INDIVIDUAL PERSPECTIVES ON DATA SHARING: HUMAN FACTORS IMPACTING THE DIGITAL ECONOMY

CASSANDRA LAROSE, CONOR FALVEY, MICHAEL SMIT
SHAYLA FITZSIMMONS, CHRISTINA MACDONALD

DALHOUSIE UNIVERSITY
CIOOS ATLANTIC

AUGUST 15, 2021

CONTENTS

EXECUTIVE SUMMARY

BACKGROUND

The recently accelerated digital transformation is straining organizations and individuals who are finding their digital skills stretched in new and interesting ways. Increasingly digital operations require increasing expertise in managing, analyzing, and sharing data. The established consensus is that certain types of data should be open by default, particularly when that data is held by publicly funded governments and organizations and is of public interest (for example, open government data and open research data). While attention is paid to structural and organizational factors that limit widespread sharing of open data, there is no synthesized understanding of how human factors and human psychology introduce or exacerbate barriers to data sharing, or how these barriers can be mitigated. Understanding and enabling open data is important for both social and economic reasons, but we also consider open data to be a data sharing canary. It warns us of potential issues in a digital economy that we expect will continue to increase its reliance on the exchange of data in various ways.

OBJECTIVES

We synthesized research conducted over the past decade to identify the barriers to data sharing introduced or influenced by individual human factors, with a particular focus on open scientific or research data and government open data. We also examined published literature for recommendations on how to mitigate these barriers, and empirical evidence of the impact of these recommendations. This narrative synthesis aims to provide the critical reflection necessary to re-focus and foster discussion of recurrent issues hampering data sharing, which if unidentified and unaddressed, will continue to prevent workers, employers and policy-makers from full and sustainable realization of the potential of the Digital Economy.

METHODS

Search keywords were identified by reviewing articles on data sharing practices and by conducting initial searches to identify keywords relating to human factors barriers towards data sharing. The final Scopus search was reviewed by all members of the research team and a random set of 100 articles were assessed for inclusion or exclusion. We met to discuss our decisions to refine inclusion/exclusion criteria until we achieved agreement. Two authors then reviewed all articles to include or exclude based on the title and abstract. Finally, full articles were reviewed for inclusion, and we completed our synthesis with 58 articles. Data extraction focused on original evidence around perceptions and attitudes toward data sharing that were, or were closely linked with, human factors; the type, subject matter, and location of data; and potential approaches to mitigating any barriers to data sharing.

RESULTS

We identified 58 articles, one third of them published in 2020, that included individual perspectives on data sharing. We identified six core themes in these perspectives. 1. Individuals describe implicit and often explicit cost-benefit analyses, where the costs are typically direct and immediate and personal,

and the benefits are typically indirect, time-delayed, and dispersed throughout a broader community. 2. A lack of organizational data sharing culture was identified, in the context of individuals feeling isolated or not wanting to act alone. 3. Individuals describe real and perceived skill and knowledge gaps, most not on working with data but rather on engaging in the open data / research data management ecosystem: terminologies, practices, repositories, metadata regimes, etc. 4. Individuals identified a range of concerns around the theme of losing control of the data, from not being appropriately credited for their work to others misusing or exploiting their data. 5. Related, data custodians take seriously their responsibility to protect individual privacy and data ownership, were not confident in their ability to anonymize data, and feared the potential consequences of accidentally sharing private data, preferring to instead not share anything. 6. Individuals were concerned that data quality issues or mistakes in analysis would be discovered thanks to the data being available, and in particular were concerned about the consequences of such mistakes.

KEY MESSAGES

The picture that emerged overall is of a field that has not yet embraced individual psychology in understanding data sharing behavior. Individual perspectives will always be intertwined with factors like organizational structures, incentive systems, resource availability, and perceptions of institutional culture (system-level factors). Understanding and mitigating human factors requires understanding the institutional and organizational barriers, which are built explicitly and implicitly by humans. Yet there has been almost no work done to disentangle human and system factors. The role of our human brains and emotions is not commonly discussed in the literature, and never with the support of experts in psychology.

When individual perspectives are clear, there are often misalignments between system-level incentives and individual perspectives. For example, while individuals described concern regarding errors or quality issues being discovered, this type of discovery is described as a benefit of sharing open data at a social and institutional level. Although one can imagine this concern is tied to human factors such as imposter syndrome (Clance & Imes, 1978) or performance anxiety (Holland et al., 2016), no explicit link has been established.

Various strategies for addressing individual and system-level factors were identified, with the most common being advice to build a "culture" of data sharing and little detail. We found no evidence of the effectiveness of such mitigations, but include a detailed list of some recurring suggestions from the expert authors of the papers included in this synthesis.

This synthesis identified substantial gaps in research. Increased collaboration with experts in human psychology (motivation, fear), change management, incentive systems, persuasive computing, and other related fields is required to understand the human element of data-sharing cultures. Studies examining the effectiveness of various interventions to mitigate barriers to data sharing are necessary.

BACKGROUND

The recently accelerated digital transformation is straining organizations and individuals who are finding their digital skills stretched in new and interesting ways. Across sectors and fields, we have seen a transition from data scarcity to data abundance, which requires a fundamental shift in how we view data: from a resource to be protected to a resource we can only effectively utilize through collaboration. Yet the psychology of scarcity is a powerful, enduring force on organizations established when data was still measured using kilobytes. The growing consensus is some data should be open by default, particularly when that data is held by publicly-funded governments and organizations and is of public interest. While there remains data which can (or must) be closely held and protected, and institutional and system barriers present challenges, it's important to understand human factors that present barriers to widespread data sharing. This is particularly important in areas where management practices expect data sharing: open government data and open research data.

While attention is paid to structural and organizational factors that limit widespread sharing of open government data and open research data ("open data"), there is no synthesized understanding of how human factors present or exacerbate barriers to data sharing, or how these barriers can be mitigated. Understanding and enabling open data is important for both social and economic reasons, but we also consider open data to be a data sharing canary. It warns us of potential issues in a digital economy that we expect will continue to increase its reliance on the exchange of data in various ways. In this report, we synthesize research conducted over the past decade to identify the barriers to open data sharing caused or exacerbated by human factors, and mitigation strategies for each identified barrier. We describe promising areas for further study.

DIGITAL REVOLUTION AND DATA LITERACY

The digital revolution, which started with the Internet and continues with mobile devices, social networking, big data, and computing clouds, has left few aspects of global societies and economies untouched. Digitization of information has proceeded rapidly: in 2012, analysts estimated 90% of the world's data had come into existence within the previous 2 years (Vesset, 2014). These global networks of commerce, communication, and transport coupled with advanced information communication technologies (ICTs) that make up the Digital Economy have allowed for a dramatic increase in collection, analysis, sharing, and use of data (Pentland, 2013).

In the face of this data abundance, organizations in all sectors are struggling with managing the volume of data being produced, the velocity at which it is growing and the variety of its formats. Just as regulations were required to address the water and air pollution caused by the Industrial Revolution, there is a growing realization that there are side-effects from collecting so much data ("information pollution") and that measures need to be put in place to properly harness the power of these data flows.

The goal is to transition from being data-rich to being information-rich and knowledge-rich, for which both data scientists and people capable of working effectively with data are needed. Indeed, while digital technologies often reshape the way work is done, it is the way in which individuals and organizations execute this transition that makes the difference in whether they achieve what they've set

out to. This set of skills, referred to as *data literacy*, is the ability to collect, manage, evaluate and apply data in a critical manner (Ridsdale et al., 2015), and was examined in depth in a previous Knowledge Synthesis project.

The COVID-19 pandemic and the massive shift online necessitated by physical distancing requirements has accelerated the transition by workers and industries to the digital economy, and has spurred innovation while bringing into sharp focus significant inadequacies, including the lack of data literacy skills in the workforce (Deloitte, 2020). In 2015, Canada faced a shortage of 150,000 people with the data analytics skills required by the market (Steele, 2015). By 2031, that gap is forecast to widen to two million people (Miner, 2014).

DATA SHARING

'Data sharing', including knowledge of methods and platforms for sharing data and how to share data legally and ethically, is identified as a key competency of data literacy (Ridsdale et al., 2015). This is not surprising as the debate on access to and ownership of data has become the most visible regulatory discussions in both recent legal scholarship and current regulatory policies (Richter and Slowinski, 2018). We are particularly interested in data sharing that involves making data in the public interest openly and freely accessible. (Other versions of data sharing, including contractual quid pro quo sharing between two parties, are important but not in scope.)

Public data sharing can be guided, among other things, by open access principles as well as FAIR principles. 'Open Data' is defined as data that anyone can access, use and share. It has an open license and is human readable and machine readable. While not equal to Open, the FAIR data principles (data that are findable, accessible, interoperable, and reusable) (Wilkinson et al., 2016) play an essential role in the objectives of open science to improve and accelerate scientific research, to increase the engagement of society, and to contribute significantly to economic growth.

Public bodies are among the largest creators and collectors of data in many different domains (Janssen, 2011) and the availability of open data has grown significantly as pressure has been placed on all kinds of public organizations to release their raw data. For example, the Government of Canada released an Open Government directive in 2014 with the objective "to maximize the release of government information and data of business value to support transparency, accountability, citizen engagement, and socio-economic benefits through reuse, subject to applicable restrictions associated with privacy, confidentiality, and security" (Government of Canada, 2014).

Strong data management and data sharing practices are essential to open science, and are increasingly being identified as essential ingredients for conducting good research in general (Vertesi and Dourish, 2011; Gallagher et al., 2015). The principal funders of research in Canada, the Tri-Agencies, have an Open Access Policy on Publications (Government of Canada, 2019) and have developed a draft Tri-Agency Research Data Management Policy, which aims to support Canadian research excellence by fostering sound digital data management and data stewardship practices (Government of Canada, 2018).

An important consideration is data held by Indigenous peoples, including both traditional ecological knowledge and data collected using Western scientific methods. Considering human factors in this context, including OCAP and the varied interests of different organizations and Nations, without excluding Indigenous peoples from management decisions and developments, is a complex and nuanced issue which requires leadership and involvement of Indigenous peoples. It is out of scope for this project, but is essential partnered work (Proulx et al., 2021).

DATA SHARING BENEFITS AND BARRIERS

There is widespread agreement on the benefits of data sharing (Tenopir et al., 2011). Openly shared data aims for organizations to become more transparent and thereby accountable to citizens, to increase organizational efficiency and effectiveness by better decisions or to realize economic activity (Eckartz et al., 2014). There are also incentives to share scientific data publicly: requirements from journals and/or funding agencies; participation in networks or clusters; advancing the state of science and enabling replicability; data publications and associated citations; and others. The annual opportunity cost of not having FAIR research data is estimated to be at least €10.2 bn for the European scientific system (European Commission, 2019).

Despite these stated benefits, issues surrounding data sharing persist (Hipsley and Sherratt, 2019). In a study among environmental scientists, Tenopir et al. (2011) found that less than 6% of the surveyed researchers make all of their data available. Andreoli-Versbach and Mueller-Langer (2014) found that 81% of empirical economists do not voluntarily share their data. In a study among researchers from different disciplines, Fecher et al. (2015) found that only 13% had actually made their own data publicly available in the past.

