community-of-practice lens to examine how cultures of data reuse and associated data competencies develop in three scientific fields: quantitative social science, archaeology, and zoology. They found that the distinctive cultures of data reuse in these three fields were fostered through mentor/mentee relationships formed between graduate students and their advisors. These findings are significant in identifying how field-specific data reuse cultures are developed as part of a cognitive apprenticeship in which novice researchers learn data-sharing cultures and norms of their fields. The authors make the case that “learning to reuse data is a form of legitimate peripheral participation, used by novice researchers to gain entry into their chosen community of practice” (Kriesberg et al., 2013, p. 4). They argue that data reuse is “a critical component of the process of acculturation for novice researchers into communities of practice because

data reuse is predicated on understanding what constitutes data within the context of a discipline, and norms for its collection and interpretation” (Kriesberg et al., 2013, p. 2).

7 | DATA ETHICS AND GOVERNANCE CULTURES

In this section, we consider the cultures of data ethics and data governance as additional aspects of data culture/s. We summarize key issues in the literature on data governance and ethics, and suggest that the relationship between data culture and data governance is underexplored. We then briefly consider recent changes in data use and governance from a power perspective, exploring how these changes can be seen as tracking shifts in the dominant voices and the positioning of ownership surrounding data use and ownership, while highlighting where loss of power has occurred and how legislative efforts have sought to redress the power imbalances that can be seen as having arisen through the use of data analytics.

Contemporary data governance (also called Data Governance 2.0) is a relatively recent phenomenon—linked with the rise of big data and its perceived value to organizations—that has shifted responsibility for data from being seen as part of the IT function to a high-level governance mechanism based around distributed accountability and collaboration across organizational silos (Famularo, 2019). In 2018, the Wall Street Journal reported a “global reckoning” on data governance had been triggered by a combination of massive data breaches, resulting reputational damage for companies, and implementation of the European Union’s General Data Protection Regulation (GDPR) compliance requirements (Famularo, 2019). Data governance can therefore be seen as being at an early stage of maturity.

Abraham et al. (2019) synthesized the definitions of data governance used by the authors of 145 papers from both academic and practitioner literature, resulting in the following:

Data governance specifies a cross-functional framework for managing data as a strategic enterprise asset. In doing so, data governance specifies decision rights and accountabilities for an organisation’s decision making about its data. Furthermore, data governance formalises data policies, standards, and procedures and monitors compliance (Abraham et al., 2019, pp. 425–426).

Alhassan et al. (2016) identified three key areas of data governance activities—defining, implementing, and

monitoring—which operate across five decision domains: data principles, data quality, metadata, data access, and the data lifecycle.

Data governance is an underdeveloped and under-researched area (Al-Ruithe et al., 2019). Perhaps due to this early stage of maturity, the concept of culture does not appear prominently in the data governance literature. By assigning decisions rights and responsibilities relating to data, data governance could be seen as a formal attempt to establish a trusted culture of data stewardship and curation. However, the relationship between organizational culture and data governance seems likely to be more complex than this and may be bi-directional. A review of critical success factors in data governance by Al-Ruithe et al. (2019) identifies organizational culture change as a critical success factor for data governance—in other words, changes in organizational culture may be necessary to align organizational culture to these data governance objectives. Others see strong data governance as a prerequisite for a so-called “data-driven culture” that is based around performing data analytics to gain strategic insight and drive innovation (e.g., Berndtsson et al., 2018). A data-driven culture is defined as a culture that is “characterized by a decision process that emphasise[s] testing and experimentation, where data outweighs opinions, and where failure is accepted—as long as something is learnt from it” (Berndtsson et al., 2018). A 2018 survey of business executives (Davenport & Bean, 2018) suggests that while start-ups may have data-driven cultures from the outset, established firms make slow progress transforming towards data-driven cultures, despite significant concerns about the potential for disruption from start-ups. The importance of fostering a culture that is suitable to support organizational data governance efforts is reinforced by Al-Ruithe et al.’s (2019) systematic review of data governance that notes that until recently [data] governance has been “mostly informal with very ambiguous and generic regulations, in siloes around specific enterprise repositories, lacking structure and the wider support of the organization” (Al-Ruithe et al., 2019, p. 839).

