Information and Organization 33 (2023) 100450

Contents lists available at ScienceDirect

Information and Organization

journal homepage: www.elsevier.com/locate/infoandorg

Data governance and digital innovation: A translational account of practitioner issues for IS research

Gregory Vial [*]

HEC Montreal, Department of Information Technology, 3000, chemin de la Côte-Sainte-Catherine, Montréal, Québec, H3T 2A7, Canada

ARTICLE INFO

Keywords:
Data governance
Digital innovation
Paradoxes
Governing data
Taxonomies
Typologies
Servitization

ABSTRACT

There is widespread agreement in research and practice that data governance is an instrumental element to help organizations leverage and protect data. IS research has observed that our practical and our scientific knowledge of data governance remains limited, and the increasing ability for organizations to generate, acquire, store, transform, process and analyze data calls for us to further identify and address issues on the topic. Striving to contribute to answer this pressing need, we argue that understanding the nature and the implications of governance *mechanisms* is of high importance as it is these mechanisms that effectively instantiate data governance in an organization. Building on our experience preparing and teaching workshops to 102 executives on the topic, we adopt a position of engaged scholarship and provide a translational account of our pedagogical experience on data governance, highlighting four outstanding themes for IS research. We argue that these four themes—(1) embracing data governance without compromising digital innovation; (2) enacting data governance through repertoires of mechanisms; (3) moving away from data governance toward *governing data*; and (4) moving away from a view of data at rest to adopt a service-based perspective on data governance—are highly relevant for practice and research. In our view, studying these themes will contribute to inform practitioners who often struggle with the implementation of comprehensive data governance programs and frameworks. At the same time, the ability to leverage theory to study these themes can help research generate novel theoretical contributions on data governance, helping future research on the topic.

1. Introduction

As organizations increasingly seek to use data to support digital innovation (Davidson, Winter, Wessel, & Winter, 2021), practitioner interest in data governance as a means to leverage and to protect data to create value (Tallon, Ramirez, & Short, 2013) has risen accordingly (Holt, 2021; Judah & White, 2020; Vial, 2020). To support these efforts, researchers have observed that there is a need to further generate theoretical and practical contributions on the topic of data governance (Alhassan, Sammon, & Daly, 2016; Benfeldt Nielsen, 2017; Benfeldt, Persson, & Madsen, 2020). Seeking to answer these calls, and inspired by the tradition of engaged scholarship (Mathiassen & Nielsen, 2008; Van de Ven, 2007), we rely on our experience preparing and delivering practitioner workshops on data governance to extend past contributions highlighting the importance of governance *mechanisms* (Peterson, 2004; Tallon et al., 2013).

Drawing from these experiences, we outline four overarching themes that we believe carry interesting potential to further our theoretical and practical understandings of the relationship between data governance mechanisms and digital innovation, and explore

* Corresponding author.
E-mail address: gregory.vial@hec.ca.

https://doi.org/10.1016/j.infoandorg.2023.100450
Received 31 October 2021; Received in revised form 6 December 2022; Accepted 24 January 2023
Available online 4 February 2023

the potential for IS research to bridge the gap between research and practice on the topic. These themes focus on (1) embracing data governance without compromising digital innovation; (2) enacting data governance through repertoires of mechanisms; (3) moving away from data governance toward *governing data*; and (4) shifting from a view of data at rest to adopt a service-based perspective on data governance.

In the following sections, we first present the approach undergirding this essay, followed by a brief overview of data governance, its current relevance, and the link between data governance and digital innovation. We then detail our four research themes, discussing their relevance for practice and the ability for IS research to inform their implications based on the use of established theoretical foundations, and provide concluding remarks.

2. Approach to engaged scholarship

In this essay, we build on pedagogical experience preparing and delivering six half to full-day workshops on the topic of data governance (see Appendix for additional details). These workshops were originally proposed by our institution's executive education institute in 2018 to answer a request expressed by practitioners to (1) learn on this topic and (2) assess whether their practices were consistent with successful patterns observed in other firms as well as those relayed in scholarly publications.

2.1. Designing the pedagogical experience: Translation from research to practice

To prepare material that was both rigorous and relevant for practitioners, we strived to build a *translational* (Barrett & Oborn, 2018) account of data governance leveraging academic literature. To identify key pedagogical material, we performed searches in online databases looking for peer-reviewed publications using the keyword "data governance", focusing on sources reporting empirical evidence as well as those providing frameworks that synthesize findings and contributions, and offering an opportunity to easily structure exchanges on this topic (e.g., Benfeldt et al., 2020; Khatri & Brown, 2010; Otto, 2011a, 2011b; Tallon et al., 2013). For example, the data governance framework proposed by Khatri and Brown (2010:149), the dimensions of data quality proposed by Strong et al. (1997:104) and the information governance model inductively developed by Tallon et al. (2013:168) all provide information that is both relevant and easily relatable for practitioners. These were supplemented with additional practitioner works (e.g., DAMA International, 2017; Eryurek, Gilad, Lakshmanan, Kibunguchy-Grant, & Ashdown, 2021; Ladley, 2019; Seiner, 2014) which we expected would be primary reference material for practitioners. Since 2019, the workshops have drawn a total of 102 participants (see Table 1) working in organizations of various sizes (from start-ups to small and medium enterprises, to firms employing more than 10,000 people) and operating in a variety of industries (e.g., artificial intelligence, financial services, utilities, and government agencies, among others).

2.2. Learning from the pedagogical experience: Translating back from practice to research

The workshops provide a platform to disseminate and transfer scientific knowledge, as well as to reflect and exchange with the instructor[1] and other participants on the design, implementation, and management of data governance. Each workshop offers an opportunity to engage with participants to illuminate common pain points experienced by practitioners.

In delivering each workshop, we adopt the position of an observer gifted with the opportunity to engage with practitioners—"knowledgeable agents" (Gioia, Corley, & Hamilton, 2013:17) who are best suited to relay their own concerns and to draw attention toward important issues for research—immersed in a phenomenon, and to reflect back on those experiences, thereby actively deepening our understanding of practitioner concerns on a topic (Makin, 2021; Marabelli & Vaast, 2020). During and following each workshop, we take handwritten notes of salient points brought up by participants, allowing us to account for the emergence of new issues that require a revision for the next workshop, and highlighting areas where existing research may not offer clear guidance. Following each workshop, participants are also invited to complete a survey inquiring on their experience, including their overall appreciation for the workshop, its material, the instructor, and to provide additional comments and questions that could inform future workshops.

While not designed as a research project, this process reflects an iterative approach to designing, evaluating, and revising an artifact—the workshop and its accompanying teaching material—consistent with the spirit of design science research (Hevner, 2007:89; Hevner, March, Park, & Ram, 2004:78). Over time, the translational account originally envisioned to create the workshops became the starting point of an iterative process wherein each workshop is an occasion to confront research with practice and to leverage the teaching experience to eventually generate a translation back from practice to research, as illustrated in Fig. 1.

3. Data governance and digital innovation

3.1. Defining data governance

As is the case for other concepts that have emerged from the realm of practice, there is no single definition for data governance

[1] All workshops were prepared and delivered by the author.

Table 1
Overview of workshop participants.

