Garbage in, garbage forever: why data needs a cold-chain mentality

Giulio Toscani
UPF Barcelona School of Management, Barcelona, Spain

Transforming Government: People, Process and Policy

Received 15 January 2026
Revised 23 February 2026
4 May 2026
12 May 2026
Accepted 13 May 2026

Abstract

Purpose – This paper aims to address the systemic issue of data deterioration caused by fragmented corporate data practices. It introduces the "data cold chain" as a novel regulatory framework for government-led standardization, designed to ensure data integrity and usability as information traverses organizational and sectoral boundaries.

Design/methodology/approach – This study uses a qualitative approach, drawing on semistructured interviews with 49 experts from 27 countries. These participants include government officials, corporate data officers and technology providers, ensuring a broad geographical and industrial perspective on data lifecycle management.

Findings – The research identifies four distinct features of data deterioration within corporate ecosystems and reveals how fragmented regulatory environments lead to failures in public service delivery. The findings demonstrate that internal governance is insufficient for cross-border interoperability, necessitating a shift toward binding, universal data integrity variables, such as provenance completeness and schema stability, enforced by central public authorities.

Research limitations/implications – This study shifts the locus of responsibility for data quality from individual firms to public institutions. It suggests that data usability and interoperability should be regulated with the same rigor as data privacy, framing the "data cold chain" as a necessary enabling infrastructure for digital governance.

Practical implications – This paper provides policy recommendations for regulatory bodies, such as National Data Offices, to mandate specific technical standards. It also highlights the use of public procurement rules as a nonregulatory lever to incentivize private-sector compliance and drive the adoption of harmonized data practices.

Originality/value – By applying the analogy of a "cold chain" from perishable supply chains to the digital economy, this paper offers a transformative perspective on the government's role in data governance. It moves beyond voluntary industry guidelines to propose a mandatory framework that supports both business innovation and public welfare.

Keywords Data governance, Regulatory standardization, Interoperability, Data cold chain, Public policy, Corporate data harmonization

Paper type Research paper

1. Introduction

Corporate data fragmentation poses significant challenges to economic coordination, regulatory oversight and public service delivery (Pemmasani and Abd Nasaruddin, 2022). While individual companies may optimize internal data practices, the absence of universally enforced standards creates structural inefficiencies at the ecosystem level, impeding cross-border trade, financial transparency, supply chain visibility and evidence-based policymaking (Fan *et al.*, 2014). This lack of universal standards leads to systemic data deterioration that directly impacts societal outcomes. In the public sector, this deterioration manifests as failures in service delivery, exemplified by errors in social welfare eligibility assessments or delayed public health

Transforming Government: People, Process and Policy
Emerald Publishing Limited
1750-6166
DOI 10.1108/TG-01-2026-0021

responses, often caused by incompatible data formats between government agencies and private providers (Van de Walle, 2016; Toscani, 2026a; Toscani, 2026b). While prior literature has predominantly addressed data quality through localized coding practices or internal governance (Miller and Mork, 2013; Gebru *et al.*, 2021), this study argues that the current lack of a "cold chain" for data results in a profound loss of public value. From health-care information exchanges to smart city platforms and climate monitoring systems, the inability to seamlessly integrate heterogeneous corporate data undermines both private innovation and public welfare (Kazantsev *et al.*, 2023). This paper contends that government intervention is essential to establish a "data cold chain": a standardized framework designed to preserve data integrity as information moves across organizational and sectoral boundaries. Just as food safety regulations mandate strict temperature controls in perishable supply chains to ensure consumer safety (Singh *et al.*, 2018), governments must enforce binding data standards to prevent value deterioration in digital ecosystems.

Unlike voluntary industry initiatives, which frequently suffer from fragmented adoption and competitive reluctance (Studer *et al.*, 2005), government-led standardization offers unique institutional advantages. Regulatory authorities can establish mandatory compliance mechanisms that transcend individual corporate interests, create neutral arbitration platforms for multistakeholder governance and overcome the coordination failures inherent in market-driven approaches (Xi, 2024). Furthermore, government oversight ensures public accountability for data practices that fundamentally affect citizen outcomes. The data cold chain framework thus extends beyond technical interoperability to address the complex sociotechnical dimensions of governance (Markus and Topi, 2015). It recognizes that data deterioration stems not only from incompatible formats, but from organizational silos, misaligned incentives and the absence of shared quality metrics (Toscani, 2026a, 2026b). As the institutional architect, the government is uniquely positioned to align these elements through policy, regulation and coordinated investment in public data infrastructure. Consequently, this study seeks to answer the following research question:

RQ1. How can governments establish and govern a "data cold chain" framework to prevent data deterioration and ensure integrity across fragmented corporate and sectoral boundaries?