Literature in the area of IT governance, a more established discipline than data governance, yet one which draws on a similar cross-organizational approach, casts further light on the role of culture as a facilitator of effective governance. In a literature review, Rowlands et al. (2014) highlight several key relational factors—commitment, involvement, and trust—that can impact IT governance, and argue that mechanisms geared towards fostering these relational factors need to be actively built into IT governance frameworks to foster an effective collaboration culture in IT governance. Further, the industry body ISACA positions organizational culture, ethics,

and behavior as a key success factor in governance activities: they feature as a key component in its IT governance framework—COBIT® 2019 (Control Objectives for Information Technologies (COBIT), 2022). In contrast, there is little academic research that investigates the interaction of culture and IT governance (Rowlands et al., 2014). Based on Detert et al.'s (2000) theoretical framework for organizational culture, Rowlands et al. (2014) propose an eight-dimensional IT governance culture framework to guide future research in this area. The proposed dimensions are: (a) the basis of truth and rationality in an organization, (b) the nature of time and time horizon, (c) motivation, (d) stability versus change/innovation, (e) orientation to work, (f) task or process, (g) isolation versus collaboration/cooperation, control, coordination, and responsibility, and (h) orientation and focus—internal and/or external. It seems reasonable to suggest that these dimensions of culture (or similar ones) may also be relevant to data governance culture. We suggest that this is a valuable area for future research.

Data ethics is another concept of relevance to data culture. Data ethics is a field in flux with data ethics codes that span many domains (Stark & Hoffmann, 2019). Further, the term data ethics has no agreed definition (Hasselbalch, 2019). It has been variously conceptualized as guidelines for the ethical use of data such as professional codes of ethics (Stark & Hoffmann, 2019), structured ways of understanding what is ethical as the basis for gathering and using data, a form of work in data-driven cultures (e.g., Wehrens et al., 2021), and a social movement geared at redressing power imbalance created by the use of big data (e.g., Hasselbalch & Tranberg, 2016). In the context of data science, Stark and Hoffmann (2019) describe data ethics as involving a series of conversations that “represent an effort to better grapple with the consequences of the language we use for understanding and working with data—‘big’ or otherwise—today, and how our discourses around data cultures shape their material, cultural, and political impact” (Stark & Hoffmann, 2019, p. 3). Data ethics is thus positioned as an iterative endeavor to understand, redress and avoid the unintended negative cultural impacts of data use. In response to claims of structural discrimination and racism from movements such as Black Lives Matter, research projects including MIT's Initiative on Combatting Systemic Racism (ICSR) are seeking to harness computational tools to provide data ethicists with the capability to identify existing structural biases in data, and thus the ability to work toward racial equity and social justice as outcomes of data use (Murray, 2022).

Roche and Jamal (2021) conducted a Systematic Literature Review of research into big data ethics, “the practice of applying ethical considerations or decision-making about how large datasets are used and the impact

the use of this dataset has on individuals and society” (Roche & Jamal, 2021, p. 328). (The review drew on only 14 papers, reflecting the emergent nature of this research area). They found that the application of data ethics to big data is often bound up with issues such as data governance, cyber security, and data privacy, which may result in data ethics not being adequately prioritized and ethical risks not being fully documented (Roche & Jamal, 2021, p. 328).

Professional codes relating to data ethics can be seen as contributing to cultures of use of data. Such codes have been found to differ between fields and professions. The ethical dimension of data culture/s, therefore, appears to be profession-specific. For example, Stark and Hoffmann (2019) found that codes of ethics for those working in data science tend to focus on the prevention of environmental harm and protecting the health and safety of populations, while ethics codes for computing and statistics place the main emphasis on privacy and freedom of speech, and on data as being confidential and requiring safeguarding (Stark & Hoffmann, 2019, p. 13). In the context of the developing and using artificial intelligence (AI), where (in the case of deep learning) the learning of AI depends on large sets of training data, ethics are based around high-level principles that guide (among other things) how data is used by AI. For example, one set of principles for ethical AI includes beneficence, nonmaleficence, autonomy, justice, and explicability (Floridi et al., 2018). In the context of a data-driven healthcare system, data ethics has been positioned by Wehrens et al. (2021) as a form of discursive work performed by health practitioners as they consider what they ought to do and what is good or worthwhile, while negotiating tensions in data use. This work includes balancing of the different “goods” involved in data use (e.g., scientific, economic, public, and professional), applying ethical “fixes” for data use through institutional policies and methods (such as ethics review boards and anonymization), and collective deliberation.

An alternative higher-level perspective views data ethics as a contemporary social movement that is leading to a worldwide paradigm shift in data culture. Hasselbalch and Tranberg (2016) state that, “Across the globe, we're seeing a data ethics paradigm shift take the shape of a social movement, a cultural shift and a technological and legal development that increasingly places the human at the centre” (Hasselbalch & Tranberg, 2016, p. 10). Evidence of this shift can be seen in GDPR and other privacy legislation that improves citizens' rights to have control over data and its use by organizations and government in today's so-called Big Data Society (Hasselbalch, 2019).