Sector	Industry	Number of organizations	Number of participants	Illustrative job titles
Private	AI solutions	1	1	Principal advisor, data governance; Head advisor, HR governance; Director, HR analytics, Head of IT, President; VP of IT and business processes; Information management analyst; Director of finance.
	Banking & finance	3	8	
	Digital services	1	1	
	Entertainment	3	3	
	Finance advisory	1	1	
	Food processing	1	1	
	Healthcare	1	1	
	Insurance	3	3	
	IT consulting services	6	6	
	Manufacturing	2	2	
	Publishing	2	2	
	Retail	1	1	
	Software provider	2	3	
	Aerospace	1	2	
	Government agency	5	38	
Public	Municipal government	2	1	Strategic planner; Director of IT; Solutions architect; Head of planning and development; Director, data strategy & governance.
	Postsecondary institution	2	5	
	Regulatory agency	1	2	
	Transport agency	2	2	
	Utility	1	6	
	Cultural sector	1	7	
Non-profit	Fundraising	1	1	Director, innovation, and data analytics; Data scientist; Head of IT.
Undisclosed	Undisclosed	Not available	5	Advisor, governance; Head of IT, Functional analyst.

Notes:

- Details on each workshop participant are withheld to protect their anonymity. Job titles were provided by participants upon registration; we include some to illustrate the audience attending the workshops.
- Some participants did not disclose their organization when they registered. They are reported in the "Undisclosed" row. Most did, however, provide a job title upon registration.

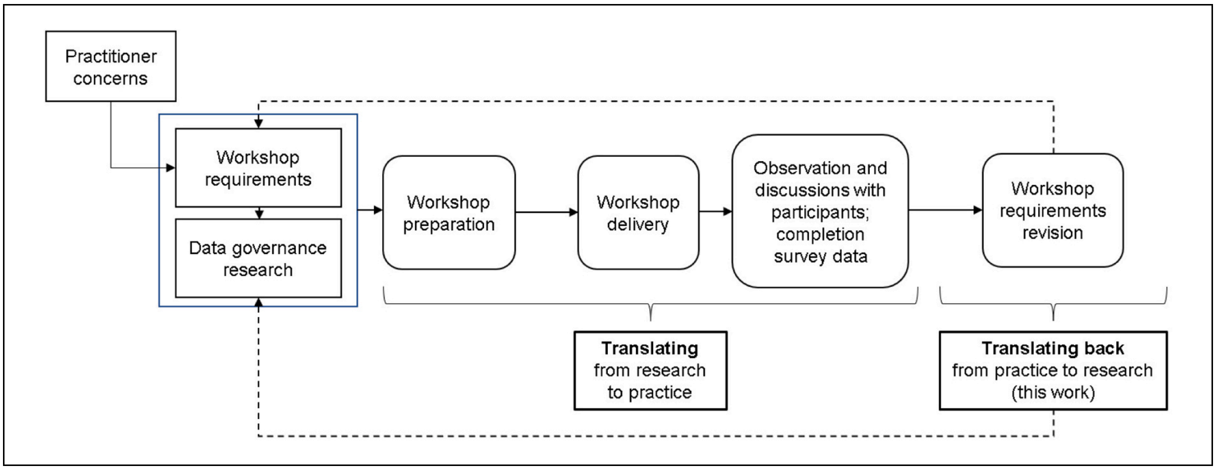

Fig. 1. Building translational accounts of data governance.

(Alhassan et al., 2016; Benfeldt et al., 2020; Benfeldt Nielsen, 2017; Davidson et al., 2021). During the workshops, we provide an initial definition of data governance as "a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods”[2] (The Data Governance Institute, 2021). This definition, like others found in practice (e.g., Holt, 2021; Ladley, 2019) and in research (Benfeldt et al., 2020; Khatri & Brown, 2010) builds on the broader concept of governance and, more specifically, the concept of IT governance which “determines who makes the IT decisions and assigns accountability for the outcomes” (Weill & Ross, 2005:26).

This view of data governance has the undeniable advantage of being coherent with other governance concepts that have long been used in organizations, as observed in other works (e.g., Tallon et al., 2013:150). For practitioners who have experience with IT governance, this means applying the same logic to a different object. In line with this view, Khatri and Brown (2010) built on the five key decision domains of IT governance developed by Weill and Ross (2005) and proposed data principles, data quality, metadata, data access, data lifecycle as important decision domains for data governance. The authors further suggest that an organization can design a *data governance matrix* for each decision domain (p. 151). This approach is practical because it allows managers to not only design, but also to diagnose data governance issues in their organizations by creating a level of abstraction that facilitates the structuring of ideas. For example, a lack of centralization for data quality standards may cause issues when data engineers seek to leverage data to create predictive models, only to find that the quality of data sources is inconsistent due to differing data quality standards across departments.

At the same time, this degree of abstraction is not without issues. Specifically, a recurring criticism reported by workshop participants is that it may be too abstract to effectively help managers make important decisions about *what* to do as well as *how* to do it. In short, it may be helpful to explain governance as an organizational element, but it is difficult to relate to *governing*. This view is echoed in the work of Tallon et al. (2013) who proposed an inductive theory of information governance. Based on interviews conducted in 2008 and 2009 with thirty executives, the authors identify twelve data governance *practices* that we can relate to other data governance frameworks (e.g., “Enforce retention and archiving” as a practice related to the Data Lifecycle decision domain of the framework proposed by Khatri & Brown). These practices are classified according to three overarching categories of governance *mechanisms* (Peterson, 2004), each with their own degree of formalism relative to the others (see Fig. 2):

- *Structural* mechanisms focus specifically on the design of formal organizational elements to establish the decision rights and accountability of actors, e.g., through official roles (e.g., naming a Data Protection Officer per GDPR requirements) or data ownership responsibilities (Tallon et al., 2013:166).
- *Procedural* mechanisms emphasize the operational means that are put in place to ensure compliance with governance principles through the performance of specific tasks that can be aided by IT, such as establishing and monitoring access (Tallon et al., 2013:164).
- *Relational* mechanisms provide a less formalized means of ensuring that data governance principles are understood and enforced by actors. For example, consistent with the idea of user education presented by Tallon et al. (2013:165), one participant working in a large financial institution informed us that they put a mentoring program in place to ensure that junior data analysts understand the importance of preserving the security and the integrity of data early on. They found that this program was more effective than other types of mechanisms implemented in the past that emphasized control and compliance.

3.2. *Why data governance matters in the digital age*

It may appear surprising to see that data governance has become a “hot topic” of interest in recent years. Indeed, while some data governance frameworks date back to the early 2000s and research had followed suit (e.g., Khatri & Brown, 2010; Otto, 2011b), it appeared to remain largely peripheral to practitioner concerns up until the mid 2010s or so. Two interrelated factors explain the (re) emergence of data governance as a topic of high interest in practice. The first is the increasing ability for organizations to access, store, process and use data at unprecedented capacity and pace. Digitalization (Legner et al., 2017) is paving the way for a world where everything we do, every interaction we have, whether with other humans or machines, or even between machines, provides an opportunity to generate vast quantities of data at relatively low cost thanks to the ubiquity of digital technologies (Loebbecke & Picot, 2015; Newell & Marabelli, 2015). However, at such scale, data management techniques alone are not sufficient. Data governance provides an important buffer to organize the interface between the strategic view of data[3] as assets and the daily management of data that aims at operationalizing this view (Tallon et al., 2013).