Significant research has been conducted to determine the underlying barriers to research data sharing (Kucera, 2017; Martin et al., 2013; Janssen, 2011). Certainly there are types of data, such as human data or species data, that have associated bioethical concerns such as privacy, anonymity and non-maleficence that prevent open sharing (Bezuidenhout, 2019). For data that can be openly shared, however, Hipsley and Sherratt (2019) argue that common barriers such as data management, curation, enforcement, and associated costs are largely technical, meaning their solutions lie in improved technology and support to guide producers through the data sharing experience. For example, practical aids for data sharing are already in place, including data management training resources (e.g., Portage), repository registries (e.g., www.re3data.org), and public (e.g., figshare, Zenodo, CIOOS) or national (e.g., Scholars Portal Dataverse, Federated Research Data Repository) data sharing platforms. These technical factors are important, and are being addressed in various ways by many groups, including the authors and other collaborators. Addressing them is necessary, but not sufficient.

OBJECTIVES

Achieving the level of data sharing required to advance the digital economy in Canada requires a thoughtful and critical understanding of these human factors. Janssen (2011) and Bezuidenhout (2019) affirm that insights into user's perspectives are necessary to support the adoption of open data systems and changes in data sharing practices, and that the solutions may not be so straightforward as creating open data incentives and policies (Fecher et al., 2015). Understanding what we currently know about

human factors and clearly identifying gaps in our understanding is an essential first step. We know there is empirical data on data sharing success stories (e.g. Martone et al., 2018), but are there documented instances where worst fears have been realized (e.g., researchers were scooped or where data was misused or misinterpreted)? What steps have been taken to reduce these barriers to data sharing, and is there empirical evidence of the impact? Broad adoption of data sharing practices, and achieving the benefits of open data policies, requires this understanding.

This report synthesizes research conducted over the past decade to identify the barriers to data sharing caused or exacerbated by human factors, with a particular focus on scientific or research data and government open data. Included is research that assesses the validity of these concerns, the recommendations that have been developed to mitigate such concerns, and empirical evidence of the impact of these recommendations. We provide the critical reflection necessary to re-focus and foster discussion of recurrent issues hampering data sharing, which if unidentified and unaddressed, will continue to prevent workers, employers and policy-makers from full and sustainable realization of the potential of the Digital Economy.

We used a systematic narrative synthesis methodology based on the systematic approaches described by Popay et al. (2006), Petticrew & Roberts (2005), and Grimshaw (2010). A narrative synthesis is one of the most common types of knowledge synthesis (Grimshaw, 2010) and is appropriate for capturing barriers to data sharing in their original context. It is particularly suited to synthesizing heterogeneous studies and case studies, which are common in the formal literature.

Our specific review questions were:

1. How do human factors introduce or influence barriers to sharing open data?
2. What interventions have been attempted to mitigate barriers introduced by human factors? What evidence is there on the effectiveness of these interventions?

METHODS

We conducted a narrative synthesis review of relevant scholarly literature to inform our narrative synthesis. While an initial protocol was defined a priori, we iteratively modified this protocol based on the unexpectedly high number of results. Where relevant and appropriate, this report follows the PRISMA Scoping Review checklist (Trico et al., 2018).

To be included in the synthesis, we required papers to be peer-reviewed; to report original results from primary data collection from a population that included data owners, managers, custodians, or others involved in the decision to share data; to describe human factors; to be in English; and to include research data or government open data. We did not require a particular data collection method or study approach, we did not constrain the size of the study, we welcomed both quantitative and qualitative research, and we included case studies. Consistent with the rise of expectations around open data, we included results published in the past decade, but did not consider any other recency criteria.

We first conducted an initial limited search of Scopus, Web of Science, Google Scholar, and papers we had cited ad hoc in previous papers to identify appropriate keywords and relevant fields of study. This led to the iterative development of a search strategy by two Information Management students (CL and CF), which was reviewed by an Information Management professor (MS) and data management experts in research and practice (MS, SF, CM).

Scopus, Web of Science, and Google Scholar were considered as potential databases to search for the research synthesis. Google Scholar was noted as returning large numbers of results on initial exploratory searches, but the inability to utilize filters and advanced search strategies made its use resource prohibitive, as suggested by Gusenbauer & Haddaway (2020). Web of Science and Scopus both allow for more advanced search strategies and filtering. Scopus was ultimately chosen for the current project. Scopus has been assessed as a suitable tool for conducting a research synthesis (Gusenbauer & Haddaway, 2020), and is among the most comprehensive of the journal indexing databases while still including only peer-reviewed articles (Martin-Martin et al, 2021) and supporting the sophisticated query and filtering tools required (Gusenbauer & Haddaway, 2020). None of the tools completely capture all literature on a subject, but Scopus is more complete in the general disciplines relative to our search (Gusenbauer & Haddaway, 2020). Scopus is still limited by the decisions of the Scopus Content Selection and Advisory Board.

Search terms pertaining to human factors were brainstormed. These were then tested by searching for each candidate term alongside the phrase "open data" in order to estimate the number and relevance of the results returned. Because the search terms being tested were vague (e.g., "barriers", "incentives", "practices", "data", etc.), and because open data is often referenced in descriptions of methodology within abstracts, the results were generally quite noisy. We had to limit our search to "research data sharing", not just "research data".

The most successful balance of sensitivity and relevance came from using proximity operators. After some experimentation, we concluded that searching for terms within 15 words of each other is a rough proxy for terms appearing in the same sentence, while a distance of 5 words or less approximates terms appearing in the same phrase or list. Our final query is shown in Table 1. It requests papers that mention open data and research data, human factors, and barriers/motivators (plus alternate phrasings of each concept) within relatively close proximity within the title, abstract, or keywords. Results were limited to those published within the last decade (2011-2021 inclusive, with 2021 an incomplete year). The search used in subsequent analysis was completed on April 20, 2021, with the results exported to a spreadsheet. No forward/backward citation searching was used. The results of this search are described here and reported in a PRISMA flow diagram (Figure 1).

The search returned 1,340 articles after removal of duplicates. A random selection of 50 articles was screened by five researchers to determine whether the article full text should be reviewed, or if it could be excluded on the basis of the title and abstract only. Each paper where we did not agree was discussed until agreement was reached and inclusion criteria could be clearly specified. Two researchers (CL and CF) then independently reviewed 100 candidate titles and abstracts each, with their conclusions reviewed by a third researcher (MS)) who found 98.5% agreement (all papers included that on further reflection should not have been). The remainder of the articles were reviewed by a single researcher for inclusion based on title and abstract.

Following this exclusion process, 140 articles remained for review. These were then included or excluded based upon their content after reading in full. At the end of this review process, 58 articles were included in the synthesis, appraised, and used to develop the themes. 82 were excluded for various reasons including methodology (e.g., literature reviews, opinion pieces), language (e.g., articles not written in English), and lack of human factors identification. We examined literature reviews for relevance, and discuss these further in the Results section.

Data extraction was completed by two researchers with no overlap, and spot-reviewed by a third researcher. Data extracted included brief summary of relevance to our review, type of data, methods used, population, sample size (if relevant), geographic area, human factor barriers identified, mitigations identified, and key properties of the researcher. Where possible, verbatim text describing barriers and mitigating factors was extracted. This information was combined with bibliographic information provided by Scopus.

In the human factor barriers data extraction step, factors that may prevent actors from sharing data were identified and grouped in an initial review to allow for an informal coding process to identify factors or themes that were common to the articles. Two researchers reviewed this extracted data (8,385 words) and identified the factors and concepts that appeared repeatedly. After identifying a number of common themes, with discussion of the team, categories for these barriers were identified and described in the following sections. Mitigations for these factors were also extracted (2,819 words) and synthesized into categories.

RESULTS

In our 58 selected articles, 32 included results relating to open research data, and 29 on open government data (3 included both). One third (n=19) were published in 2020 (Figure 2). Our search retrieved an increasing number of articles each year (Figure 2), but a disproportionate number of 2020 articles met the inclusion criteria (10%, compared to less than 5% of articles in all other). This suggests an increase in interest in the subject of human factors as a barrier to data sharing within the narrow parameters of our search.

Most of the articles focus on a single country (41), ten articles focus on 2-6 countries, five are global (only one translated to additional languages), and two focus on LMIC. A total of 40 countries are the focus of at least one article, with UK and Germany receiving the most attention. Notably absent is China. Given the context of this study, it is also noteworthy that none focused on Canada. A table of all the articles returned, and which of our synthesized themes the paper speaks to, is in Appendix A.

The human factors that pose barriers to data sharing are entangled with the institutional barriers. In some cases, they are explicitly linked. For example, individual perspectives on the value of sharing data directly influence the extent to which they perceive institutional barriers (Wang et al., 2019). The themes identified are described through a human factors lens, but are related to the themes in reviews focused on institutional barriers (e.g. Zuiderwijk et al., 2020; Haini et al., 2020; Perrier et al., 2020).
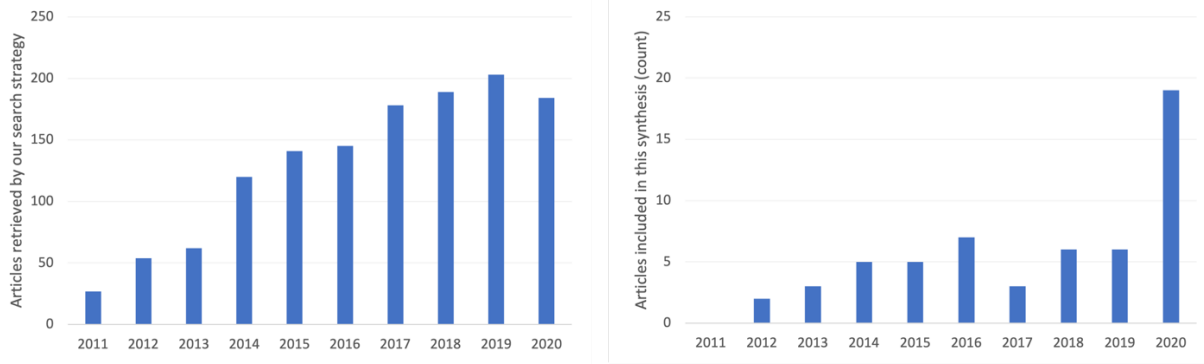
Figure 1. The search shows increasing interest in the keywords in our search strategy over the last decade (a). One third of articles included in this synthesis were published in 2020 (b).

Country	Count
UK	7
Germany	6
Spain	4
Taiwan	4
USA	3
Austria	3
Greece	3
Kenya	3

Figure 2. Countries that were the focus of 3 or more articles.

PERCEIVED COSTS VERSUS PERCEIVED VALUE

The most common barriers identified were lack of time and money. Across the literature, regardless of method, participants noted a lack of time to share data or to properly prepare datasets for sharing. (The level of preparation required might be related to concerns about data quality, see the

Concern regarding Data Quality / Mistakes section.) Yet a picture also emerged of an implicit cost-benefit trade-off.

Particularly for open government data, there was evidence of unfunded mandates: the responsibility of data sharing activities was added without associated resources of new staff or current job duties being reduced or shifted to allow for this extra duty. Shao et al. (2019) found that data sharing may be added to the duties of the government department, but that the associated tasks of implementing this mandate were not added to job descriptions. The result is employees are asked to do the work, but there is no formal accountability, unlike most of their other daily tasks. With the tasks not part of the formal incentive system in place, they are done ad hoc, if at all (Shao et al., 2019). Similarly, Wang et al. (2019) found that immediate managers were not supportive, regardless of levels of support at senior levels of government.