From the perspective of power dynamics, the evolution of data governance and ethics outlined above, can

indeed be seen as reflecting shifts in the locus of power and perceived ownership of data over time. For example, as data has become seen as more important to organizations, society, and governments, data governance has shifted to an organizational and enterprise level, with power being moved away from the IT department and distributed to multiple voices and owners. At societal level, the growth in value of data and analytics capabilities (e.g., in response to increased data, data granularity, and insights from longitudinal observations and connections with other data sets) has created new benefits for governments and for-profit organizations that can leverage datasets to predict or alter behavior data for good (e.g., when used for by government “Nudge Units” such as the UK) or ill (e.g., when used by Cambridge Analytica to help influence the outcome of US elections). At the same time, the increasing use of these capabilities to convert data into “value” has started to erode the privacy and (arguably) ownership rights of individuals and groups. The voices of individuals and marginalized groups such as Indigenous data owners were initially lost in this shift of power around data use. The implementation of GDPR policies and parallel initiatives around the world can be seen as aiming to redress this power imbalance. However, Roche and Jamal (2021) note that there is a regulatory gap in managing big data ethics because the GDPR only applies to identifiable data (and data that can be used to re-identify individuals). A recurring theme in the literature about big data ethics is the significant power that is held by the owners of big datasets, and the implications of this power for individuals and society, notably for disadvantaged and minority groups whose experiences may not be captured in datasets (Roche & Jamal, 2021). The issue of who should hold responsibility for data ethics at a societal level is therefore complex. In the next section, we explore Indigenous perspectives on data use and governance, an issue which spans values, legal, and ethical dimensions of data culture.

8 | INDIGENOUS PERSPECTIVES ABOUT DATA CULTURE/S

Data sovereignty is a relatively new concept that has developed within the literature on data management. A recent review by Hummel et al. (2021) defines data sovereignty as being involved with, or identified with the “control of data flows via national jurisdiction” (Hummel et al., 2021, p. 1) and identifies that data sovereignty has many dimensions. This leads to the possibility that the understanding of data sovereignty is unclear, which could result in misunderstandings or disagreement about

how it is defined and represented in research outputs (Hummel et al., 2021).

Data sovereignty can also refer to the measures and regulations put in place by countries (or territories) to control data that has been generated or managed through their information-related infrastructure (Peterson et al., 2011). The development of international and regional cloud-based storage (by private companies) has made the control and ownership of this data even more complex.

The concerns of Indigenous people at a global level about the collection, management, and application of data relating to their communities has led to the development of an International Indigenous Data Sovereignty movement, as an attempt to ensure that Indigenous peoples are the primary beneficiaries of data relating to them (whether collected by others or self-generated). Over the last decade, there has been an increasing level of scholarship developing in this area. The issues associated with Indigenous data sovereignty rights can be viewed through a wider lens that relates to decolonization and the development of self-determination, and challenges past narratives that have misrepresented the realities of Indigenous lives (Smith, 2021).

Indigenous data sovereignty is described by Kukutai and Taylor (2016) as being “multifaceted [...] and involving a wide-ranging set of issues, from legal and ethical dimensions around data storage, ownership, access, and consent, to intellectual property rights and practical considerations about how data are used in the context of research, policy, and practice.” They further reinforce this by stating that “Indigenous data sovereignty thus refers to the proper locus of authority over the management of data about Indigenous peoples, their territories and ways of life.” A communique from an Indigenous data sovereignty summit defined Indigenous data sovereignty as “a global movement concerned with the right of Indigenous peoples to govern the creation, collection, ownership and application of their data” (Indigenous Data Sovereignty—Communique, 2018). Indigenous data sovereignty is seen by the Global Indigenous Data Alliance (2019) as reinforcing the “right to engage in decision-making in accordance with Indigenous values and collective interests.” This level of autonomy enables Indigenous people to set their own agenda for information, using their data to achieve their own goals and build their own narrative, not one that others set for them. The concept of Indigenous data sovereignty and the rights associated with its application are closely linked to the autonomous approaches to the decolonization of research involving Indigenous peoples. The importance of these to research outcomes is to enable “a more critical understanding of the underlying assumptions, motivations, and

values that inform research practices” involving Indigenous communities (Smith, 2021). Contextualizing these within an Indigenous data sovereignty framework creates an environment where Indigenous researchers assert Indigenous peoples’ rights to own, access, and regulate data sets made about them, arguing that Indigenous peoples have an inherently datafied way of being (Carroll et al., 2019). Although the disciplinary scope of the literature on this topic is starting to broaden, the main contributions to scholarly communication comes from those interested in data from a quantitative and demographic perspective.

The issue of Indigenous data problematics is identified by Walter et al. (2021) as being common, especially in the Anglo settler-colonial CANZUS countries (Canada, Australia, New Zealand, and the United States). Indigenous peoples in these territories have been identified with a negative and deficit-focused perspective that is not contextualized within an Indigenous world view, thus resulting in an inaccurate portrayal of Indigenous communities and the realities of their lives (Walter & Suina, 2019).