The second factor is the realization that data are no longer the mere output of some process aided by technological artifacts. As Fig. 3 illustrates, organizations are now incorporating a view of data as both outputs of traditional processes and as inputs for new (or reengineered) processes that leverage digital technologies. For instance, in the past an organization would typically consider that the value of the data generated by their enterprise resource planning (ERP) system was lower than the products it helped to manufacture and sell. Nowadays, the same organization will often spend significant time and effort to try and leverage the data generated by their ERP through predictive or prescriptive analytics to support other parts of their business. The monetization of these data (Wixom & Ross, 2017)—whether they are used internally or sold to third parties—can sometimes create far more value for the organization than

[2] Although we acknowledge that conceptual differences exist between *data* and *information*, literature on data governance does not typically distinguish between the two as information depends on data (Tallon et al., 2013). We use the term “information” to maintain consistency with relevant sources where applicable (Legner et al., 2020; Tallon et al., 2013) but the meanings of the two terms are the same in this context.

[3] In this essay we distinguish between references to data as a concept (using the singular form) and actual data (using the plural form), consistent with the etymology of the word *data* as a plural form.

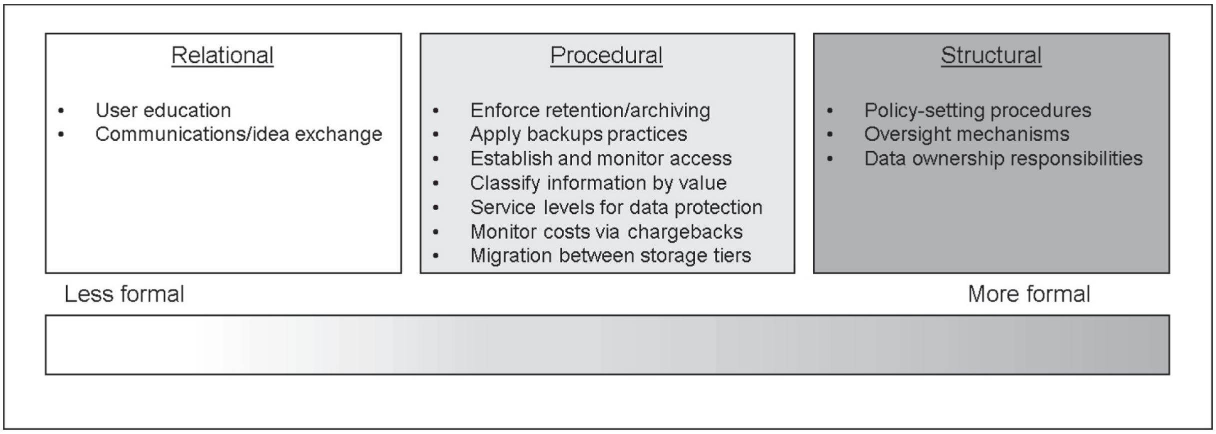

Fig. 2. Types of governance mechanisms (adapted from Tallon et al., 2013:168).

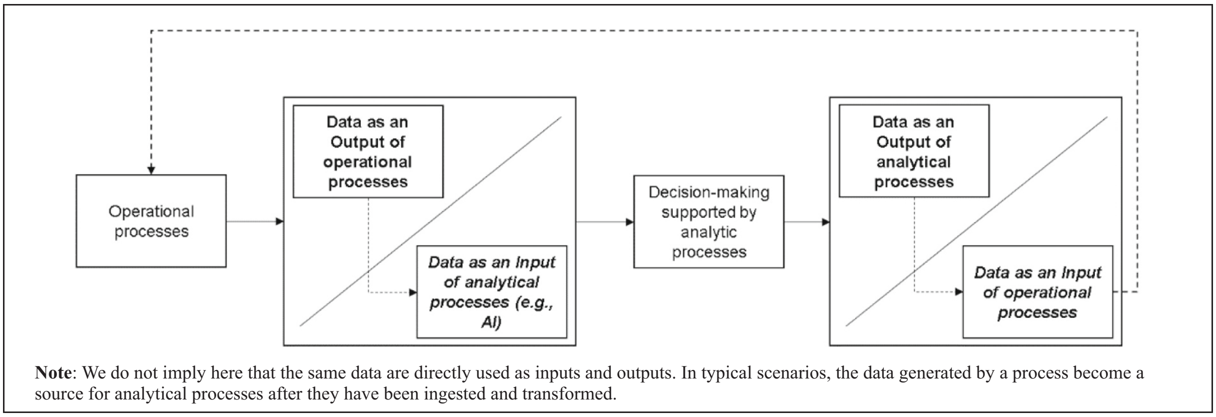

Fig. 3. The dual role of data as output and input.

the product itself. For instance, the virtual assistant Alexa is sold by Amazon at a low cost, but its use also provides crucial data to further enhance Amazon Web Services offerings (e.g., for the provision of speech to text services).

3.3. Innovating while protecting: The duality of data governance

IS literature argues that data governance seeks to achieve two objectives: "(1) to maximize the value of information to the organization by ensuring that information is reliable, secure, and accessible for decision making and (2) to protect information so that its value to the organization is not diminished through technology or human error, loss of timely access, inappropriate use, or misadventure" (Tallon et al., 2013:142). In theory, these two objectives make sense, especially in the context of digital innovation where the ubiquity and the generativity of digital technologies support the large-scale generation of data. In practice however, their simultaneous achievement is often experienced as highly challenging.

The ability to maximize the value of data—an important aspect often sought by digital innovation initiatives—is perceived as a desirable, albeit risky endeavor due in no small part to the legal and the ethical implications associated with this practice (Buytendijk & Heiser, 2016; Günther, Mehrizi, Huysman, & Feldberg, 2017; Vial, 2019). At the same time, the degree of risk exposure afforded by data—both in terms of probability and magnitude—means that organizations often implement structural and procedural mechanisms that emphasize the second objective of data governance—to *protect* data—while limiting the potential to support digital innovation as a form of uncertain, exploratory endeavor that involves a part of serendipity (Nambisan, Lyytinen, Majchrzak, & Song, 2017). This is especially true since the potential to innovate using data is emergent and uncertain (Bourton, Wigley, & Williams, 2018), with the implication that data governance mechanisms in place at time t may not anticipate potential use cases for data at time $t + n$. As a result, data governance can be perceived as series of measures that hinder digital innovation because it constrains the ability to find innovative uses for data (Vial, Jiang, Giannelia, & Cameron, 2021). In line with this idea, Tallon et al. (2013) observed that while they did not observe "a sense of resentment among users toward information governance policies or practices where resentment may be an indication of over-governance" (p. 168), "over-governance could limit information-led innovation, motivating users to work around policies and to take unnecessary risks with their information" (p. 167), highlighting the potential existence of "a curvilinear link

Table 2
A pragmatic outlook on research themes for data governance in organizations.