In one large survey, 57% of respondents reported the lack of a mandate as a key reason why they don't share their research data (Nicholas et al., 2020), another recurring theme (e.g. Schmidt et al., 2016; Roa et al., 2020; Safavorv, 2020).

In the absence of formal incentives or allocation of resources, the inherent value of the task becomes a more important factor. Yang et al. (2020) report "perceived value of open data" as having some of the most predictive power for the availability of open government data in Taiwan, and Gonzalez-Zapata and Heeks (2016) report a lack of understanding of the value or purpose of open data by government employees. Nuhu et al. (2020) report that a basic philosophical question, is data owned by the citizens or the government, is an important driver of data sharing in Ghana. If government employees consider data to be owned by the citizens, they are more likely to participate in making the data available. Wang et al. (2019) report that how an individual perceives and experiences barriers depends on their overall attitude toward open government data. Stieglitz et al. (2020) report that perceived career disadvantages and advantages were significant drivers of research data sharing intentions.

This cost-benefit calculation is sometimes explicit. For example, Yang et al. (2021) quote a participant as saying "If no one is interested in using the datasets we open, it doesn't make sense for us to implement open data…". 28% of respondents in another study identified "wasting time making [data] available" as a concern about data sharing (Aleixandre-Benavent et al., 2020). Stieglitz et al. (2020) include the role of perceived value as one of their hypothesis, and confirmed through survey research (n=955) that perceived value has a positive effect on researchers' intentions to adopt open data practices. (They also reported that professors perceived a higher level of effort in sharing data than research associates.) This blend of implicit and explicit cost-benefit analysis is consistent with other human factors research, including customer purchasing decisions (e.g. Zeithaml, 1988).

The lack of time and funding to share data is a real concern, and it is present throughout the literature on open data sharing. Yet allocating more time and funding won't be sufficient to address this barrier while open data remains a low priority relative to the other expectations of governments and researchers. The broader challenge is that humans will require extrinsic motivation (mandates, accountability) or intrinsic motivation (belief in the value of open data, the importance of transparency, and public ownership). Intrinsic motivation is preferable, because individual perspectives on the value or

necessity of open data influence not just behavior, but also how they perceive all of the barriers described in this report.

CULTURE (ORGANIZATIONAL NORMS)

The term "culture" was used frequently to describe barriers to sharing data in research and government contexts. From context, authors and participants were referring to their perception of the unwritten norms, attitudes, values, and overall environment of a workplace. There is at times a sense that the informal culture becomes codified through policies and practices. There is overlap with the cost/value barrier in that culture will inform one's understanding of the value of open data.

The most basic sense of culture is that sharing data is not a norm. Some studies identified a lack of positive examples (e.g Yang et al., 2020; Roa et al., 2020). Others directly indicated it wasn't the norm for their discipline (e.g. Abdullahi and Noorhidawati, 2020). Among articles addressing data sharing in government organization, a repeating theme was lack of interest or cooperation within the organization. Albano and Reinhard (2014) reported that civil servants had a low level of interest in terms of taking part in open data initiatives and credit "institutional factors such as structural and political factors".

Often the culture was not directly related to open data, but rather to related constructs like a culture of transparency in government. Cahlikova et al. (2020) credit a "culture of secrecy" for "reserved attitudes of senior officials and politicians toward open data". Wirtz et al. (2016) describe identify "perceived organizational transparency" as a driver of personal assessments on whether or not to share open data: if data custodians believe the government is not transparent, they are less likely to make government open data available. This chilling effect from political leadership has been recognized as a disinformation risk in the post-truth era (Colborne and Smit, 2020).

Some authors describe cultural factors that would make any kind of change challenging. Khurshid et al. (2020) suggest that a bureaucratic culture and a lack of political leadership are barriers to the growth of government open data, with a bureaucracy that is slow to change. Beno et al. (2017) identify resistance to change as a barrier for both academic and government open data, and report that "private entities were more likely to rate "resistance to change, risk averse culture" as a significant barrier".

One difficulty in the literature is the loaded term "culture" is rarely unpacked, even when change to culture is recommended (e.g. Gonzalez-Zapata & Heeks, 2020). it is helpful to know how a culture is perceived, organizational culture is well-known to be difficult to change. Unpacking how this perception is formed would be valuable, but this is a research gap. One source of detail on this environmental factor is Yang et al. (2020) who identified "facilitating conditions" as one of the factors that played a statistically significant role in decisions to make government open data available, especially in government units with less than two years of experience in open data.

There is also some evidence that we project our own beliefs onto the culture we perceive, with individual perspectives on who owns government data impacting how one perceives a data sharing imperative (Nuhu et al. (2020), Wang et al. (2019). Another study suggested that a culture of data sharing might not be sufficient: they describe respondents who report a belief that their workplace has a

strong culture of data sharing and that they are expected to share data, but they do not personally take part in data sharing (Jeng et al., 2016).

The picture that emerges on culture is unclear and unconvincing, but on balance it appears that facilitating data sharing requires that it be perceived as the organizational norm, but culture alone is not sufficient to address human barriers. An open data culture is necessary, but not sufficient.

REAL OR PERCEIVED SKILL AND KNOWLEDGE GAPS

A lack of technical skills, training, or general ability to share data was mentioned throughout the literature, for both types of open data we considered. Knowledge gaps occur when individuals don't know something they believe they need to know to share data. Skill gaps occur when individuals don't have experience or training to perform a task involved in the data sharing process. In some cases these gaps are a direct barrier to data sharing; in other cases, the barrier results from a fear of negative consequences or outcomes. For example, a perceived knowledge gap in understanding relevant legislation around privacy and data sharing contributes to a fear that one will share something that should not be shared publicly (Barry and Bannister, 2014). Mosconi et al. (2019) identify knowledge gaps as one of three critical gaps in opening science.

A common example is sufficient technical skill (e.g. Toots et al., 2017; Sá et al., 2016). In the context of research open data, there was less concern about technical skill working with data and more concern that terms like FAIR and the general vocabulary and landscape of open research were unfamiliar to researchers (e.g. Delikoura and Kouis, 2021). The lack of understanding or ability to provide appropriate metadata, or a need for appropriate standards was noted throughout the literature (e.g., Martin et al., 2013; Schmidt et al., 2015; Suhr et al., 2020). Schmidt et al. (2015) included respondents working in both government and research data, and noted that an important barrier was the difficulty in meeting standards around the release of data, which are higher than required for effective short-term internal use. The need to provide context around data came through consistently, and like many other barriers the authors suggested lack of training was a cause.

In addition to those who reported on their own need for training or additional skills development, Tenopir et al. (2020) suggest that some people working with data may not understand that they need support or know where to look for it.

Concern about skills and training was a consistent theme in articles focused on government open data in LMI countries (e.g. Safarov, 2020; Luthfi, 2020), though Abdullahi and Noorhidawati (2020) specifically noted the absence of concern about technical training and infrastructure in their interviews with researchers. The articles implied that respondents were comparing their resources to their peers in other nations. Yet the research in the other nations reports similar skill gaps.

Skill gaps are often mentioned in the same sentence as a lack of training or support (e.g. Delikoura and Kouis, 2021), and training is suggested as a mitigation for this and many of the other barriers. Yet Vicente-Saez et al. (2020) identify a lack of standards for training and skills; while many perceive a skill gap, it's not clear what a reasonable expectation should be. While there is some evidence on what data skills are broadly useful in the context of data literacy (Ridsdale et al., 2015) and which are most useful

for making use of open data (Smit et al., 2018) or in specific fields (Wilson et al., 2017), we're not aware of a systematic look at the skills required for sharing open data. The skills required, the level of training for each skill, and how to best deliver this training (In advance? Just in time? Online?) remains an open question.

FEAR OF MISUSE / LOSS OF CERTAINTY ON HOW DATA IS USED

Perhaps the most prevalent barrier is concern about the potential for various types of misuse of data. The potential misuses can be organized along a spectrum of severity, and the misuses can be intentional or incidental. While some articles frame this as concern about loss of control, we suggest a more accurate synthesis of the concerns is a loss of certainty: acknowledgment of the risk of data being used in ways the researcher does not approve of. Often these expectations are consistent with the researcher's personal philosophies, which of course are influenced by disciplinary norms, career status, and diverse other factors.

On the less-severe end of the spectrum, researchers were concerned that they **would not be appropriately credited** for the work they put into collecting and sharing data. Authors anticipate that others who use their data will fail to provide fair attribution including data citation, and their work will not be acknowledged by existing incentive / performance systems (Schmidt et al., 2015). This is connected to the challenge of culture: we do know there are issues with data citation and appropriately tracking and crediting the value created by data sharing (e.g., Silvello, 2018).

Related to appropriate credit is concern about "scooping": the fear that **others will benefit from my data before I do**. Again unique to research data, a consistent concern is the perceived risk of releasing data before publishing, or before publishing multiple papers from a dataset. The concern that one would put the effort into collecting and organizing data but not realize the full benefit of this work is recurring.

Several articles suggest this relates to loss of control (e.g. Mallasvik & Martins, 2020) : once research data is released publicly, there are few feasible mechanisms for ensuring the data is used appropriate and within norms for research, or within how the researcher would like it to be used. Melero & Navarro-Molina C. (2020) reported researchers identifying issues around data ownership and data and intellectual property rights when considering data sharing. Commercial exploitation of data was also mentioned as a concern, with different attitudes toward other researchers using data versus corporations profiting from data (e.g. Delikoura & Kouis, 2021). The picture that emerged across the articles discussing this concern is not that researchers needed control or feared losing personal control, but rather they did not have confidence that their data would be used for broadly- and personally-defined "good". We summarize this concern as a loss of certainty.

Many articles reported on concern about data misuse without defining precisely what types of misuse were a concern, and are reported in a **generalized data misuse** category (e.g. Elsayed & Saleh, 2018; Jao et al., 2018; Schmidt et al., 2015). A commonly mentioned potential misuse of data is the concern that data would be used without consideration of the proper context. This was often mentioned in terms of research data (Kurata et al., 2017; Mallasvik & Martins, 2020; J.T.; Hodonu-Wusu et al., 2020) and in context with a perceived skill gap around the provision and use of metadata for datasets to be shared. Hall (2013) reports that researchers fear that without proper context for the data they share,

there is a possibility that it could be misused and politicized inappropriately (e.g. Colborne & Smit, 2020).

Relatedly, Huang et al. (2020) interviewed government workers in Taiwan and assessed their willingness to comply with open government data policies based on a model relating to individual and organizational accountability. They note that people were less likely to comply fully if they had a high perception of risk to individual accountability (i.e., they will be blamed or punished for mistakes or misuse of open data). Those with a high perception of risk to organizational accountability were also less likely to fully comply, as they may have a fear that the government's reputation could suffer if data of low quality was released.

Jao et al. (2015) report concerns about sharing data it could lead to stigmatisation of specific populations. How science will be perceived when it enters the public discourse was not a traditional concern of most researchers, but there is growing recognition of this consideration (e.g. Llorente et al., 2019).

One surprising finding is the suggestion that fear of losing control of data can be a driver of data sharing, not a barrier. The perceived interest in one's data, and the potential increase in potential citations, may be a motivator for research data sharing Melero & Navarro-Molina (2020). Stieglitz et al. (2020) confirmed the influence of fear of data misuse (broadly defined) on data sharing intentions, but found it to be a positive influence in contrast to previous studies (Cragin et al., 2010). They identify this as an unusual contradiction, but suggest researchers may perceive that data that has been appropriately processed with metadata will counter data misuse because metadata provides clear context and citation paths.