Walter and Suina (2019) build on this by linking the practice of Indigenous data sovereignty to Indigenous data governance, which asserts Indigenous interests related to data. They associate this with Indigenous decision-making across the data ecosystem, from data conception to control of access to, and usage of, data. This they state makes Indigenous decision-making a prerequisite for ensuring Indigenous data reflects Indigenous priorities, values, culture, life worlds, and diversity.

The rights of Indigenous peoples to sovereignty over their data have been closely associated with the self-determination rights stated in the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP). This is reflected in “a call for the decolonization of existing nation-state statistical systems” (Kukutai & Taylor, 2016, p. 15).

Although the term Indigenous data sovereignty has so far been the focus of this section, it acts as the umbrella term for these self-determination issues regarding data and makes it an issue that is globally important. Within each of the CANZUS nations, collectives exist that focus on the data sovereignty interests of their Indigenous populations, including Te Mana Raraunga Māori Data Sovereignty Network in New Zealand, United States Indigenous Data Sovereignty Network, Maiam nayri Wingara (Aboriginal and Torres Strait Islander Data Sovereignty Collective) in Australia and the First Nations Information Governance Centre in Canada (see: <https://indigenousdatalab.org/networks/>).

For instance, in New Zealand, Māori data sovereignty is viewed through a lens that asserts its importance in

keeping with the tino-rangatiratanga (self-determination) rights expressed as guaranteed in the Treaty of Waitangi/ Te Tiriti o Waitangi signed between the British Crown and rangatira Māori (Māori chiefs) in 1840 (New Zealand History, 1840). Māori Data Sovereignty Network (2018) believes that Māori should have complete control over the collection, analysis, storage, and use of Māori data. Similar assertions are made in the communique issued by the Maiam nayri Wingara Indigenous Data Sovereignty Collective and the Australian Indigenous Governance Institute after a summit (Māori Data Sovereignty Network, 2018), where it was stated that Indigenous peoples have the “right to exercise control of the data ecosystem including creation, development, stewardship, analysis, dissemination and infrastructure” and reserves the right for Indigenous peoples to not engage in any data processes that are inconsistent with the principles espoused in the communique. In Canada, the concept of Indigenous Data Sovereignty is based on the OCAP (Ownership, Control, Access, and Possession) principles first developed in 1998 and designed as a model to align the world views, knowledge, and cultural protocols of First Nations communities with approaches to information governance that respect these (McCreedy et al., 2018).

Although the principles discussed above have been developed, it is clear from the literature searches of the information studies disciplines that Indigenous data sovereignty is an area that has received very little attention and needs to be explored further, particularly as libraries and archival institutions’ collections include research data that has been developed through scholarly activities involving Indigenous peoples and their communities. However, it should be noted that there is a paucity of Indigenous-focused outputs to draw on in general in the literature base.

Miner (2022) focuses on the use of Indigenous data in the mapping of Missing and Murdered Indigenous Women in the United States and Canada, where data-sharing protocols restrict the use, mapping, and analysis of data without the expressed authorization of the Sovereign Bodies Institute. These protocols and processes are defined by Miner as being “tactical cartography,” and this approach challenges the data that is collected and shared by government agencies, which consistently fails to represent Indigenous realities.

The Indigenous Archives Collective (2021) has released its Position Statement on the Right of Reply to Indigenous Knowledge and Information held in Archives. This statement calls for the return (either in digital or physical repatriation) of Indigenous archival collections to Aboriginal and Torres Strait Island communities and the creation of Indigenous-led archives. It indicates that this should be informed by the involvement of

Indigenous data sovereignty experts in projects, particularly when digitization of archives is likely to cause data that was previously embedded in the paper to be extracted and subject to analysis using technology. The involvement of these experts, in their view, would be to advise and to protect this data from unwanted exposure or exploitation.

Montenegro (2019) focuses on metadata standards and the limitations of their relationship with Indigenous knowledge, finding that the universality and (supposed) neutrality of Dublin Core standards are too restrictive for the attribution needs that Traditional Knowledge requires, with the metadata often decontextualizing or causing inaccuracies to be recorded about Indigenous people's histories and realities. As a response to the generalizability of Dublin Core, Montenegro promotes Traditional Knowledge labels. These labels were created to be used by Indigenous communities to counter the Western notions of collecting, authorship, and ownership that prevent Indigenous communities from being able to exercise legal ownership or control over collections. Indigenous data sovereignty in this sense is identified as being exercised by using these labels to provide more information about the proper use, guidelines for action, and responsible stewardship of the items being labeled, and these can be adapted locally to reflect local languages, and their unique epistemological concepts and definitions. These labels, therefore, enable local Indigenous communities to provide a form of control over their histories and cultural representations.