	Theme 1 Embracing data governance without compromising digital innovation	Theme 2 Enacting data governance through repertoires of mechanisms	Theme 3 From data governance to governing data	Theme 4 From systems to data to services
Practical issue	Organizations often emphasize data protection at the expense of data-driven value creation. This leads to perceptions that data governance hinders rather than enables digital innovation.	Data governance frameworks are presented as generic programs, but the enactment of data governance varies widely across organizations and contexts.	Data governance is often viewed as series of mechanisms implemented in organizations, at the expense of understanding the process of *governing* data.	Many organizations still conceive data as assets "at rest" in systems. Being able to innovate with data while protecting data requires engaging with the nature of data as assets *in flux*.
Relevance	Data governance aims at leveraging data to create value while ensuring its protection; maintaining a balance between these two seemingly contradictory objectives is challenging.	The duality of data governance, coupled with its contingent nature, require different combinations of mechanisms to protect data while fostering digital innovation.	The *design* of data governance only provides a partial account of data governance in an organization; the instantiation of this design in practice is important to understand *how* an organization protects and leverages data for digital innovation.	Organizations traditionally adopt a view of data based on the existence of physical and/or functional silos. Data-based digital innovation requires the removal of those silos, or at least the ability for data to seamlessly across those silos.
Relevance of data governance mechanisms	Structural mechanisms (e.g., security and access policies) are traditionally perceived as restrictive, while the potential for procedural mechanisms (e.g., data dictionaries, data lineage facilities) as well as relational mechanisms (e.g., employee coaching) to balance the dual objectives of data governance remains understudied.	Accounting for the relative contributions of each type of governance mechanisms to the achievement of both objectives of data governance is important to uncover and understand *patterns* of structural, procedure, and relational data governance mechanisms.	Data governance mechanisms are traditionally planned but studying their enactment in practice is important to understand their contributions to the dual objectives of data governance (e.g., employee coaching—a relational mechanism—may foster better alignment with policies and guidelines that exist as structural mechanisms).	Current trends emphasize the importance of services to enable data-based digital innovation. The provision of those services relies on structural, procedural, and relational mechanisms (e.g., dynamic access quotas, best practices) that depart significantly from the more traditional view of IT governance based on the existence of physical IT artifacts.
Potential practical solution	Implement data governance mechanisms that are designed with the explicit goal of balancing data protection and digital innovation (e.g., when defining the roles and responsibilities of data stewards).	Build data governance profiles to help an organization determine how it fares based on its own objectives, but also to compare itself to other organizations for benchmarking purposes.	Consider implementing data governance mechanisms in an iterative manner, within the context of data-driven initiatives to ensure that those mechanisms are quickly evaluated and adjusted if needed.	Implement digital services embedding procedural mechanisms (e.g., via APIs) that are consistent with data governance requirements for the organization.
How IS research can help	Research can draw attention to the possibility to espouse the two objectives of data governance as paradoxical, fostering the implementation of mechanisms (e.g., data stewards) that can help to reconcile both objectives.	Research can help to develop a conceptualization of data governance as repertoires of mechanisms that form configurations that contribute to the achievement of organizational outcomes.	Like strategy, data governance incorporates both planned and emergent components, calling for approaches that are closer to the practice of governing data and its impact on everyday work.	Research provides the conceptual scaffolding and the empirical evidence supporting an approach to governing data based on the provision and the orchestration of digital services.
Potential conceptual foundation (s)	**Paradoxes and tensions**; **Paradoxical thinking** (e.g., Gaim, Wåhlin, Pina e Cunha, & Clegg, 2018; Lewis, 2000; Poole & Van de Ven, 1989; Putnam, Fairhurst, & Banghart, 2016; Quinn & Cameron, 1988)	**Taxonomies** (Nickerson, Varshney, & Muntermann, 2013; Oberländer, Röglinger, & Rosemann, 2021); **Typologies** (e.g., Doty & Glick, 1994; Gregor, 2006)	**Practice perspective** (e.g., Jarzabkowski, Lê, & Feldman, 2012; Peppard, Galliers, & Thorogood, 2014); **Ostensive and performative aspects of organizational elements** (e.g., Latour, 1986) **and routines** (e.g, Feldman & Pentland, 2003);	**Orchestration** (e.g., Maruping & Matook, 2020); **Servitization** (e.g., Schüritz, Seebacher, Satzger, & Schwarz, 2017)

between information governance and firm performance" (p. 168).

This idea of having to strike a balance between value creation and data protection is a recurring issue reported by our workshop participants. Senior executives sometimes view data governance as a necessary element to "reign in" rogue digital innovation initiatives that carry more potential for liability than to create and capture value. Even in instances where the emphasis is set on value creation through the enforcement of enterprise-wide data quality standards, digital innovation can be stifled. For example, the implementation of centralized data repositories that involves the creation of complex data transformation and quality assurance pipelines may appear enticing to address data quality issues. At the same time, the processes that support this objective are often cumbersome and lead to a degree of centralization that can slow down the pace at which new sources of data or data transformation pipelines can be implemented. Data lakes are sometimes built and managed as traditional enterprise data warehouses, leading to scenarios where data consumers cannot easily and rapidly explore the potential for data to support digital innovation (e.g., Inmon, 2016; Tyagi & Demirkan, 2016).

4. Translating practitioner issues into research themes

Our experiences translating scholarly knowledge on data governance for practitioners allows us to make two key observations. First, academic literature provides an important foundation to structure discourse on data governance. In practitioner literature, the concept often remains vague, falling quickly into series of prescriptions to implement data governance programs that rely heavily on structural mechanisms, albeit without offering the opportunity to reflect on the rationale behind the implementation of data governance based on the need to encourage digital innovation and to protect data. Empirical evidence as well as conceptual frameworks offer insightful, actionable knowledge that not only contributes to establishing common grounds between the realms of research and practice, but also among practitioners who are otherwise unable to communicate with one another unless they refer to the same data governance framework.

At the same time, our experience preparing and delivering these workshops reveals that there is still much to be learned on this topic and that research has an opportunity to leverage theory to bridge this gap. To that effect, we present four overarching themes that have emerged over time based on our pedagogical experience (see Table 2). Each theme is based on the existence of recurring practical issues reported during our workshops. Within each theme, we explore the relevance of issues for practice, highlight the implications of this theme for governance mechanisms, and explain how IS research can contribute to inform this theme through the leveraging of relevant theoretical foundations.

4.1. Theme #1: Embracing data governance without compromising digital innovation

The first theme highlights the overemphasis often put toward a view of data governance based on compliance and controls (e.g., DAMA International, 2017; ISO, 2017) to maximize data protection at the expense of value creation. This view is reflected in works that draw parallels with other, highly institutionalized fields where compliance and control drive the entire execution of work (e.g., Ladley, 2019:17).

We argue that an interesting way to address this dominant perception requires that practitioners embrace the dual objectives of data governance as paradoxical (Winter & Davidson, 2019). Indeed, consistent with conceptual arguments on the existence of paradoxes as an inherent element of organizational life (Lewis, 2000; Quinn & Cameron, 1988; Smith & Lewis, 2011), we can conceptualize the two overarching objectives of data governance as "elements that seem logical in isolation but absurd and irrational when appearing simultaneously" (Lewis, 2000:760). The crux of the issue of course is that favoring one element (e.g., protecting data) may be done at the expense of the other (e.g., digital innovation). To resolve paradoxes, organizational scholars argue that actors need to rethink their approach to dealing with these issues and to engage in paradoxical thinking (Westenholz, 1993), that is, to embrace conflicting requirements of the constituting elements of a paradox and to look for ways to accommodate them both (e.g., Gregory, Keil, Muntermann, & Mähring, 2015), e.g., by striving to achieve ambidexterity (Andriopoulos & Lewis, 2009).

To engage with this theme, we could for example study the emerging role of the data steward and the ability for this role to embrace both data protection and digital innovation. Data stewards have been proposed to replace the traditional role of data owner. However, this role can be implemented in different ways. For instance, Dyché and Polsky (2016) proposed five models of data stewardship that reflect varying degrees of centralization (Khatri & Brown, 2010; Weill & Ross, 2005) for this role. Our discussions with workshop participants on these patterns highlight the fact that in practice, data stewards may be more or less inclined (and able) to encourage innovation based on their decision rights and accountability (e.g., the role may be temporary when data stewards are assigned on a per project basis). Nevertheless, the nature of this role and the combination of structural, procedural, and relational mechanisms that support its actual performance are relatively unclear in many organizations and in research.