Stieglitz et al. (2020) had a sufficiently large sample to distinguish among disciplines and career status. They reported that younger researchers had a greater intention of sharing data (consistent with other articles, including Tenopir et al. (2018)). They also found researchers in the humanities were less concerned about data misuse than life scientists (Stieglitz et al., 2020). In the public sector, cultural / memory institutions expressed particular concerns about how the resources they stewarded would be used (Estermann, 2013).

PERCEIVED RESPONSIBILITY FOR SENSITIVE OR PERSONAL DATA

Some data cannot be shared because it is sensitive: it includes information that is personal, or about sacred sites, or identifying animal nesting / feeding areas, or deserving special protection because of past data misuse (e.g. Indigenous Traditional Knowledge, see Cochran et al., 2008; Fidel et al., 2014; Proulx et al., 2021). There are necessary and legitimate limits on the open sharing of government and research data; we are not concerned with necessary, unavoidable barriers to sharing data.

Considering human factors and addressable barriers, concern about releasing sensitive data was a recurring theme in the literature but often with insufficient detail to assess if the concern was legitimate or spurious. In some cases, the concern was associated with knowledge/skill gaps, including about relevant privacy legislation (e.g. Albano & Reinhard, 2014; Barry and Bannister, 2014) or institutional policies (e.g. Zuiderwijk, 2014).

Concern about de-identified research data was also common, including concern about a lack of ability to correctly de-identify data (e.g. Delikoura & Kouis, 2021) and concern that end users would be able to combine different datasets to create identifying information that government did not intend to release (e.g. Barry & Bannister, 2014). This re-identification concern is fueled in part by awareness of the increasing power of machine learning but incomplete understanding of its capabilities and limitations.

What emerges is a picture of responsibility and accountability: both a perceived duty to research participants and subjects (or those served by government), and the understanding that mistakes would have consequences for both the data owner and the data subject (e.g. Yang and Wu, 2021). This is consistent with the norms in research and government. Combined with the skill gaps around de-identifying data, the result is an addressable barrier to sharing data that is very difficult to distinguish from situations where the nature of the data does strictly limit sharing.

CONCERN REGARDING DATA QUALITY / MISTAKES

One barrier mentioned often throughout the literature was the possibility that mistakes in data could be identified if data is shared openly. Issues around data quality were also identified as a barrier. These appeared in studies addressing both government open data as well as research data.

Berrone et al. (2016) surveyed government workers and noted that several managers reported worry that they may release data of questionable or low quality, such as incorrect data around specific procedures or the duration of government-issued licenses. This suggests an issue with the currency of data as well as perhaps a need for more resources to ensure that data is reviewed and updated where necessary.

While much of the discussion around the potential for mistakes to be found within data or around users identifying data quality issues seemed to focus on the impact it may have on the researcher or government entity, some surrounded the actual mechanisms for collecting data and making it available. One study interviewed researchers in biochemistry in Kenya and South Africa. Bezuidenhout (2019) reported researchers suggesting they worried that the data they shared would be viewed as lower quality by virtue of their geographic location and possibly less expensive equipment use than research done by researchers with more funding available to them. It's noted that these labs were not affiliated with large research consortia or other large funding sources in the area. Bezuidenhout notes that, "the awareness that their data were created using older equipment and methodologies, and less expensive reagents was raised by a number of participants as reasons why, even if they were to share, "no one would care"." (p. 22-23). Despite these concerns about data quality, it is noted that the researchers interviewed still report the risk of being scooped as having an impact on their willingness to share data (Bezuidenhout, 2019).

Another note regarding this barrier is that some reported that avoidance or the act of not sharing any data may be preferable to releasing data that has mistakes or is of low quality. Barry and Bannister (2014) reported data providers may opt to not share data than to share and open themselves to the risk of having people who access their data find mistakes they may have made.

MITIGATIONS

The barriers are heavily interwoven, so rather than report mitigations for each individual barrier, we produce a synthesized set of best practices from the literature, most addressing multiple barriers. 25 of the papers included in this synthesis did not include original possible mitigations in their scope of work (though many of these cited other papers that had considered mitigations).

One of the difficulties of mitigating the cost-benefit calculation barrier is that most costs (in terms of time and responsibilities) is realized by data owners, while the primary direct benefit is achieved by data users. Sharing data is, at present, an act of altruism and generosity. All participants benefit in a general culture of data sharing, as most are both data owners and data users, but in the first half of the adoption curve we cannot rely on this being true. Suggestions to address this imbalance can be organized into two broad categories: decreasing perceived costs and increasing perceived benefits. Many of the following approaches do one or both of these things.

Encouraging open data use by data owners. When individuals are data users, the benefit of data sharing is more clear (e.g. Abdullahi & Noorhidawati, 2020). The implied *quid pro quo* of sharing data because you have yourself benefited from others sharing their data taps into basic human psychology. This leads toward a culture of data sharing, where it's generally expected because it's good for everyone (including the data owner).

Encouraging open data use as a way to demonstrate value. Reducing barriers to using data makes re-use more common, and people using data is a demonstration of value. This can include examples, use cases, and steps for how to implement a use case in your own organizations (Zuiderwijk, 2014); sharing the code, notebooks, and other resources needed to make effective use of the data; and improving data search and discovery (Phillips & Smit, 2021).

Selective release of open data. Conradie & Choenni (2014) (cited in Beno et al., 2017) suggest there is not enough open government data use (benefit) to justify the costs of a blanket expectation of sharing, at least in the context of local government. To maximize return on investment, they address the cost side of the equation, suggesting only the data most in demand be released. Determining what data is in demand is a challenge, and raises the possibility of inequity to the "long tail" of data users using infrequently accessed datasets. It's also not clear how this would work for research data, where sharing isn't just about use but also about replication, transparency, and collegiality.

Making benefits more apparent. In the open government context, this meant evidence that sharing data would improve agency governance or better serve the public, which might include success stories or examples of how data was used.

Support improving the quality of data during the entire research process. The less work one has to do to prepare data for sharing, the more likely it is to happen (Zabijakin-Chatleska & Cekikj, 2020), and the more confidence a researcher has that it will be useful and used responsibly (Tenopir et al., 2018). Research data sharing cannot focus on the "last mile" of getting data into a repository.

Peer-led data sharing. One of the drivers of data sharing identified was seeing peers or leaders embrace data sharing. If your immediate context makes it appear that a healthy data culture is expected, you are more likely to participate in data sharing. In contrast, a third party advocating for increased data sharing in a field reinforces the belief that no one else is doing it, and may serve to decrease interest in data sharing. This does not preclude advocacy, but rather suggests the approach to advocacy should be one of invitation to join a healthy, vibrant data-sharing community. The desire to belong to a peer community, and to help people in one's community, was described as a strong motivating factor (e.g. Mallasvik & Martins, 2020).

Critically examine our reward system in the context of human psychology. The weight placed on seeing peers share data suggests the value of making data sharing visible, whether informally, casually, or through formal recognition mechanisms. While we often think of reward in terms of career progression or access to funding or publications or other substantive forms of recognition, at the level of individual psychology humans are motivated by surprisingly small rewards that generate a rush of dopamine (Arias-Carrión et al., 2010). Using this knowledge to make changes to human behavior through small adjustments with corresponding rewards has been widely studied, particularly in the

context of information systems (e.g. persuasive technology or captology) (Orji & Moffatt, 2018). Anyone who uses social media has encountered attempts to use this approach to enhance engagement (and ultimately increase profits for the social media company). Data sharing is set up to require substantial effort followed by time-delayed and indirect rewards, which means even if there are benefits we don't attribute them directly to the work we did to share data. We should be re-thinking how we offer micro-rewards and "nudge" behavior in desired directions, but the literature does not include evidence that this conversation is underway.

A careful balance of extrinsic (mandates) and intrinsic motivation. Data sharing mandates are commonly cited as a possible mitigation effort, along with forms of pressure that are difficult to ignore (like publishers, funding agencies, public opinion). This form of validation is inherently extrinsic; as described in the review of barriers, intrinsic motivation is preferable because individual perspectives alter how one assesses the value and the costs/barriers of sharing data. Extrinsic motivation may change behavior - and survey respondents indicated a lack of mandate was a big contributor in their decision to not share data - but long-term change and a culture of data sharing require intrinsic motivation. There is significant evidence that ideology - whether the value of reproducibility described earlier, or a general commitment to open science - is a key motivator (Vicente-Saez et al., 2020). There is some evidence that these ideals are more motivating later in one's career (Elsayed & Saleh, 2018). There is some evidence that the drivers of data sharing exert influence at different stages: early on, when there is little data sharing, leadership, authority, and mandates are more influential; later in data sharing projects, the behavior of peers is more influential (Yang & Wu, 2021). This suggests an initial and short-term role for extrinsic motivation may be desirable. We also note that the value of intrinsic motivations has been described since at least 2016 (Schmidt et al.), but while there has been progress the culture of data sharing has not been established purely through intrinsic motivations.

Building a data-sharing culture. Several articles describe the need for a "culture of data sharing", but don't tackle what it would take to build such a culture. Cahlikova & Mabillard (2020), for example, simply recommend "incentive-based policies". Yang and Wu (2020) describe a need for "facilitating conditions", a set of resources and training that are made available, but don't define what these should look like. Some suggest incentives might be financial (Abdullahi & Noorhidawati, 2020), but a survey of food scientists (n=101) found "the most important motivations for publishing research data were … facilitating the reproducibility of the research, increasing the likelihood of citations of the article, and compliance with funding body mandates" (Melero & Navarro-Molina, 2020). Reproducibility as an essential value shows up frequently (e.g. Nicholas et al., 2020).

Training. In the barriers section, we discussed the lack of a clear standard for the necessary skills for managing research data. There was some evidence that in the context of managing research data, the perceived skill gap is often larger than the actual skill gap, and that much of the gap is jargon (Aleixandre-Benavent et al., 2020). This has three big implications: first, the scale of the intervention needed re: training is likely small and might be suitable for a nudging-type approach; second, the intervention should focus on RDM-specific skills and jargon like data anonymization, FAIR principles, data repositories (Delikoura & Kouis, 2021); and third, clearly defining the skills required could itself be a mitigation of a training barrier.

Sharing data as a means for building connections and networks. This was seen as motivating in some studies (e.g. Bezuidenhout & Chakauya, 2018; Roa et al., 2020), which suggests a transition from the psychology of scarcity (hoarding, value is in the resource I hold) to the psychology of plenty (sharing, value is in how the resource connects me to others). It's unclear how to operationalize this mitigating factor.

LIMITATIONS

The search strategy used in the current synthesis identified a larger number of studies than was anticipated. However, it is still a relatively small number of studies, which doesn't allow for robust comparison using statistics or even qualitative comparison depending on the different methods, geographies, fields of study, questions, and other parameters of the studies.

For example, while there appears to be a number of articles addressing open data in Brazil, and two addressing Kenyan data, this is not a large enough sample to draw larger conclusions about those populations.

We included all of the articles in this synthesis on a mostly equal basis: we did not weight based on the strength and generalizability of the evidence, mostly because there was very little research that was broadly generalizable. This was necessary, but limits the synthesis to being suggestive of possible research directions based on preliminary evidence.