At a panel session at the ASIST Conference in 2021, Patin et al. (2021) provided an outline as part of their discussion. This panel looked at principles of alternative ways of knowing, which included a discussion about Indigenous data sovereignty matters. Aspects of data sovereignty they emphasized were the need for self-governance of information to enable critical Indigenous nation building, and how Indigenous data sovereignty provides a framework for Indigenous communities to assert and assign their own data protocols that specifically relate to their ways of knowing. Their summary of the contribution their panel makes to this dialogue acknowledges that there needs to be a rethink of how ownership, stewardship, and access to Indigenous knowledge, and how these concepts relate to Indigenous data sovereignty principles.

The application of pre-determined principles is addressed by Adelson and Mickelson (2022), where they discuss the relationship between a Medical anthropologist (Adelson) and the Whapmagoostui First Nation on the Miiyupimatisiun Research Data Archives Project. In their article, the authors focus on the challenges and opportunities of transferring the ownership and control

of research data collected and collated by Adelson to the Whapmagoostui First Nation using the OCAP principles.

Although the literature on Indigenous data sovereignty in the information studies literature area is low, it should not be dismissed as not being relevant to our disciplines. As Adelson and Mickelson (2022) point out, the principles related to Indigenous data sovereignty require us to address fundamental questions about data ownership rights and access to this data, and this needs to be addressed one community at a time to ensure that any agreement establishes who in the community is responsible for managing the data, where the data will be stored, and how access to the archive will be provided and/or restricted. For library and information institutions in countries where reconciliation between Indigenous and non-Indigenous authorities continue to develop, demands for Indigenous control over data in its many forms will be an issue that needs to be addressed. In other words, recognizing and understanding of data culture/s is of paramount importance for any serious consideration of decolonization initiatives.

9 | CONCLUSION

Awareness of the interactions of data with culture became apparent in the course of early attempts to exchange and share data between scientists of countries on politically divergent paths (Boyd et al., 2019). Subsequently, the rise of e-science and cyberscholarship has served to motivate ongoing attention to this area, attracting considerable research funding globally (Oliver & Harvey, 2016), ultimately resulting in the development of a new sub-discipline, data curation. However, the range of concerns about the perceived issues and challenges relating to the ambiguously defined emergent topic of data culture/s is evident from the diversity of disciplines contributing to research in this area, and is by no means limited to data curation.

This article has aimed to enrich understanding of the emergent concept of data culture/s by highlighting and exploring key concepts from the literature that relate to data culture/s. (We purposively did not explore IT-related cultural phenomena that may co-exist with data culture/s such as digital culture/s.) It is important to note that in employing a retroductive approach to this analysis, we did not set out to undertake a systematic literature review, nor did we set out to read every article that employs the term data culture/s. Rather, our goal was to surface key dimensions, inputs, and aspects of data culture/s, attributes, and dimensions of data culture/s and the nature of their influence or impact. A further limitation is that the literature reviewed was limited to material

published in English. The analysis is based on our interpretation and perspectives as information studies/systems researchers based in Australia and in Aotearoa New Zealand.

Our focus in this literature review has also been strongly influenced by the research priorities identified by information studies and information systems scholars, considering the cultural dimensions from micro (individuals' skills and competencies), meso (organizational and inter-organizational initiatives), technological (the data itself and associated infrastructure), and macro (Indigenous) perspectives. Our analysis of the literature shows that there is a considerable body of knowledge emerging about different dimensions, inputs, and aspects of data culture, but that within our specialist disciplines this work is considerably more diffuse. The analysis highlights the richness, diversity, and complexity of data culture/s, as well as the very early stage of research in this area. We therefore consider it timely to propose the adoption of the term data culture/s in information studies and information systems research, in order to facilitate the potential to identify synergies and leverage collaborative effort, to further map this construct to the literature.

Much further research is needed to progress understanding of the critical nature of data culture/s and their impact on influences, activities, and initiatives underway at every level of societal endeavor. This includes opportunities for projects that are based on the collection and analysis of empirical data to interrogate the concept and validate the relationships highlighted in this article. As a first step, the development of a conceptual model making explicit the interrelationships of the initial thematic candidate dimensions explored in this article could provide the framework for a research agenda transcending disciplinary silos apparent between cognate as well as noncognate academic disciplines.

ACKNOWLEDGMENT

Open access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST

We have no known conflict of interest to disclose.