This theme also offers interesting opportunities to further build theory on data governance. For instance, we may consider how the four approaches to study paradoxes among theories proposed by Poole and Van de Ven (1989)—opposition, spatial separation, temporal separation, synthesis—can be related to the dual objectives of data governance. To illustrate this point, consider the emergence of new forms of technological solutions that provide procedural governance mechanisms for organizations, such as Microsoft Azure Purview (Microsoft Corporation, 2021). Among other things, these tools seek to facilitate the building of comprehensive data dictionaries as well as to provide data lineage facilities that can help data consumers make sense of data—including their contextual quality and data accessibility quality (Strong et al., 1997; Wang & Strong, 1996)—and thus contributing to facilitate data-based value creation. At the same time, these tools increasingly rely on machine learning to automatically flag potentially sensitive data (e.g., Amazon.com, n.d.; Google.com, n.d.). In doing so they help to protect sensitive data and support structural mechanisms such as access and security policies. In theory, these tools may help organizations reach a new form of *synthesis* that resolves the contradiction between the two objectives of data governance. However, such tools are only starting to emerge and our knowledge of how they fit within a broader data governance initiative is limited as practitioners are still wrestling with the notion that data governance mechanisms can effectively contribute to both objectives simultaneously.

Another illustration is the implementation of different *curation zones* in data lakes (Terrizzano, Schwarz, Roth, & Colino, 2015) to help data consumers engage in unconstrained digital innovation in zones where the data are not curated while offering highly curated zones for exploitative endeavors where data transformation pipelines must be audited and data quality centralized. Conceptually, we may view this as a way to address the paradox of data governance through the creation of *separate spaces* embedded within IT. Clearly, IT has an important role to play here, but we can expect that the implementation of tools alone will not suffice. Understanding the conditions through which these tools prove useful for organizations as well as the relational mechanisms that support their adoption (e.g., user ratings of data source quality in a data dictionary) could help us extend current efforts that propose design principles (Legner, Pentek, & Otto, 2020) as well as prescriptions for practitioners (Otto, 2011b; Weber, Otto, & Österle, 2009) on the topic to generate theories focused on the paradoxical nature of data governance.

4.2. Theme #2: Enacting data governance through repertoires of mechanisms

A second theme that emerged from our teaching experience relates to the ability to understand data governance as an ensemble of mechanisms that contributes to the achievement of organizational outcomes. Research provides limited coverage on the conceptualization of data governance as a *bundle* of mechanisms implemented in an organization, with few works accounting for the contextuality of these mechanisms (Fadler & Legner, 2021) or the accumulation of knowledge on the topic of information management and data governance (Legner et al., 2020). In the realm of practice, the lack of consideration for relational mechanisms and the overemphasis put on the creation of structural and procedural mechanisms means that data governance initiatives often result in the drafting of official documentation and guides that can have legitimacy but little practical utility (Vial, 2020). Yet, as indicated in previous research (e.g., Tallon et al., 2013:144–146), data governance involves *combinations* of structural, procedural, and relational mechanisms that together should contribute to protect data and to create value. For workshop participants, this idea was often found important because characterizing these combinations could help to draw comparisons between organizations, a concern often expressed by those wondering whether their organization was effectively "doing enough" on the topic of data governance.

An interesting avenue would be to conceptualize series of data governance mechanisms put in place by organizational actors as *repertoires* of mechanisms. These repertoires could lead to the discovery of data governance archetypes that could be studied using configurational approaches such as fuzzy set qualitative comparative analysis (fsQCA) to develop taxonomies (Nickerson et al., 2013; Oberländer, Lösser, & Rau, 2019), or studied deductively to develop typological theories (Doty & Glick, 1994; Gregor, 2006; Guillemette & Paré, 2012) of data governance. It would also be useful for organizations to understand the extent to which their patterns of data governance mechanisms conform to a given archetype, which could in turn lead to specific interventions to alter the mechanisms in place to be closer or further from that archetype.

From a methodological standpoint, this endeavor could entail studying the constituting elements of the three types of data governance mechanisms. For instance, it may be informative to initially derive archetypes based on the ratio of structural to procedural to relational mechanisms in place within an organization. These could be augmented based on the decision domains they affect, as well as the locus of accountability of those decisions to build a data governance matrix (Khatri & Brown, 2010:151) that characterizes the patterns of data governance mechanisms in place within an organization.[4] During our workshops, participants working in highly regulated industries (e.g., banking) emphasized the importance of centralized, structural and procedural mechanisms they were required to put in place due to regulations, at the expense of relational mechanisms that were not perceived as useful because they bore no impact on regulatory compliance, suggesting a data governance archetype emphasizing data protection that hindered digital innovation. In contrast, other firms in related albeit less regulated related industries (e.g., fintech) relied on combinations of procedural and relational mechanisms supporting rapid growth.

Other works could further study the properties of these mechanisms and deductively derive their constituting actions and their contribution to the protection of data and the creation of value. Doing so would help us derive *ideal types* that explain the contribution of data governance to the realization of an organization's strategy. At the same time, this may provide an opportunity to contribute to the design of typological theories because the paradoxical nature of the objectives of data governance seemingly goes against the single dependent variable usually observed in these theories, highlighting the importance of *hybrid profiles* to achieve the best possible outcomes, and informing the idea that the relationship between governance and performance may be nonlinear (Tallon et al., 2013:168).

4.3. Theme #3: From data governance to governing data

Another important issue experienced by workshop practitioners relates to the enactment of data governance in practice. Extant guidelines, frameworks (e.g., Ladley, 2019; Seiner, 2014) as well as research on the topic (e.g., Khatri & Brown, 2010) represent data governance as an object identifying the loci of responsibility and accountability across decision domains, consistent with the concept of

[4] I would like to thank an anonymous reviewer for this insightful suggestion.

IT governance. While useful, this representation often results in a conceptualization of data governance as a form of planned object masterfully crafted through the design and the enforcement of mechanisms (e.g., Ladley, 2019). For reasons of legitimacy and visibility, these mechanisms are primarily structural or procedural in nature, consistent with the diffusion of regulative and normative institutions (Scott, 2008) crystallized in physical or logical artifacts such as documented procedures and technologies that act as carriers reinforcing those institutions.

In our view, it is important for such a representation to be confronted to the idea that governance is not just something that exists as a static element of organizational life crystallized in structural mechanisms, but also as a dynamic element that is implemented and should evolve in conjunction with strategy and operations, as recently suggested by authors (Benfeldt et al., 2020). This view parallels the idea that strategy exists as an object that incorporates both planned and emergent components (Mintzberg & Waters, 1990; Peppard et al., 2014). Studying the micro-level actions that compose the emergent component of strategy is crucial to understand how business and IT maintain alignment over time, consistent with works that emphasize the importance of *aligning* in practice (e.g., Marabelli & Galliers, 2017) and which have been found useful to understand the process through which firms engage with digital innovation (e.g., Chanias, Myers, & Hess, 2019; Yeow, Soh, & Hansen, 2018). We argue that the same idea could be applied to the context of data governance to move toward a view that also incorporates *governing* as a practice.