Several articles were excluded due to a language barrier. The team working on this project were only able to review articles that could be found in English. Several articles written in Portuguese had to be excluded, as well as a few in Japanese. It's possible that there were valuable insights or potential comparisons to be found in this research.

This synthesis excluded grey literature for scope reasons. it's possible that there is more information on data sharing and human factors in grey literature.

It bears repeating that as a scoping review, we are reporting what is in the literature, but not making any claims about the reliability, validity, or applicability of the findings.

IMPLICATIONS

The picture that emerges from this review of the literature is that the applicability of individual psychology to data sharing behavior is not well-understood. There is some recognition of the role individual psychology plays, but it is very difficult to separate from institutional and systemic factors, with factors like organizational structures, incentive systems, resource availability, and perceptions of institutional culture. There has been little work seeking to treat these as separate, and a clear understanding of psychology and emotion is missing from the literature.

Despite being lumped together in much of the literature, there are examples where individual- and system-level perspectives are misaligned. While individuals described concern regarding errors or quality issues being discovered, this type of discovery is described as a benefit of sharing open data at a social and institutional level. Although one can imagine this concern is tied to human factors such as

imposter syndrome (Clance & Imes, 1978) or performance anxiety (Holland et al., 2016), no explicit link has been investigated.

We frame the individual decision to share data as a cost-benefit analysis, and identify useful mitigations to overcome perceived barriers (i.e. concerns that increase the costs, and decrease the perceived benefits). These mitigations have direct implications for those engaged in supporting, exhorting, or mandating individual data owners toward sharing data. However, we found very little data on whether these suggestions are likely to be effective.

Despite the framing of sharing decisions as rational processes, the barriers identified in this review seem more closely linked to baser instincts: fear, anxiety, sense of loss, and the need for a sense of certainty and place. While there are rational reasons to be concerned about data sharing, even these can be skewed by personal opinions on the value of open data. This suggests we need to re-examine how we study individual perspectives on data sharing. It's more difficult to counter negative emotions (fear, anxiety, grief) than it is to provide information to a purely rational actor assessing cost-benefit.

The role of emotion means we need to critically examine our reward system, including how to ensure rewards are immediate and direct. Rewards for sharing data are often expressed as building a culture of data sharing, career progression, future access to funding, or future publications: all time-delayed and indirect benefits. Small but immediate and direct rewards may be more effective, but this approach needs to be explored.

We are convinced, based on the literature reviewed, that data management sharing incentives need to carefully balance extrinsic (mandates) and intrinsic motivation. Extrinsic motivation is difficult to ignore (like publishers, funding agencies, and other directives). Intrinsic motivation is preferable because individual perspectives on the value of sharing data alter how one assesses the benefits versus the costs of sharing data. Extrinsic motivation may change behavior in the short-term, but long-term change and a culture of data sharing require intrinsic motivation. There is room for both forms of motivation, but we need much more research on how to impact intrinsic motivation, and on the correct balance of motivation types.

CONCLUSION

Owners and custodians of open data have the digital skills necessary to participate in the data sharing required for a modern economy, but there are concerns about the motivation / attitude toward data sharing, and the level of systemic support for individuals. Data sharing may be impeded by a combination of natural human emotional responses and a cost-benefit analysis that assigns costs on the individual and benefits to a broad community.

We have described a set of barriers synthesized from the limited published research that examines individual perspectives, with a focus on original data directly from individual data owners and custodians. We've also synthesized a set of possible mitigations to these barriers, but found little evidence on the effectiveness of these mitigations. Most importantly, we've identified substantial research gaps.

Research examining individual perspectives on data sharing is significantly incomplete and in its infancy. A substantial push for interdisciplinary research is required across the board, but we hope to see leadership from the funding agencies interesting in making research data more open, and the field of library and information science which has been too insular in its ownership of research on this subject. Understanding motivation for data sharing would benefit from the perspectives of economists, psychologists, neuroscientists, industrial engineers, and occupational therapists, to name only a few. Filling in the many research gaps identified will require a significant and ongoing effort, but is essential to move toward a digital economy where data is a tool for collaboration and networking.

REFERENCES

Abdullahi, K. A., & Noorhidawati, A. (2020). Perceptions towards research data sharing: A qualitative study of Nigerian academics. *Malaysian Journal of Library & Information Science*, *25*(3), 103–121. https://doi.org/10.22452/mjlis.vol25no3.7

Acord, S. K., & Harley, D. (2012). Credit, time, and personality: The human challenges to sharing scholarly work using Web 2.0. New Media & Society, 15(3), pp. 379-397. doi: 10.1177/1461444812465140

Albano, C. S., & Reinhard, N. (2014). Open Government Data: Facilitating and Motivating Factors for Coping with Potential Barriers in the Brazilian Context. In M. Janssen, H. J. Scholl, M. A. Wimmer, & F. Bannister (Eds.), *Electronic Government* (pp. 181–193). Springer. https://doi.org/10.1007/978-3-662-44426-9_15

Aleixandre-Benavent, R., Vidal-Infer, A., Alonso-Arroyo, A., Peset, F., & Ferrer Sapena, A. (2020). Research Data Sharing in Spain: Exploring Determinants, Practices, and Perceptions. *Data*, *5*(2), 29. https://doi.org/10.3390/data5020029

Andreoli-Versbach, P., & Mueller-Langer, F. (2014). Open access to data: An ideal professed but not practised, Research Policy, 43(9), pp. 1621-1633. doi: 10.1016/j.respol.2014.04.008

Arias-Carrión, O., Stamelou, M., Murillo-Rodríguez, E., Menéndez-González, M., & Pöppel, E. (2010). Dopaminergic reward system: a short integrative review. *International archives of medicine*, *3*(1), 1-6. doi: 10.1186/1755-7682-3-24

Barry, E., & Bannister, F. (2014). Barriers to open data release: A view from the top. *Information Polity: The International Journal of Government & Democracy in the Information Age*, *19*(1/2), 129–152. https://doi.org/10.3233/IP-140327

Berrone, P., Ricart, J. E., & Carrasco, C. (2016). The Open Kimono: Toward a General Framework for Open Data Initiatives in Cities. *California Management Review*, *59*(1), 39–70. https://doi.org/10.1177/0008125616683703

Bezuidenhout, L., & Chakauya, E. (2018). Hidden concerns of sharing research data by low/middle-income country scientists. *Global Bioethics*, *29*(1), 39–54. https://doi.org/10.1080/11287462.2018.1441780

Bezuidenhout, L. (2019). To share or not to share: Incentivizing data sharing in life science communities. *Developing World Bioethics*, *19*(1), 18-24. https://doi.org/10.1111/dewb.12183

Borgerud, C., & Borglund, E. (2020). Open research data, an archival challenge? *Archival Science*. https://doi.org/10.1007/s10502-020-09330-3

Burgelman, J-C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schouppe, M. (2019). Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century, *Frontiers in Big Data*, 2(43). doi: 10.3389/fdata.2019.00043

Cahlikova, T., & Mabillard, V. (2020). Open Data and Transparency: Opportunities and Challenges in the Swiss Context. *Public Performance & Management Review*, *43*(3), 662–686. https://doi.org/10.1080/15309576.2019.1657914

Campbell, E. G., Bendavid, E. (2002). Data-sharing and data-withholding in genetics and the life sciences: results of a national survey of technology transfer officers. J Health Care Law Policy, 6(2), pp. 241-255.

Clance, P. R., & Imes, S. A. (1978). The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention. *Psychotherapy: Theory, research & practice*, *15*(3), 241.

Cochran, P. A., Marshall, C. A., Garcia-Downing, C., Kendall, E., Cook, D., McCubbin, L., & Gover, R. M. S. (2008). Indigenous ways of knowing: Implications for participatory research and community. *American journal of public health*, *98*(1), 22-27.

Colborne, A., & Smit, M. (2020). Characterizing disinformation risk to open data in the post-truth era. *Journal of Data and Information Quality (JDIQ)*, *12*(3), 1-13. https://doi.org/10.1145/3328747

Conradie, P., & Choenni, S. (2014). On the barriers for local government releasing open data. *Government Information Quarterly*, *31*, S10-S17. https://doi.org/10.1016/j.giq.2014.01.003

Cragin, M. H., Palmer, C. L., Carlson, J. R., & Witt, M. (2010). Data sharing, small science and institutional repositories. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *368*(1926), 4023-4038. https://doi.org/10.1098/rsta.2010.0165

Curty, R. G., Crowston, K., Specht, A., Grant, B. W., & Dalton, E. D. (2017). Attitudes and norms affecting scientists' data reuse, PloS ONE, 12(12), e0189288. doi: 10.1371/journal.pone.0189288

Damalas, D., Kalyvioti, G., Sabatella, E., & Stergiou, K. (2018). Open data in the life sciences: The 'Selfish Scientist Paradox.' *Ethics in Science and Environmental Politics*, *18*, 27–36. https://doi.org/10.3354/esep00182

Darby, R., Lambert, S., Matthews, B., Wilson, M., Gitmans, K., Dallmeier-Tiessen, S., Mele, S., & Suhonen, J. (2012). Enabling scientific data sharing and re-use. *2012 IEEE 8th International Conference on E-Science*, 1–8. https://doi.org/10.1109/eScience.2012.6404476

Delikoura, E., & Kouis, D. (2021). Open Research Data and Open Peer Review: Perceptions of a Medical and Health Sciences Community in Greece. *Publications*, *9*(2), 14. https://doi.org/10.3390/publications9020014

Deloitte. (2020). The upskilling imperative: Building a future-ready workforce for the AI age. [PDF] Retrieved from https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/deloitte-analytics/ca-covid19-upskilling-EN-AODA.pdf.

Eckartz, S., Hofman, W., & Veenstra, A. (2014). A Decision Model for Data Sharing. 13th International Conference on Electronic Government (EGOV), Dublin, Ireland (pp. 235-264). doi: 10.1007/978-3-662-44426-9_21

Elsayed, A. M., & Saleh, E. I. (2018). Research data management and sharing among researchers in Arab universities: An exploratory study. *IFLA Journal*, *44*(4), 281–299. https://doi.org/10.1177/0340035218785196

Estermann, B. (2013). Are memory institutions ready for open data and crowdsourcing?: Results of a pilot survey from Switzerland. *Proceedings of the 9th International Symposium on Open Collaboration*, 1–3. https://doi.org/10.1145/2491055.2491075

European Commission. (2019). Cost-Benefit Analysis for FAIR Research Data: Cost of Not Having FAIR Research Data in Europe. Publications Office of the European Union. doi: 10.2777/02999

Fecher, B., Friesike, S., Hebing, M., Linek, S., & Sauermann, A. (2015). A Reputation Economy: Results from an Empirical Survey on Academic Data Sharing. DIW Berlin Discussion Paper, 1454(26S). Retrieved from http://www.diw.de/sixcms/detail.php?id=diw_01.c.497452.de.

Ferguson, L. (2014). How and Why Researchers Share Data (and Why They don't). Wiley Network. Retrieved from https://www.wiley.com/network/researchers/licensing-and-open-access/how-and-why-researchers-share-data-and-why-they-dont

Fidel, M., Kliskey, A., Alessa, L., & Sutton, O. O. P. (2014). Walrus harvest locations reflect adaptation: a contribution from a community-based observation network in the Bering Sea. *Polar Geography*, *37*(1), 48-68.