ORCID

Gillian Oliver  <https://orcid.org/0000-0003-2943-5035>

Jocelyn Cranefield  <https://orcid.org/0000-0003-3690-1698>

Spencer Lilley  <https://orcid.org/0000-0002-4881-4981>

Matthew J. Lewellen  <https://orcid.org/0000-0001-6970-7715>

REFERENCES

- Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424–438.
- Adelson, N., & Mickelson, S. (2022). The Miiyupimatisiun research data archives project: Putting OCAP® principles into practice. *Digital Library Perspectives*, 508–520. <https://doi.org/10.1108/DLP-11-2021-0099>
- Alhassan, I., Sammon, D., & Daly, M. (2016). Data governance activities: An analysis of the literature. *Journal of Decision Systems*, 25(suppl 1), 64–75.
- Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2019). A systematic literature review of data governance and cloud data governance. *Personal and Ubiquitous Computing*, 23(5), 839–859.
- Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399–418.
- Barry, E., & Bannister, F. (2014). Barriers to open data release: A view from the top. *Information Polity*, 19(1–2), 129–152.
- Berndtsson, M., Forsberg, D., Stein, D., & Svahn, T. (2018). Becoming a data-driven organisation. In *Proceedings of the 26th European Conference on Information Systems (ECIS2018)*, Portsmouth, United Kingdom, June 23–28, 2018.
- Bonikowska, A., Sanmartin, C. A., & Frenette, M. (2019). Data literacy: What it is and how to measure it in the public service (No. 022; Analytical Studies: Methods and References). Analytical Studies Branch, Statistics Canada. Retrieved from https://epe.lac-bac.gc.ca/100/201/301/weekly_acquisitions_list-ef/2019/19-33/publications.gc.ca/collections/collection_2019/statcan/11-633-x/11-633-x2019003-eng.pdf
- Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6), 1059–1078. <https://doi.org/10.1002/asi.22634>
- Boyd, A., Gatewood, J., Thorson, S., & Dye, T. D. V. (2019). Data Diplomacy. *Science & Diplomacy*, 8(1) Retrieved from <http://sciencediplomacy.org/article/2019/data-diplomacy>
- Capdarest-Arest, N., & Navarro, C. E. (2021). Promoting health data fluency skills by expanding data and informatics work in libraries: The role of a health library informaticist. *Medical Reference Services Quarterly*, 40(1), 130–138. <https://doi.org/10.1080/02763869.2021.1873648>
- Carroll, S. R., Rodriguez-Lonebear, D., & Martinez, A. (2019). Indigenous data governance: Strategies from United States native nations. *Data Science Journal*, 18, 31. <https://doi.org/10.5334/dsj-2019-031>
- Condon, P. B., & Pothier, W. G. (2022). Advancing data literacy: Mapping business data literacy competencies to the ACRL framework. *Journal of Business & Finance Librarianship*, 27(2), 104–126. <https://doi.org/10.1080/08963568.2022.2048168>
- Control Objectives for Information Technologies (COBIT). (2022). ISACA. Retrieved from <https://www.isaca.org/resources/cobit>
- Davenport, T. H., & Bean, R. (2018). Big companies are embracing analytics, but most still don't have a data-driven culture. *Harvard Business Review*, 6, 1–4.
- Detert, J. R., Schroeder, R. G., & Mauriel, J. J. (2000). A framework for linking culture and improvement initiatives in organizations. *Academy of Management Review*, 25(4), 850–863.