An alternative conceptual foundation inspired by this idea would be to study *ostensive* and *performative* views (Latour, 1986) of data governance to understand differences between the performance of data governance in practice (the performative view) and the idea of data governance in principle (the ostensive view). In doing so, we may find that organizations design governance mechanisms by emphasizing the design of artifacts to enforce structural and procedural governance mechanisms while not accounting for organizational routines (Feldman & Pentland, 2003) in the performance of work (Pentland & Feldman, 2008), leading to suboptimal or undesirable outcomes. At the same time, the study of relational mechanisms (e.g., mentoring, user education) could help us better understand how structural mechanisms (e.g., official policies) are effectively implemented and supported in practice.

In sum, we believe that IS research has the potential to contribute to expand our view of data governance as something that exists, toward a view that incorporates *governing data* as a process wherein data governance mechanisms are frequently adjusted to try and sustain alignment with strategy and operations. How to conceptualize this process depends of course on the ways through which we as researchers approach the governing of data. For example, understanding why some firms opt to design structural mechanisms while others lean more favorably toward the use of relational mechanisms can help us understand *how* governing data is effectively performed in practice (e.g., due to different organizational cultures or legal environments) and to identify important contextual elements that hinder the seamless transferability of data governance across contexts advocated in practitioner works, similar to other organizational elements (Becker, 2004:651).

Another important aspect of governing data that has not yet received attention in research despite its relevance for practice is the process and the impact of the implementation of data governance mechanisms on data creators and consumers. IS research has a long tradition of studying sociotechnical systems. Understanding whether and how data governance impacts the practice of organizational work could help us highlight important issues and design interventions to alleviate their undesirable consequences. For example, if the implementation of structural mechanisms (e.g., new data access policies) is met with resistance, it could foster the emergence of workarounds that support digital innovation, albeit at a cost for the organization's legal liability. While research on big data and analytics has studied the complex paths linking data-leveraging capabilities to organizational performance, the question of governance has so far remained largely absent from these studies despite its relevance (e.g., Mikalef, Boura, Lekakos, & Krogstie, 2020). We argue that exploratory, qualitative studies studying the implementation of data governance in practice across different categories of stakeholders could help thus us further unpack the link between data and organizational performance.

Finally, organizations often adopt a program approach to implementing data governance or treat data governance as a project that has a defined start and end. While experts advocate for a view of data governance where mechanisms become engrained in the practice of everyday work, to the point where they become an integral part of how the work is performed (e.g., Ladley, 2019; Seiner, 2014), the question of *how* to implement those mechanisms remains. Indeed, the task of designing and implementing data governance based on existing frameworks (e.g., DAMA International, 2017) for an organization appears daunting at best. In contrast, our conversations with workshop participants indicate that incorporating the design and the implementation of data governance mechanisms as an integral part of data-driven initiatives (e.g., analytics projects) provides advantages because costs, benefits and issues remain "connected" to those initiatives impacted by these mechanisms. In addition, it offers an opportunity to adopt an iterative approach to implementing data governance. These opportunities remain so far anecdotal to our personal discussions with practitioners however, and research has yet to help us understand how data governance is implemented in practice.[5]

4.4. Theme #4: From systems to data to services

Our final theme expands the view of data governance as an object focused on data at rest in systems toward a view of data flowing within and across organizational boundaries through digital services. One issue commonly experienced by workshop participants is that in many organizations, data governance initiatives are implemented based on existing patterns of IT governance. However, the underlying conceptualization of IT governance is based on the notion that IT is comprised of systems—physical or virtual—that are dedicated to the achievement of specific tasks. While they can evolve, these systems include relatively stable components and elements

[5] Fortunately, there are a few initiatives such as the Competence Center Corporate Data Quality (https://www.cc-cdq.ch/) that seek to bridge this gap, although they focus on the broader role of data in organizations rather than data governance in particular.

that can be used as references to design and implement mechanisms that seek to maximize the organization's ability to generate value from their use. For example, while an ERP such as SAP can be customized, its *raison d'être* is to allow firms to digitize and integrate core business processes. In the eventuality that an issue should occur (e.g., a security breach), patching will effectively resolve the issue. There will of course be damage that cannot be easily undone (e.g., if manufacturing or production operations are affected by a system outage), but the issue, once fixed, is expected to be resolved permanently.

In contrast, data have no such fixed boundaries (Lawton, 2016). Once they are obtained, they can be easily copied, altered, falsified, and used for a purpose that is vastly different from their original intent. In the previous example, the stealing of data that results from the exploitation of a vulnerability in a badly configured ERP is usually more serious than the vulnerability itself because once the data leave the confines of the system where there were generated, they can be used for a number of unplanned uses, either immediately or after a period of time that can sometimes span months or years (Hill & Swinhoe, 2021). Data cannot be "patched" or "fixed". They do not just "sit" in a database or in files. Data can move across functions, units, and organizational boundaries. Keeping data in silos to protect them may be beneficial in the short run, but it goes against the creation of value that data governance strives to achieve. Procedural mechanisms that relate to the creation and the enforcement of security policies for data access to repositories, while possible, are impractical to design and enforce in practice because data are fragmented across the organization.

To address this issue and to foster digital innovation while trying to maintain high standards of data protection, some organizations have turned toward the design of digital services in the form of application programming interfaces (APIs) (e.g., Basole, 2016) that not only facilitate data accessibility but also enforce data protection, e.g., using access-based security as well as data usage quotas. Even in the context where legacy IT infrastructure exists (e.g., in incumbent firms, as studied by Sebastian et al., 2017), it is possible to design centrally managed digital services that integrate with legacy systems and provide a uniform experience for data consumers. Doing so reduces the time required to gain access to data, increasing data quality, and provides the means to enforce and audit data protection. From the perspective of data governance, this suggests that the implementation of such services can reflect the enactment of structural mechanisms (e.g., business rules) embedded into digital boundary resources (Ghazawneh & Henfridsson, 2013) that are leveraged by users through procedural (e.g., use of a software library) and relational (e.g., best practices, online Q&A platforms) mechanisms that work in conjunction with one another. Studies on this topic, however, have so far studied this type of digital resource without engaging specifically with data governance.

We believe that there is an opportunity for IS research to mobilize perspectives such as servitization (Schüritz et al., 2017) and service orchestration (e.g., Maruping & Matook, 2020) to understand and make prescriptions on the *design* and the *evolution* of data services. For example, popular digital platforms such as GitHub provide public APIs that enable users to consume data to conduct research or to interact with their own software projects (e.g., using webhooks triggered by repository events to execute workflows such as build and test procedures on code commit). The design of such APIs requires careful planning since the platform serves more than 40 million users worldwide. Another telling example of the relationship between services and data governance is the Cambridge Analytica scandal (Zuckerman, 2018) where, according to Facebook, millions of user profiles where leaked due to a flaw in the design of a Facebook API, while others have noted that "the problematic collection of Facebook users' personal info — and the ability to obtain unusually rich info about users' friends — is due to the **design and functionality** of Facebook's Graph API." (Albright, 2018, emphasis added). Overall, we believe that studying the design and the implementation of digital services can help us better understand how data flow within an organization while drawing attention for practitioners to move away from an artifact-centric view (Tallon et al., 2013:143; 149) toward a service-centric view of data.