Foulonneau, M. (n.d.). *Open Data: Barriers, Risks and Opportunities*. Retrieved July 5, 2021, from https://www.academia.edu/8334526/Open_Data_Barriers_Risks_and_Opportunities

Gallagher, J., Orcutt, J., Simpson, P., Wright, D., Pearlman, J., and Raymond, L. (2015). Facilitating Open Exchange of Data and Information, Earth Science Informatics, (8)(4), pp. 721-739.

Gonzalez-Zapata, F., & Heeks, R. (n.d.). The Influence of the Transparency Agenda on Open Government Data in Chile. *2016 Conference for E-Democracy and Open Government (CeDEM)*. Retrieved August 15, 2021, from https://www.academia.edu/27006606/The_influence_of_the_transparency_agenda_on_open_government_data_in_Chile

Government of Canada. (2014). Directive on Open Government. Treasury Board of Canada Secretariat. Retrieved from https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=28108.

Government of Canada. (2018). DRAFT Tri-Agency Research Data Management Policy for Consultation. Innovation, Science and Economic Development Canada. Retrieved from https://www.ic.gc.ca/eic/site/063.nsf/eng/h_97610.html.

Government of Canada. (2019). Tri-Agency Open Access Policy on Publications. Canadian Institutes for Health Research. Retrieved from https://cihr-irsc.gc.ca/e/32005.html.

Grimshaw, J. (2010). A Guide to Knowledge Synthesis: A Knowledge Synthesis Chapter. Canadian Institutes of Health Research. Retrieved from http://www.cihr-irsc.gc.ca/e/41382.html.

Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, 11(2), 181–217. https://doi.org/10.1002/jrsm.1378.

Haini, S. I., Rahim, N. Z. A., Zainuddin, N. M. M., & Ibrahim, R. (2020). Factors influencing the adoption of open government data in the public sector: a systematic literature review. *Int. J. Adv. Sci. Eng. Inf. Technol*, *10*, 611-617.

Hall, N. F. (2013). Environmental studies faculty attitudes towards sharing of research data. *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries - JCDL '13*, 383. https://doi.org/10.1145/2467696.2467773

Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Budden, A. E., Batchellar, A. L., Duke, C. S., & Porter, J. H. (2013). Big data and the future of ecology. Front Ecol Environ, 11(3), pp. 156-162. doi: 10.1890/120103

Hill, R., & Stein, C. (2019). Scooped! Estimating Rewards for Priority in Science.

Hipsley, C. A., & Sherratt, E. (2019). Psychology, not technology, is our biggest challenge to open digital morphology data. Scientific Data, 6(41). doi: 10.1038/s41597-019-0047-0

Holland, C., Lorenzi, F., & Hall, T. (2016). Performance anxiety in academia: Tensions within research assessment exercises in an age of austerity. *Policy Futures in Education*, *14*(8), 1101-1116.

Hossain, M. A., & Chan, C. (2015). *Open data adoption in Australian government agencies: An exploratory study*. 13.

Huang, H., Liao, C. Z.-P., Liao, H.-C., & Chen, D.-Y. (2020). Resisting by workarounds: Unraveling the barriers of implementing open government data policy. *Government Information Quarterly*, *37*(4), 101495. https://doi.org/10.1016/j.giq.2020.101495

Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. Government Information Quarterly, 28(4), pp. 446–456. doi: 10.1016/j.giq.2011.01.004

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. https://doi.org/10.1080/10580530.2012.716740

Jao, I., Kombe, F., Mwalukore, S., Bull, S., Parker, M., Kamuya, D., Molyneux, S., & Marsh, V. (2015a). Involving Research Stakeholders in Developing Policy on Sharing Public Health Research Data in Kenya: Views on Fair Process for Informed Consent, Access Oversight, and Community Engagement. *Journal of Empirical Research on Human Research Ethics: An International Journal*, *10*(3), 264–277.

Jao, I., Kombe, F., Mwalukore, S., Bull, S., Parker, M., Kamuya, D., Molyneux, S., & Marsh, V. (2015b). Research Stakeholders' Views on Benefits and Challenges for Public Health Research Data Sharing in Kenya: The Importance of Trust and Social Relations. *PLOS ONE*, *10*(9), e0135545. https://doi.org/10.1371/journal.pone.0135545

Jeng, W., He, D., & Oh, J. S. (2016). Toward a conceptual framework for data sharing practices in social sciences: A profile approach. *Proceedings of the Association for Information Science and Technology*, *53*(1), 1–10. https://doi.org/10.1002/pra2.2016.14505301037

Khurshid, M. M., Zakaria, N., Arfeen, M., Rashid, A., Faisal, H., & Ahmad, M. (2020). *An Intention-Adoption Behavioral Model for Open Government Data in Pakistan's Public Sector Organizations–An Exploratory Study* (pp. 377–388). https://doi.org/10.1007/978-3-030-64849-7_34

Kucera, J. (2017). Analysis of Barriers to Publishing and Re-Use of Open Government Data. IDIMT- 2017 (Schriftenreihe Informatik). Retrieved from https://idimt.org/wp-content/uploads/2017/03/2017- Topic-K-Kucera-keynote-public-draft.pdf.

Kurata, K., Matsubayashi, M., & Mine, S. (2017). Identifying the Complex Position of Research Data and Data Sharing Among Researchers in Natural Science. *SAGE Open*, *7*(3), 2158244017717301. https://doi.org/10.1177/2158244017717301

Linde, P., Norling, E., Pettersson, A., Petersson, L., Pettersson, K., Stockmann, A., & Swartz, S. (n.d.). *Researchers and Open Data – Attitudes and Culture at Blekinge Institute of Technology*. 5.

Linek, S. B., Fecher, B., Friesike, S., & Hebing, M. (2017). Data sharing as social dilemma: Influence of the researcher's personality. PLoS ONE, 12(8), e0183216. doi: 10.1371/journal.pone.0183216

Llorente, C., Revuelta, G., Carrió, M., & Porta, M. (2019). Scientists' opinions and attitudes towards citizens' understanding of science and their role in public engagement activities. *PloS one*, *14*(11), e0224262.

Longo, D. L., & Drazen, J. M. (2016). Data Sharing (Editorial). N. Engl. J. Med., 374(3), pp. 276-277. doi: 10.1056/NEJMe1516564

Luthfi, A., Janssen, M., & Crompvoets, J. (2020). Stakeholder Tensions in Decision-Making for Opening Government Data. In B. Shishkov (Ed.), *Business Modeling and Software Design* (Vol. 391, pp. 331–340). Springer International Publishing. https://doi.org/10.1007/978-3-030-52306-0_23

Ma, R., & Lam, P. T. I. (2019). Investigating the barriers faced by stakeholders in open data development: A study on Hong Kong as a "smart city." *Cities*, *92*, 36–46. https://doi.org/10.1016/j.cities.2019.03.009

Mallasvik, M. L., & Martins, J. T. (2020). Research data sharing behaviour of engineering researchers in Norway and the UK: Uncovering the double face of Janus. *Journal of Documentation*, *77*(2), 576–593. https://doi.org/10.1108/JD-08-2020-0135

Martín-Martín, A., Thelwall, M., Orduna-Malea, E. et al. Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. *Scientometrics* 126, 871–906 (2021). https://doi.org/10.1007/s11192-020-03690-4

Martin, S., Foulonneau, M., Turki, S., & Ihadjadene, M. (2013). Risk analysis to overcome barriers to open data. *Electronic Journal of e-Government*, 11(1), pp. 348-359.

Martone, M. E., Garcia-Castro, A., & VandenBos, G. R. (2018). Data sharing in psychology. *American Psychologist*, 73(2), pp. 111-125. doi: 10.1037/amp0000242

Melero, R., & Navarro-Molina, C. (2020). Researchers' attitudes and perceptions towards data sharing and data reuse in the field of food science and technology. *Learned Publishing*, *33*(2), 163–179. https://doi.org/10.1002/leap.1287

Miner, R. (2014). The Great Canadian Skills Mismatch: People Without Jobs, Jobs Without People and MORE. Miner Management Consultants. Retrieved from http://www.minerandminer.ca/data/Miner_March_2014_final(2).pdf.

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. *PLoS Med* 6(6): e1000097. doi:10.1371/journal.pmed1000097

Mosconi, G., Li, Q., Randall, D., Karasti, H., Tolmie, P., Barutzky, J., Korn, M., & Pipek, V. (2019). Three Gaps in Opening Science. *Computer Supported Cooperative Work (CSCW)*, *28*(3–4), 749–789. https://doi.org/10.1007/s10606-019-09354-z

Nicholas, D., Jamali, H. R., Herman, E., Watkinson, A., Abrizah, A., Rodríguez-Bravo, B., ... & Polezhaeva, T. (2020). A global questionnaire survey of the scholarly communication attitudes and behaviours of early career researchers. *Learned Publishing*, *33*(3), 198-211. https://doi.org/10.1002/leap.1286

Nuhu, H. (n.d.). *(PDF) Social Enablers and Constraints Related to the Publication and Use of Open Government Data in a Developing Country*. Retrieved August 15, 2021, from https://www.researchgate.net/publication/347470091_Social_Enablers_and_Constraints_Related_to_the_Publication_and_Use_of_Open_Government_Data_in_a_Developing_Country

OECD. (2019). Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re- use across Societies. OECD Publishing, Paris. doi: 10.1787/276aaca8-en

Orji, R., & Moffatt, K. (2018). Persuasive technology for health and wellness: State-of-the-art and emerging trends. *Health informatics journal*, *24*(1), 66-91. https://doi.org/10.1177/1460458216650979

Perrier, L., Blondal, E., & MacDonald, H. (2020). The views, perspectives, and experiences of academic researchers with data sharing and reuse: A meta-synthesis. *PloS one*, *15*(2), e0229182. https://doi.org/10.1371/journal.pone.0229182

Pandey, J. L. (2016). The Challenge of Making Data Publicly Accessible. BioScience, 66(7), p. 535. doi: 10.1093/biosci/biw073

Pentland, A. S. (2013). The data-driven society. Scientific American, 309(4), pp. 78-83. doi: 10.1038/scientificamerican1013-78

Perception of Key Barriers in Using and Publishing Open Data | JeDEM - eJournal of eDemocracy and Open Government. (n.d.). Retrieved August 15, 2021, from https://jedem.org/index.php/jedem/article/view/465

Perry, J. S. (2017). What is big data? More than volume, velocity, and variety… IBM Developer Blog. Retrieved from https://developer.ibm.com/blogs/what-is-big-data-more-than-volume-velocity-and-variety/

Petticrew, M., & Roberts, H. (2008). Systematic Reviews in the Social Sciences: A practical guide. John Wiley & Sons, Inc.