- Famularo, A. (2019, March 11). *Council Post: The evolution of data governance*. Forbes Retrieved from <https://www.forbes.com/sites/forbestechcouncil/2019/03/11/the-evolution-of-data-governance/>
- Faniel, I. M., & Zimmerman, A. (2011). Beyond the data deluge: A research agenda for large-scale data sharing and reuse. *International Journal of Digital Curation*, 6(1), 58–69.
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., & Rossi, F. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707.
- Fortun, K. (2016). Scaling and visualizing multi-sited ethnography. In *Multi-sited Ethnography* (pp. 87–100). Routledge.
- Ginman, M. (1993). Information and business performance. In J. L. Olaisen (Ed.), *Information management: A Scandinavian approach* (pp. 79–94). Scandinavian U.P.
- Global Indigenous Data Alliance. (2019). Global indigenous data Alliance. Retrieved from <https://www.gida-global.org>
- Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, 5(2), 1–13. <https://doi.org/10.1177/2053951718786316>
- Hasselbalch, G. (2019). Making sense of data ethics. The powers behind the data ethics debate in European policymaking. *Internet Policy Review*, 8(2), 1–19.
- Hasselbalch, G., & Tranberg, P. (2016). *Data ethics: The New Competitive Advantage*. Publishare.
- Howarth, D., & Stavrakakis, Y. (2000). Introducing discourse theory and political analysis. In D. Howarth, A. J. Norval, & Y. Stavrakakis (Eds.), *Discourse theory and political analysis: Identities, hegemonies and social change*. Manchester University Press.
- Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. *Big Data & Society*, 8(1), 2053951720982012.
- Indigenous Archives Collective. (2021). The indigenous archives collective position statement on the right of reply to indigenous knowledges and information held in archives. *Archives and Manuscripts*, 49(3), 244–252. <https://doi.org/10.1080/01576895.2021.1997609>
- Indigenous Data Sovereignty—Communique. (2018). Indigenous Data Sovereignty Summit (Maiam nayri Wingara Indigenous Data Sovereignty Collective and the Australian Indigenous Governance Institute). Retrieved from <https://aigi.com.au/wp-content/uploads/2019/10/Communique-Indigenous-Data-Sovereignty-Summit.pdf>
- International Open Data Charter. (2015). *Open Data Charter—Principles*. Retrieved from <https://opendatacharter.net/principles/>
- Jacobsen, L. L., Fedorova, I., & Lajus, J. (2021). The seismograph as a diplomatic object: The soviet–American exchange of instruments, 1958–1964. *Centaurus*, 63(2), 277–295. <https://doi.org/10.1111/1600-0498.12393>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29(4), 258–268.
- Kennedy-Clark, S., & Reimann, P. (2022). Knowledge types in initial teacher education: A multi-dimensional approach to developing data literacy and data fluency. *Learning: Research and Practice*, 8(1), 42–58. <https://doi.org/10.1080/23735082.2021.1957140>
- Kirkwood, R. J. (2016). Collection development or data-driven content curation? An exploratory project in Manchester. *Library Management*, 37(4/5), 275–284. <https://doi.org/10.1108/LM-05-2016-0044>
- Koltay, T. (2015). Data literacy: In search of a name and identity. *Journal of Documentation*, 71(2), 401–415. <https://doi.org/10.1108/JD-02-2014-0026>
- Koltay, T. (2017). Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science*, 49(1), 3–14. <https://doi.org/10.1177/0961000615616450>
- Kowalczyk, S., & Shankar, K. (2011). Data sharing in the sciences. *Annual Review of Information Science and Technology*, 45(1), 247–294. <https://doi.org/10.1002/aris.2011.1440450113>
- Kriesberg, A., Frank, R. D., Faniel, I. M., & Yakel, E. (2013). The role of data reuse in the apprenticeship process. *Proceedings of the American Society for Information Science and Technology*, 50(1), 1–10.
- Kukutai, T., & Taylor, J. (2016). *Indigenous data sovereignty: Toward an agenda*. ANU Press.
- MacMillan, D. (2014). Data sharing and discovery: What librarians need to know. *The Journal of Academic Librarianship*, 40(5), 541–549.
- Māori Data Sovereignty Network. (2018). Principles of Māori data sovereignty (Brief #1). Te Mana Rauranga. Retrieved from <https://static1.squarespace.com/static/58e9b10f9de4bb8d1fb5ebbc/t/5bda208b4ae237cd89ee16e9/1541021836126/TMR+Ma%CC%84ori+Data+Sovereignty+Principles+Oct+2018.pdf>
- Mecredy, G., Sutherland, R., & Jones, C. (2018). First nations data governance, privacy, and the importance of the OCAP® principles. *International Journal of Population Data Science*, 3(4). <https://doi.org/10.23889/ijpds.v3i4.911>
- Miner, J. D. (2022). Informatic tactics: Indigenous activism and digital cartographies of gender-based violence. *Information, Communication & Society*, 25(3), 431–448. <https://doi.org/10.1080/1369118X.2020.1797851>
- Montenegro, M. (2019). Subverting the universality of metadata standards the TK labels as a tool to promote indigenous data sovereignty. *Journal of Documentation*, 75(4), 731–749. <https://doi.org/10.1108/JD-08-2018-0124>
- Muganda, N. (2013). Linking reasoning to theoretical argument in information systems research. In *Proceedings of the 19th Americas Conference on Information Systems, AMCIS 2013—Hyperconnected World: Anything, Anywhere, Anytime*, Vol. 3, pp. 2311–2321.
- Murray, S. (2022, March 16). *How artificial intelligence can help combat systemic racism*. MIT News | Massachusetts Institute of Technology Retrieved from <https://news.mit.edu/2022/how-ai-can-help-combat-systemic-racism-0316>
- New Zealand History. (1840). Treaty of Waitangi [Māori version].
- Oliver, G., Cranefield, J., Lilley, S., & Lewellen, M. (2023). Data cultures: A scoping literature review. *Information Research*, 28(3), 950.
- Oliver, G., & Harvey, R. (2016). *Digital curation* (2nd ed.). American Library Association Retrieved from <https://lccn.loc.gov/2015043274>
- Organisation for Economic Co-operation and Development (OECD). (2007). *OECD Principles and Guidelines for Access to Research Data from Public Funding*. Organisation for Economic Co-Operation and Development Retrieved from <http://www.oecd.org/dataoecd/9/61/38500813.pdf>