5. Concluding remarks

In this essay, we have used our experience preparing and delivering data governance workshops to outline key issues experienced by practitioners on the topic. At the heart of these issues is the reality that data governance mechanisms aim to achieve both value creation and data protection, two seemingly incompatible objectives accurately showcasing the benefits (e.g., Nambisan et al., 2017; Svahn, Mathiassen, & Lindgren, 2017) and the pitfalls of digital innovation (e.g., Loebbecke & Picot, 2015; Newell & Marabelli, 2015) previously pointed in IS research. Confirming observations from recent works that have bridged the gap between academia and practice on the topic of data governance (Alhassan, Sammon, & Daly, 2018; Benfeldt et al., 2020), our work further translates these practical issues into research themes that can be studied using conceptual foundations that will help bring about novel insights that have clear scientific and practical utility (Corley & Gioia, 2011). In our view, IS research is particularly well-suited to contribute to these themes because IS researchers work at the intersection of social and technical elements of organizational life, consistent with the nature of data governance mechanisms. In addition, data governance initiatives increasingly involve the implementation of technological solutions, an area of enduring interest for IS researchers that can contribute to temper some of the claims formulated by vendors who currently emphasize, perhaps too optimistically, the role of these solutions as a panacea. Overall, we hope that our work can contribute to the study of data governance as it continues to gain prominence in the digital age.

Author statement

VIAL, Gregory: Sole author of the work; responsible for all aspects of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.infoandorg.2023.100450.

References

Albright, J. (2018). The Graph API: Key points in the Facebook and Cambridge Analytica debacle. Retrieved September 15, 2020 from https://medium.com/tow-center/the-graph-api-key-points-in-the-facebook-and-cambridge-analytica-debacle-b69fe692d747.

Alhassan, I., Sammon, D., & Daly, M. (2016). Data governance activities: An analysis of the literature. *Journal of Decision Systems, 25*(sup1), 64–75.

Alhassan, I., Sammon, D., & Daly, M. (2018). Data governance activities: A comparison between scientific and practice-oriented literature. *Journal of Enterprise Information Management, 31*(2), 300–316.

Amazon.com. Amazon Macie - Amazon Web Services. Retrieved September 23, 2021 from https://aws.amazon.com/macie/.

Andriopoulos, C., & Lewis, M. W. (2009). Exploitation-exploration tensions and organizational ambidexterity: Managing paradoxes of innovation. *Organization Science, 20*(4), 696–717. https://doi.org/10.1287/orsc.1080.0406

Barrett, M., & Oborn, E. (2018). Bridging the research-practice divide: Harnessing expertise collaboration in making a wider set of contributions. *Information and Organization, 28*(1), 44–51. https://doi.org/10.1016/j.infoandorg.2018.02.006

Basole, R. C. (2016). Accelerating digital transformation: Visual insights from the API ecosystem. *IT Professional, 18*(6), 20–25. https://doi.org/10.1109/MITP.2016.105

Becker, M. C. (2004). Organizational routines: A review of the literature. *Industrial and Corporate Change, 13*(4), 643–678.

Benfeldt Nielsen, O. (2017). A comprehensive review of data governance literature. *Selected Papers of the IRIS, 8*, 120–133.

Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data governance as a collective action problem. *Information Systems Frontiers, 22*(2), 299–313.

Bourton, S., Wigley, C., & Williams, S. (2018). Embrace the uncertainty of AI. Retrieved June 11, 2021 from https://www.mckinsey.com/business-functions/organization/our-insights/the-organization-blog/embrace-the-uncertainty-of-ai.

Buytendijk, F., & Heiser, J. (2016). Privacy and ethical concerns can make big data analytics a big risk too. *Gartner*. https://www.gartner.com/en/documents/3194817.

Chanias, S., Myers, M. D., & Hess, T. (2019). Digital transformation strategy making in pre-digital organizations: The case of a financial services provider. *The Journal of Strategic Information Systems, 28*(1), 17–33.

Corley, K. G., & Gioia, D. A. (2011). Building theory about theory building: What constitutes a theoretical contribution? *Academy of Management Review, 36*(1), 12–32. https://doi.org/10.5465/amr.2009.0486

DAMA International. (2017). *DAMA-DMBOK: Data management body of knowledge* (2nd ed.). Technics Publications.

Davidson, E., Winter, J., Wessel, L., & Winter, S. (2021). Data governance, digital innovation and grand challenges. Retrieved August 3, 2021 from https://www.journals.elsevier.com/information-and-organization/call-for-papers/data-governance-digital-innovation-and-grand-challenges.

Doty, D. H., & Glick, W. H. (1994). Typologies as a unique form of theory building: Toward improved understanding and modeling. *Academy of Management Review, 19*(2), 230–251.

Dyché, J., & Polsky, A. (2016). *5 models for data stewardship. I*. SAS Institute. https://www.sas.com/content/dam/SAS/en_au/doc/whitepaper1/5%20Models%20Data%20Stewardship%20White%20Paper.pdf.

Eryurek, E., Gilad, U., Lakshmanan, V., Kibunguchy-Grant, A., & Ashdown, J. (2021). *Data governance: The definitive guide*. O'Reilly.

Fadler, M., & Legner, C. (2021). Toward big data and analytics governance: Redefining structural governance mechanisms. In *Hawaii International Conference on System Sciences*.

Feldman, M. S., & Pentland, B. T. (2003). Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly, 48*(1), 94–118.

Gaim, M., Wåhlin, N., Pina e Cunha, M., & Clegg, S. (2018). Analyzing competing demands in organizations: A systematic comparison. *Journal of Organization Design, 7*(1), 1–16. https://doi.org/10.1186/s41469-018-0030-9

Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third-party development: The boundary resources model. *Information Systems Journal, 23*(2), 173–192.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods, 16*(1), 15–31.

Google.com. Cloud data loss prevention. Retrieved March 12, 2021 from https://cloud.google.com/dlp.

Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly, 30*(3), 611–642.

Gregory, R. W., Keil, M., Muntermann, J., & Mähring, M. (2015). Paradoxes and the nature of ambidexterity in IT transformation programs. *Information Systems Research, 26*(1), 57–80.

Guillemette, M. G., & Paré, G. (2012). Toward a new theory of the contribution of the IT function in organizations. *MIS Quarterly*, 529–551.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems, 26*(3), 191–209. https://doi.org/10.1016/j.jsis.2017.07.003

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(2), 4.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75–105.

Hill, M., & Swinhoe, D. (2021). The 15 biggest data breaches of the 21st century. Retrieved October 2, 2021 from https://www.csoonline.com/article/2130877/the-biggest-data-breaches-of-the-21st-century.html.

Holt, A. (2021). *Data governance: Governing data for sustainable business*. The Chartered Institute for IT: BCS.

Inmon, B. (2016). *Data lake architecture: Designing the data lake and avoiding the garbage dump*. Technics Publications.

ISO. (2017). *ISO/IEC 38505–1:2017: Information technology — Governance of IT — Governance of data — Part 1: Application of ISO/IEC 38500 to the governance of data*. International Organization for Standardization.

Jarzabkowski, P. A., Lê, J. K., & Feldman, M. S. (2012). Toward a theory of coordinating: Creating coordinating mechanisms in practice [article]. *Organization Science, 23*(4), 907–927. https://doi.org/10.1287/orsc.1110.0693

Judah, S., & White, A. (2020). *The state of data and analytics governance is worse than you think*. Gartner.

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM, 53*(1), 148–152.

Ladley, J. (2019). *Data governance: How to design, deploy, and sustain an effective data governance program*. Academic Press.

Latour, B. (1986). The powers of association. In J. Law (Ed.), *Power, actions and belief – A new sociology of knowledge* (pp. 264–280). Routledge & Kegan Paul.

Lawton, J. (2016). Four critical success factors to effectively managing data as an asset. *The Journal of Government Financial Management, 65*(3), 34–39.