Phillips, D., & Smit, M. (2021). Toward Best Practices for Unstructured Descriptions of Research Data. *Proceedings of the Association for Information Science and Technology*, *58*(1), 303-314. https://doi.org/10.1002/pra2.458

Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. PeerJ, 1, e175. doi: 10.7717/peerj.175

Poelen, J. H., Simons, J. D., & Mungall, C. J. (2014). Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets. *Ecological Informatics*, *24*, 148–159. https://doi.org/10.1016/j.ecoinf.2014.08.005

Poldrack, R. A., & Gorgolewski, K. J. (2014). Making big data open: data sharing in neuroimaging. Nature neuroscience, 17(11), pp. 1510-1517. doi: 10.1038/nn.3818

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme, Version 1, b92. doi: 10.13140/2.1.1018.4643

Proulx, M., Ross, L., Macdonald, C., Fitzsimmons, S., & Smit, M. (2021). Indigenous Traditional Ecological Knowledge and Ocean Observing: A Review of Successful Partnerships. *Frontiers in Marine Science*, *8*, 806. https://doi.org/10.3389/fmars.2021.703938

Richter, H. & Slowinski, P. R. (2018). The Data Sharing Economy: On the Emergency of New Intermediaries. IIC-International Review of Intellectual Property and Competition Law, 50(1), pp. 4-29. doi: 10.1007/s40319-018-00777-7

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin S., and Wuetherick, B. (2015). Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report. Dalhousie University. doi: 10.13140/RG.2.1.1922.5044

Roa, H. N., Loza-Aguirre, E., & Flores, P. (2020). Drivers and Barriers for Open Government Data Adoption: An Isomorphic Neo-Institutional Perspective. In K. Arai, S. Kapoor, & R. Bhatia (Eds.), *Advances in Information and Communication* (Vol. 1129, pp. 589–599). Springer International Publishing. https://doi.org/10.1007/978-3-030-39445-5_43

Roy, J., Sharaput, M., & Toze, S. L. (2019). Building Digital Capacity: Report on the Training Needs Analysis for the Government of Canada. http://hdl.handle.net/10222/76333

Sá, C., & Grieco, J. (2016). Open Data for Science, Policy, and the Public Good. *Review of Policy Research*, *33*(5), 526–543. https://doi.org/10.1111/ropr.12188

Safarov, I. (2020). Institutional Dimensions of Open Government Data Implementation: Evidence from Transition Countries. *Public Performance & Management Review*, *43*(6), 1359–1389. https://doi.org/10.1080/15309576.2020.1805336

Sayogo, D. S., & Yuli, S. B. C. (2018). Critical Success Factors of Open Government and Open Data at Local Government Level in Indonesia: *International Journal of Electronic Government Research*, *14*(2), 28–43. https://doi.org/10.4018/IJEGR.2018040103

Scheliga, K., & Friesike, S. (2014). Putting open science into practice: A social dilemma? First Monday, 19(1). doi: 10.5210/fm.v19i9.5381

Schmidt, B., Gemeinholzer, B., & Treloar, A. (2016). Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey. *PLOS ONE*, *11*(1), e0146695. https://doi.org/10.1371/journal.pone.0146695

Shao, D. D., & Saxena, S. (2019). Barriers to Open Government Data (OGD) initiative in Tanzania: Stakeholders' perspectives. *Growth and Change*, *50*(1), 470–485. https://doi.org/10.1111/grow.12282

Silvello, G. (2018). Theory and practice of data citation. *Journal of the Association for Information Science and Technology*, *69*(1), 6-20. https://doi.org/10.1002/asi.23917

Smit, M., Ridsdale, C., & Colborne, A. (2018). Proficient Use of Open Data Requires These Core Information Skills: An Open Data Community Perspective. In *Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI*. https://doi.org/10.29173/cais990

Steele, P. (2015). Closing Canada's Big Data Talent Gap [Report]. Canada's Big Data Consortium. Retrieved from https://www.ryerson.ca/content/dam/provost/PDFs/Big_Data_Talent_Gap.pdf.

Suhr, B., Dungl, J., & Stocker, A. (2020). Search, reuse and sharing of research data in materials science and engineering—A qualitative interview study. *PLOS ONE*, *15*(9), e0239216. https://doi.org/10.1371/journal.pone.0239216

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE*, 6(6), e1101. doi: 10.1371/journal.pone.0021101

Tenopir, C., Christian, L., Allard, S., & Borycz, J. (2018). Research Data Sharing: Practices and Attitudes of Geophysicists. *Earth and Space Science*, *5*(12), 891–902. https://doi.org/10.1029/2018EA000461

Toots, M., McBride, K., Kalvet, T., & Krimmer, R. (2017). Open Data as Enabler of Public Service Co-creation: Exploring the Drivers and Barriers. *2017 Conference for E-Democracy and Open Government (CeDEM)*, 102–112. https://doi.org/10.1109/CeDEM.2017.12

Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and Explanation. *Ann Intern Med.* 2018;169:467–473. doi: 10.7326/M18-0850.

Vertesi, J., & Dourish, P. (2011). The value of data: considering the context of production in data economies. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 533-542). https://doi.org/10.1145/1958824.1958906

Vesset, D., Olofson, C. W., Schubmehl, D., McDonough, B., Woodward, A., Stires, C., Fleming, M., Nadkarni, A., Zaidi, A., & Dialani, M. (2014). IDC FutureScape: Worldwide Big Data and Analytics 2015 Predictions [Report]. International Data Corporation.

Vicente-Saez, R., Gustafsson, R., & Van den Brande, L. (2020). The dawn of an open exploration era: Emergent principles and practices of open science and innovation of university research teams in a digital world. *Technological Forecasting and Social Change*, *156*, 120037. https://doi.org/10.1016/j.techfore.2020.120037

Wang, V., Shepherd, D., & Button, M. (2019). The barriers to the opening of government data in the UK: A view from the bottom. *Information Polity: The International Journal of Government & Democracy in the Information Age*, *24*(1), 59–74. https://doi.org/10.3233/IP-180107

Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U., & Dwivedi, Y. K. (2017). Open data and its usability: an empirical view from the Citizen's perspective. Information Systems Frontiers, 19, pp. 285- 300. doi: 10.1007/s10796-016-9679-1

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. doi: 10.1038/sdata.2016.18

Wilms, K., Brenger, B., López, A., & Rehwald, S. (2018). *Open Data in Higher Education – What Prevents Researchers from Sharing Research Data?* In *ICIS 2018 Proceedings*. 9. https://aisel.aisnet.org/icis2018/education/Presentations/9

Wilson, L., Colborne, A., & Smit, M. (2017). Preparing data managers to support open ocean science: Required competencies, assessed gaps, and the role of experiential learning. In *2017 IEEE International Conference on Big Data (Big Data)* (pp. 3984-3993). https://doi.org/10.1109/BigData.2017.8258412

Wirtz, B. W., Piehler, R., Thomas, M.-J., & Daiser, P. (2016). Resistance of Public Personnel to Open Government: A cognitive theory view of implementation barriers towards open government data. *Public Management Review*, *18*(9), 1335–1364. https://doi.org/10.1080/14719037.2015.1103889

Wolkovich, E. M., Regetz, J., & O'Connor, M. I. (2012). Advances in global change research require open science by individual researchers. Global Change Biology, 18(7), pp. 2012-2110. doi: 10.1111/j.1365-2486.2012.02693.x

Yang, T.-M., & Wu, Y.-J. (2016). Examining the socio-technical determinants influencing government agencies' open data publication: A study in Taiwan. *Government Information Quarterly*, *33*(3), 378–392. https://doi.org/10.1016/j.giq.2016.05.003

Yang, T.-M., & Wu, M.-C. (2020). An Exploration of Factors Influencing Taiwan Government Agencies’ Open Data Participation: A Multi-Group Analysis Perspective. *The 21st Annual*

International Conference on Digital Government Research, 356–358. https://doi.org/10.1145/3396956.3396966

Yang, T.-M., & Wu, Y.-J. (2021). Looking for datasets to open: An exploration of government officials' information behaviors in open data policy implementation. *Government Information Quarterly*, *38*(2), 101574. https://doi.org/10.1016/j.giq.2021.101574

Zabijakin-Chatleska, V., & Cekikj, A. (2020). Attitudes and Practices of Data Sharing and Data Preservation Among Social Science Researchers in the Republic of North Macedonia. *Balkan Social Science Review*, *15*(15), 251–275. https://doi.org/10.46763/BSSR20150252zc

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. *Journal of marketing*, *52*(3), 2-22. https://doi.org/10.1177/002224298805200302

Zuiderwijk, A., Janssen, M., Choenni, S., & Meijer, R. (2014). Design principles for improving the process of publishing open data. *Transforming Government: People, Process and Policy*, *8*(2), 185–204. http://dx.doi.org.ezproxy.library.dal.ca/10.1108/TG-07-2013-0024

Zuiderwijk, A., & Janssen, M. (2014). Barriers and Development Directions for the Publication and Usage of Open Data: A Socio-Technical View. In M. Gascó-Hernández (Ed.), *Open Government: Opportunities and Challenges for Public Governance* (pp. 115–135). Springer. https://doi.org/10.1007/978-1-4614-9563-5_8

Zuiderwijk, A., Shinde, R., & Jeng, W. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. *PloS one*, *15*(9), e0239283. https://doi.org/10.1371/journal.pone.0239283

Zuiderwijk, A., & Spiers, H. (2019). Sharing and re-using open data: A case study of motivations in astrophysics. *International Journal of Information Management*, *49*, 228–241. https://doi.org/10.1016/j.ijinfomgt.2019.05.024

PRISMA 2009 Flow Diagram
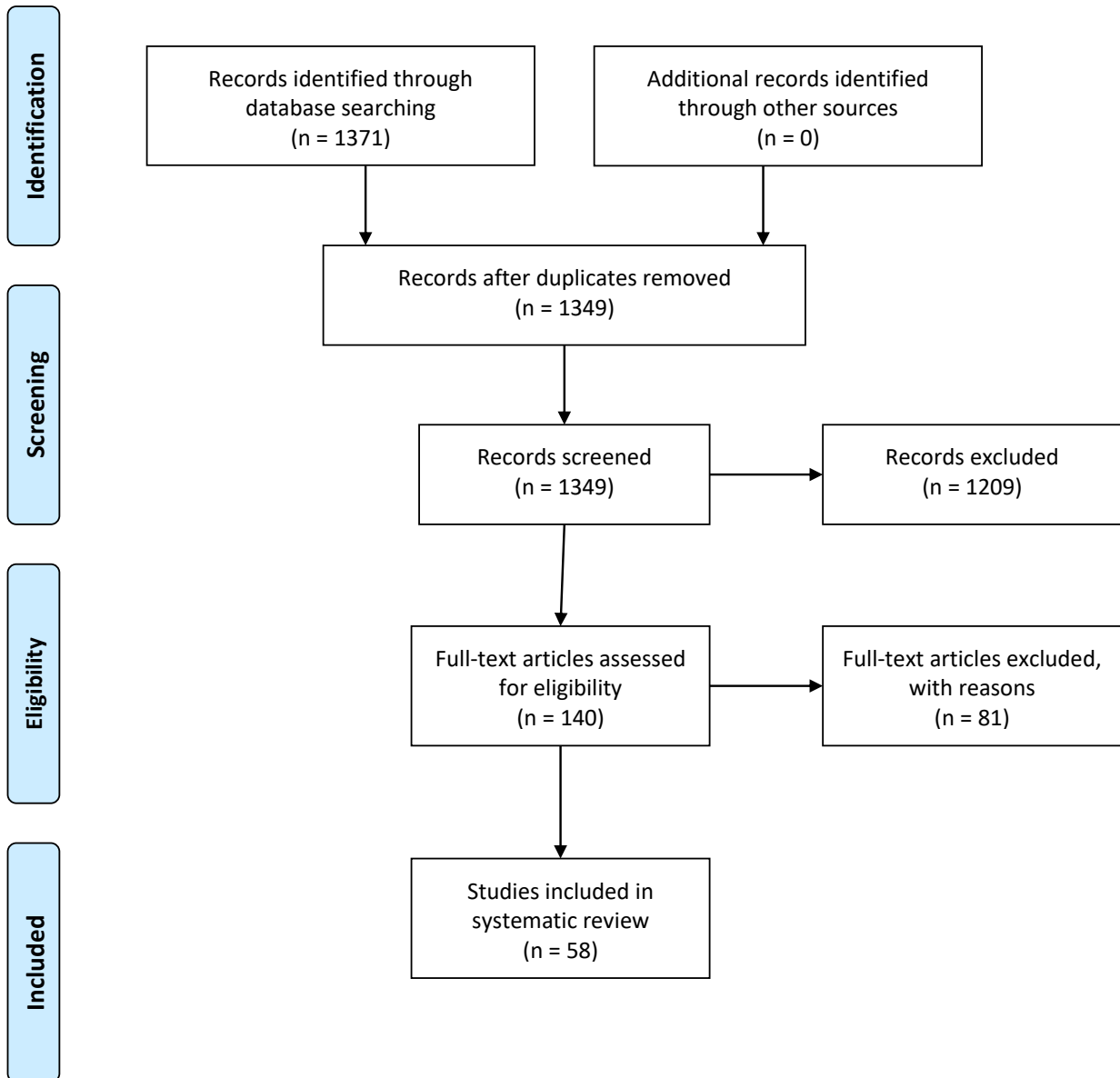
Identification

| Records identified through database searching (n = 1371) | Additional records identified through other sources (n = 0) |