- Organisation for Economic Co-operation and Development (OECD). (2022). Future of education and skills 2030. Retrieved from <https://www.oecd.org/education/2030-project/>
- Pampel, H., & Dallmeier-Tiessen, S. (2014). Open research data: From vision to practice. In *Opening science* (pp. 213–224). Springer.
- Pangrazio, L., & Sefton-Green, J. (2020). The social utility of ‘data literacy’. *Learning, Media and Technology*, 45(2), 208–220. <https://doi.org/10.1080/17439884.2020.1707223>
- Patin, B., Oliphant, T., Allard, D., Gray, L., Clarke, R. I., Tacheva, J., & Lar-Son, K. (2021). At the margins of epistemology: Amplifying alternative ways of knowing in library and information science. *Proceedings of the Association for Information Science and Technology*, 58(1), 630–633. <https://doi.org/10.1002/pr2.515>
- Peterson, Z. N., Gondree, M., & Beverly, R. (2011). A position paper on data sovereignty: The importance of geolocating data in the cloud. In *Proceedings of the 3rd USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 11)*.
- Poirier, L., & Costelloe-Kuehn, B. (2019). Data sharing at scale: A heuristic for affirming data cultures. *Data Science Journal*, 18(1). <https://doi.org/10.5334/dsj-2019-048>
- Pothier, W. G., & Condon, P. B. (2020). Towards data literacy competencies: Business students, workforce needs, and the role of the librarian. *Journal of Business & Finance Librarianship*, 25(3–4), 123–146. <https://doi.org/10.1080/08963568.2019.1680189>
- Roche, J., & Jamal, A. (2021). A systematic literature review of the role of ethics in big data. In H. Jahankhani, A. Jamal, & S. Lawson (Eds.), *Cybersecurity, Privacy and Freedom Protection in the Connected World. Proceedings of the 13th International Conference on Global Security, Safety and Sustainability, London, January 2021* (pp. 327–342). Springer.
- Rowlands, B., Haes, S. D., & Grembergen, W. V. (2014). Exploring and developing an IT governance culture framework. *ICIS 2014 Proceedings*, 7 Retrieved from <https://aisel.aisnet.org/icis2014/proceedings/ISGovernance/7>
- Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, 22(1), 1–24.
- Smith, L. T. (2021). *Decolonizing methodologies: Research and Indigenous peoples* (3rd ed.). Zed Books.
- Stark, L., & Hoffmann, A. L. (2019). Data is the new what? Popular metaphors & professional ethics in emerging data culture. *Journal of Cultural Analytics*, 4, 1–22. <https://doi.org/10.22148/16.036>
- Sternkopf, H., & Mueller, R. M. (2018, January 3). Doing good with data: Development of a maturity model for data literacy in non-governmental organizations. In *Proceedings of the 51st Hawaii International Conference on System Sciences*, Hawaii, USA. Retrieved from <http://hdl.handle.net/10125/50519>
- The Research Data Alliance (RDA). (2022a, May). RDA in a Nutshell. Retrieved from <https://www.rd-alliance.org/about-rda>
- The Research Data Alliance (RDA). (2022b, May). The RDA Vision and Mission. Retrieved from <https://www.rd-alliance.org/about-rda>
- The United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG). (2014). A World that Counts: Mobilising the Data Revolution for Sustainable Development. Retrieved from <https://www.undatarevolution.org/report/>
- Thessen, A. E., & Patterson, D. J. (2011). Data issues in the life sciences. *ZooKeys*, 150, 15–51.
- Walter, M., Lovett, R., Maher, B., Williamson, B., Prehn, J., Bodkin-Andrews, G., & Lee, V. (2021). Indigenous data sovereignty in the era of big data and open data. *Australian Journal of Social Issues*, 56(2), 143–156. <https://doi.org/10.1002/ajs4.141>
- Walter, M., & Suina, M. (2019). Indigenous data, indigenous methodologies and indigenous data sovereignty. *International Journal of Social Research Methodology*, 22(3), 233–243. <https://doi.org/10.1080/13645579.2018.1531228>
- Wehrens, R., Stevens, M., Kostenzer, J., Weggelaar, A. M., & de Bont, A. (2021). Ethics as discursive work: The role of ethical framing in the promissory future of data-driven healthcare technologies. *Science, Technology & Human Values*, 1–29. <https://doi.org/10.1177/01622439211053661>
- Wilson, T. D. (2002). The nonsense of knowledge management. *Information Research*, 8(1), 1–33.
- Yan, J., Steinke, T., & Brazeau, M. (2002). Creativity in math? Say it isn't so! Building data literacy in high schools. *School Libraries in Canada*, 22(1), 28–30.

How to cite this article: Oliver, G., Cranefield, J., Lillie, S., & Lewellen, M. J. (2023). Understanding data culture/s: Influences, activities, and initiatives. *Journal of the Association for Information Science and Technology*, 1–14. <https://doi.org/10.1002/asi.24737>