Legner, C., Eymann, T., Hess, T., Matt, C., Böhmann, T., Drews, P., Mädche, A., Urbach, N., & Ahlemann, F. (2017). Digitalization: Opportunity and challenge for the business and information systems engineering community. *Business & Information Systems Engineering, 59*(4), 301–308.

Legner, C., Pentek, T., & Otto, B. (2020). Accumulating design knowledge with reference models: Insights from 12 years' research into data management. *Journal of the Association for Information Systems, 21*(3), 735–770.

Lewis, M. W. (2000). Exploring paradox: Toward a more comprehensive guide. *Academy of Management Review, 25*(4), 760–776.

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems, 24*(3), 149–157.

Makin, S. (2021). The research-practice gap as a pragmatic knowledge boundary. *Information and Organization, 31*(2), Article 100334.

Marabelli, M., & Galliers, R. D. (2017). A reflection on information systems strategizing: The role of power and everyday practices. *Information Systems Journal, 27*(3), 347–366.

Marabelli, M., & Vaast, E. (2020). Unveiling the relevance of academic research: A practice-based view. *Information and Organization, 30*(3), Article 100314.

Maruping, L. M., & Matook, S. (2020). The evolution of software development orchestration: Current state and an agenda for future research. *European Journal of Information Systems, 29*(5), 443–457.

Mathiassen, L., & Nielsen, P. A. (2008). Engaged scholarship in IS research. *Scandinavian Journal of Information Systems, 20*(2), 1.

Microsoft Corporation. (2021). Azure Purview - unified data governance solution. Retrieved October 5, 2021 from https://azure.microsoft.com/en-ca/services/purview/.
Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2020). The role of information governance in big data analytics driven innovation. *Information & Management, 57*(7), Article 103361.
Mintzberg, H., & Waters, J. (1990). Studying deciding: An exchange of views between Mintzberg and Waters, Pettigrew, and Butler. *Organization Studies, 11*(1), 001–006.
Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital innovation management: Reinventing innovation management research in a digital world. *MIS Quarterly, 41*(1), 223–238.
Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'. *The Journal of Strategic Information Systems, 24*(1), 3–14. https://doi.org/10.1016/j.jsis.2015.02.001
Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems, 22*(3), 336–359.
Oberländer, A. M., Lösser, B., & Rau, D. (2019). Taxonomy research in information systems: A systematic assessment. In *European Conference of Information Systems, Stockholm, Sweden*.
Oberländer, A. M., Röglinger, M., & Rosemann, M. (2021). Digital opportunities for incumbents–a resource-centric perspective. *The Journal of Strategic Information Systems, 30*(3), Article 101670.
Otto, B. (2011a). A morphology of the organisation of data governance. In *European conference on information systems, Helsinki, Finland*.
Otto, B. (2011b). Organizing data governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems, 29*(1), 3.
Pentland, B. T., & Feldman, M. S. (2008). Designing routines: On the folly of designing artifacts, while hoping for patterns of action. *Information and Organization, 18*(4), 235–250.
Peppard, J., Galliers, R. D., & Thorogood, A. (2014). Information systems strategy as practice. *The Journal of Strategic Information Systems, 23*(1), 1–10.
Peterson, R. (2004). Crafting information technology governance. *Information Systems Management, 21*(4), 7–22.
Poole, M. S., & Van de Ven, A. H. (1989). Using paradox to build management and organization theories. *Academy of Management Review, 14*(4), 562–578. https://doi.org/10.5465/amr.1989.4308389
Putnam, L. L., Fairhurst, G. T., & Banghart, S. (2016). Contradictions, dialectics, and paradoxes in organizations: A constitutive approach. *Academy of Management Annals, 10*(1), 65–171. https://doi.org/10.5465/19416520.2016.1162421
Quinn, R. E., & Cameron, K. S. (1988). *Paradox and transformation: Toward a theory of change in organization and management*. Ballinger.
Schüritz, R. M., Seebacher, S., Satzger, G., & Schwarz, L. (2017). Datatization as the next frontier of Servitization-understanding the challenges for transforming organizations. In *International conference of information systems, Seoul*. South Korea.
Scott, W. R. (2008). *Institutions and organizations: Ideas and interests*. SAGE Publications.
Sebastian, I. M., Ross, J. W., Beath, C., Mocker, M., Moloney, K. G., & Fonstad, N. O. (2017). How big old companies navigate digital transformation. *MIS Quarterly Executive, 16*(3), 197–213.
Seiner, R. S. (2014). *Non-invasive data governance: The path of least resistance and greatest success*. Technics Publications.
Smith, W. K., & Lewis, M. W. (2011). Toward a theory of paradox: A dynamic equilibrium model of organizing. *Academy of Management Review, 36*(2), 381–403.
Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. *Communications of the ACM, 40*(5), 103–110.
Svahn, F., Mathiassen, L., & Lindgren, R. (2017). Embracing digital innovation in incumbent firms: How Volvo cars managed competing concerns. *MIS Quarterly, 41*(1), 239–253.
Tallon, P. P., Ramirez, R. V., & Short, J. E. (2013). The information artifact in IT governance: Toward a theory of information governance. *Journal of Management Information Systems, 30*(3), 141–178.
Terrizzano, I., Schwarz, P. M., Roth, M., & Colino, J. E. (2015). Data wrangling: The challenging journey from the wild to the lake. In *7th biennial conference on innovative data systems research, Asilomar, CA*.
The Data Governance Institute. (2021). Definitions of data governance. Retrieved July 12, 2021 from https://datagovernance.com/the-data-governance-basics/definitions-of-data-governance/.
Tyagi, P., & Demirkan, H. (2016). *Data Lakes: The biggest big data challenges*. Analytics Magazine. http://analytics-magazine.org/data-lakes-biggest-big-data-challenges/.
Van de Ven, A. H. (2007). *Engaged scholarship: A guide for organizational and social research*. Oxford University Press.
Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems, 28*(2), 118–144.
Vial, G. (2020). Data governance in the 21st-century organization. Retrieved March 2, 2021 from https://sloanreview.mit.edu/article/data-governance-in-the-21st-century-organization/.
Vial, G., Jiang, J., Giannelia, T., & Cameron, A.-F. (2021). The data problem stalling AI. *MIT Sloan Management Review, 62*(2), 47–53. https://doi.org/https://sloanreview.mit.edu/article/the-data-problem-stalling-ai/.
Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5–33.
Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all - a contingency approach to data governance. *Journal of Data and Information Quality, 1*(1), 1–27.
Weill, P., & Ross, J. (2005). A matrixed approach to designing IT governance. *MIT Sloan Management Review, 46*(2), 26–34.
Westenholz, A. (1993). Paradoxical thinking and change in the frames of reference. *Organization Studies, 14*(1), 37–58.
Winter, J. S., & Davidson, E. (2019). Big data governance of personal health information and challenges to contextual integrity. *The Information Society, 35*(1), 36–51.
Wixom, B. H., & Ross, J. W. (2017). How to monetize your data. *MIT Sloan Management Review, 58*(3).
Yeow, A., Soh, C., & Hansen, R. (2018). Aligning with new digital strategy: A dynamic capabilities approach. *The Journal of Strategic Information Systems, 27*(1), 43–58.
Zuckerman, E. (2018). *This is so much bigger than Facebook*. The Atlantic. Retrieved September 14, 2020 from https://www.theatlantic.com/technology/archive/2018/03/data-misuse-bigger-than-facebook/556310/.