Records after duplicates removed (n = 1349)

Screening

Records screened (n = 1349) → Records excluded (n = 1209)

Eligibility

Full-text articles assessed for eligibility (n = 140) → Full-text articles excluded, with reasons (n = 81)

Included

Studies included in systematic review (n = 58)

34

Authors	Title	Data Type	Data Quality / Mistakes	Time / Resources	Culture	Training / Skills	Risk of Misuse	Risk of Scooping	Sensitivity of Data	Lack of Credit
Abdullahi K.A., Noorhidawati A.	Perceptions towards research data sharing: A qualitative study of Nigerian academics	Research data		x	x				x	x
Albano C.S., Reinhard N.	Open government data: Facilitating and motivating factors for coping with potential barriers in the Brazilian context	Government open data	x		x					
Aleixandre-Benavent R., Vidal-Infer A., Alonso-Arroyo A., Peset F., Sapena A.F.	Research data sharing in spain: Exploring determinants, practices, and perceptions	Research data		x			x	x	x	x
Barry E., Bannister F.	Barriers to open data release: A view from the top	Government open data (Ireland)	x		x	x	x		x	
Beno M., Figl K., Umbrich J., Polleres A.	Perception of key barriers in using and publishing open data	Government and research and private	x	x			x			
Berrone P., Ricart J.E., Carrasco C.	The open kimono: Toward a general framework for open data initiatives in cities	Government open data	x				x		x	
Bezuidenhout L.	To share or not to share: Incentivizing data sharing in life science communities	Research data			x			x		
Bezuidenhout L., Chakauya E.	Hidden concerns of sharing research data by low/middle-income country scientists	Research data (life sciences)		x				x		
Borgerud C., Borglund E.	Open research data, an archival challenge?	Research data								
Cahlikova T., Mabillard V.	Open Data and Transparency: Opportunities and Challenges in the Swiss Context	Government open data	x		x					
Damalas D., Kalyvioti G., Sabatella E.C., Stergiou K.I.	Open data in the life sciences: The 'Selfish Scientist Paradox'	Research data (life sciences)						x		x
Darby R., Lambert S., Matthews B., Wilson M., Gitmans K., Dallmeier-Tiessen S., Mele S., Suhonen J.	Enabling scientific data sharing and re-use	Research data	x	x		x	x	x	x	

Author	Title	Category								
Delikoura E., Kouis D.	Open research data and open peer review: Perceptions of a medical and health sciences community in Greece	Research (health sciences)			x	x		x	x	x
Donald Shao D., Saxena S.	Barriers to Open Government Data (OGD) initiative in Tanzania: Stakeholders' perspectives	Government open data	x	x	x	x				
Elsayed A.M., Saleh E.I.	Research data management and sharing among researchers in Arab universities: An exploratory study	Research data			x	x		x		x
Estermann B.	Are memory institutions ready for open data and crowdsourcing? Results of a pilot survey from switzerland	Public (memory institutions)								x
Gonzalez-Zapata F., Heeks R.	The influence of the transparency agenda on open government data in Chile	Government open data				x				
Hall N.	Environmental studies faculty attitudes towards sharing of research data	Research data							x	x
Hodonu-Wusu J.O., Noorhidawati A., Abrizah A.	Malaysian researchers on open data: The first national survey on awareness, practices and attitudes	Research data						x	x	
Hossain M.A., Chan C.	Open data adoption in Australian government agencies: An exploratory study	Government open data								
Huang H., Liao C.Z.-P., Liao H.-C., Chen D.-Y.	Resisting by workarounds: Unraveling the barriers of implementing open government data policy	Government open data	x					x		
Janssen M., Charalabidis Y., Zuiderwijk A.	Benefits, Adoption Barriers and Myths of Open Data and Open Government	Government open data	x	x						
Jao I., Kombe F., Mwalukore S., Bull S., Parker M., Kamuya D., Molyneux S., Marsh V.	Research stakeholders' views on benefits and challenges for public health research data sharing in Kenya: The importance of trust and social relations	Research data						x	x	x
Jao I., Kombe F., Mwalukore S., Bull S., Parker M., Kamuya D., Molyneux S., Marsh V.	Involving Research Stakeholders in Developing Policy on Sharing Public Health Research Data in Kenya: Views on Fair Process for Informed Consent, Access Oversight, and Community Engagement	Research data								x
Jeng W., He D., Oh J.S.	Toward a conceptual framework for data sharing practices in social sciences: A profile approach	Research data				x				

Authors	Title	Category									
Khurshid M.M., Zakaria N.H., Arfeen M.I., Rashid A., Shehzad H.M.F., Ahmad M.N.	An Intention-Adoption Behavioral Model for Open Government Data in Pakistan's Public Sector Organizations–An Exploratory Study	Government open data	x			x	x				
Kurata K., Matsubayashi M., Mine S.	Identifying the complex position of research data and data sharing among researchers in natural science	Research data						x			
Linde P., Norling E., Pettersson A., Petersson L., Pettersson K., Stockmann A., Swartz S.	Researchers and open data - Attitudes and culture at blekinge institute of technology	Research data		x				x			
Luthfi A., Janssen M., Crompvoets J.	Stakeholder tensions in decision-making for opening government data	Government open data		x		x					
Ma R., Lam P.T.I.	Investigating the barriers faced by stakeholders in open data development: A study on Hong Kong as a "smart city"	Government open data		x		x	x		x		
Mallasvik M.L., Martins J.T.	Research data sharing behaviour of engineering researchers in Norway and the UK: uncovering the double face of Janus	Research data		x						x	x
Martin S., Foulonneau M., Turki S., Ihadjadene M.	Open Data: Barriers, risks and opportunities	Government open data	x								
Melero R., Navarro-Molina C.	Researchers' attitudes and perceptions towards data sharing and data reuse in the field of food science and technology	Research data			x		x	x			
Mosconi G., Li Q., Randall D., Karasti H., Tolmie P., Barutzky J., Korn M., Pipek V.	Three Gaps in Opening Science	Research data			x	x	x				
Nicholas D., Jamali H.R., Herman E., Watkinson A., Abrizah A., Rodríguez-Bravo B., Boukacem-Zeghmouri C., Xu J., Świgoń M., Polezhaeva T.	A global questionnaire survey of the scholarly communication attitudes and behaviours of early career researchers	Research data	x	x	x	x		x			

Author	Title	Type								
Nuhu H., Van Belle J.-P., Turpin M.	Social Enablers and Constraints Related to the Publication and Use of Open Government Data in a Developing Country	Government open data								
Poelen J.H., Simons J.D., Mungall C.J.	Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets	Research data (bio)						x		
Roa H.N., Loza-Aguirre E., Flores P.	Drivers and Barriers for Open Government Data Adoption: An Isomorphic Neo-Institutional Perspective	Government open data			x					
Sá C., Grieco J.	Open Data for Science, Policy, and the Public Good	Government open data	x			x				
Safarov, I	Institutional Dimensions of Open Government Data Implementation: Evidence from Transition Countries	Government open data	x	x	x	x	x			
Sayogo D.S., Yuli S.B.C.	Critical success factors of open government and open data at local government level in Indonesia	Government open data	x	x			x			
Schmidt B., Gemeinholzer B., Treloar A.	Open data in global environmental research: Findings from the community	Research data						x	x	x
Schmidt B., Gemeinholzer B., Treloar A.	Open data in global environmental research: The Belmont Forum's open data survey	Research data and government - any open data related to global environmental change		x			x	x	x	x
Stieglitz S., Wilms K., Mirbabaie M., Hofeditz L., Brenger B., López A., Rehwald S.	When are researchers willing to share their data? – Impacts of values and uncertainty on open data in academia	Research data								x
Suhr B., Dungl J., Stocker A.	Search, reuse and sharing of research data in materials science and engineering—A qualitative interview study	Research data	x	x			x		x	
Tenopir C., Christian L., Allard S., Borycz J.	Research Data Sharing: Practices and Attitudes of Geophysicists	Research data (geophysics)	x					x		x
Toots, M., Mcbride, K., Kalvet, T., & Krimmer, R.	Open data as enabler of public service co-creation: Exploring the drivers and barriers	Government open data	x		x	x			x	

Author	Title	Data type								
Vicente-Saez R., Gustafsson R., Van den Brande L.	The dawn of an open exploration era: Emergent principles and practices of open science and innovation of university research teams in a digital world	Research data					x			
Wang V., Shepherd D., Button M.	The barriers to the opening of government data in the UK: A view from the bottom	Government open data		x	x		x			
Wilms K., López A., Brenger B., Rehwald S.	Open data in higher education - What prevents researchers from sharing research data?	Research data (Germany)							x	
Wirtz B.W., Piehler R., Thomas M.-J., Daiser P.	Resistance of Public Personnel to Open Government: A cognitive theory view of implementation barriers towards open government data	Government open data			x					
Yang T.-M., Wu M.-C.	An exploration of factors influencing Taiwan government agencies' open data participation: A multi-group analysis perspective	Government open data			x					
Yang T.-M., Wu Y.-J.	Looking for datasets to open: An exploration of government officials' information behaviors in open data policy implementation	Government open data	x	x		x	x		x	
Yang T.-M., Wu Y.-J.	Examining the socio-technical determinants influencing government agencies' open data publication: A study in Taiwan	Government open data					x			
Zabijakin-Chatleska V., Cekikj A.	Attitudes and practices of data sharing and data preservation among social science researchers in the Republic of North Macedonia	Research data					x		x	x
Zuiderwijk A., Janssen M., Choenni S., Meijer R.	Design principles for improving the process of publishing open data	Government open data					x		x	
Zuiderwijk A., Spiers H.	Sharing and re-using open data: A case study of motivations in astrophysics	Research data					x			
Zuiderwijk, Janssen	Barriers and development directions for the publication and usage of open data: A socio-technical view	Research and public sector open data (Netherlands)	x				x		x	