2. Literature review

Market forces alone have proven insufficient to achieve interoperability at scale. However, innovations emerged through prolonged negotiation among industry stakeholders, often taking decades to achieve widespread adoption. In rapidly evolving digital contexts, such timescales are untenable (Bidgood *et al.*, 1997; Walravens, 2010). Government intervention accelerates standardization through several mechanisms. Mandating compliance through regulation proves more effective than relying on voluntary adoption, as demonstrated by cold chain regulation in China, where mandatory standards outperform recommendations (Zhao *et al.*, 2018). Governments provide neutral governance structures that balance competing private interests, enabling coordination that market participants cannot achieve independently (Borgogno and Colangelo, 2019). Public institutions coordinate investments in shared infrastructure that individual firms systematically underinvest in, recognizing that these investments benefit competitors as much as themselves (Bhatt *et al.*, 2017; Toscani and Prendergast, 2022). Finally, the government ensures equitable access to standardized data ecosystems, preventing monopolistic control by dominant platforms. The European Union (EU)'s General Data Protection Regulation (GDPR) exemplifies how government action can reshape entire data ecosystems, forcing thousands of organizations to harmonize practices around common principles (Calzati and van Loenen, 2023). Similarly, financial reporting standards enforced by securities regulators create transparency that markets alone would not produce (Armstrong *et al.*, 2010). These examples demonstrate that regulatory intervention, when properly designed, catalyzes coordination that voluntary mechanisms cannot achieve (Zurstrassen, 2025). Corporate data governance, when left unregulated, optimizes for firm-level efficiency rather than ecosystem-level interoperability (Shao *et al.*, 2024). Three structural barriers impede voluntary harmonization. Competitive disincentives lead companies to view proprietary data architectures as competitive advantages, actively resisting standardization that could benefit rivals (Jussen *et al.*, 2024). Coordination failures emerge even when standardization would benefit all parties, as collective action problems prevent the emergence of shared frameworks without external coordination (Skopik *et al.*, 2016). Short-term optimization pressures cause corporate decision-making to prioritize immediate operational needs over long-term ecosystem health, leading to fragmented practices that accumulate over time (Nadkarni and Prügl, 2021). These barriers are not theoretical abstractions, but practical realities observed across industries. Health-care providers develop incompatible electronic health record systems despite clear benefits from interoperability. Financial institutions maintain idiosyncratic transaction formats that complicate regulatory oversight (Xi, 2024). Supply chain participants deploy proprietary tracking systems that prevent end-to-end visibility (Hernández *et al.*, 2024). In each case, rational firm-level decisions produce collectively suboptimal outcomes that only government coordination can address (Thakur-Weigold and Miroudot, 2024). The cold chain in perishable goods management offers a compelling model for government regulation of data quality. Food safety authorities mandate continuous temperature monitoring from production to consumption, ensuring that products remain within safe thresholds throughout distribution quality (Wang *et al.*, 1995). They require mandatory documentation of handling at each transfer point, creating accountability for quality maintenance, because liability frameworks assign clear responsibility for quality failures, enabling enforcement when deterioration occurs (Reimann, 2003). Regular inspections and enforcement mechanisms ensure compliance rather than relying on voluntary adherence to best practices. Temperature serves as a singular, universally accepted metric (Singh *et al.*, 2018), a quality indicator that all stakeholders monitor using standardized instrumentation. This contrasts sharply with fragmented data quality metrics across organizations, where each entity defines quality differently, making cross-organizational assessment impossible (Wang *et al.*, 1995). The power of the cold chain model lies not in technological sophistication, but in regulatory simplicity: one metric, universally measured, consistently enforced. Applying the cold chain model to corporate data requires the government to undertake four core functions. First, define core data quality metrics analogous to temperature thresholds that all participants must monitor (Heinrich *et al.*, 2018). Second, mandate continuous monitoring as data moves across organizational boundaries, ensuring visibility into degradation points (Adepoju *et al.*, 2022). Third, establish clear liability for data degradation at each handoff point, creating accountability that drives careful stewardship (Chan *et al.*, 2019). Fourth, create enforcement mechanisms with meaningful consequences for noncompliance, ensuring that standards shape actual behavior rather than remaining aspirational guidelines (Tanzi and Pitea, 2009). The government possesses institutional capabilities that private actors lack, making public leadership essential for ecosystem-wide standardization. Regulatory authority provides the power to mandate compliance across all market participants, preventing free-riding on standardization efforts that plague voluntary initiatives. A long-term perspective enables the government to prioritize ecosystem health over quarterly performance metrics that constrain corporate decision-making, allowing investments in infrastructure that mature over decades.

Convening power grants the government the legitimacy to assemble diverse stakeholders, including competitors, to negotiate shared standards that serve collective interests (Gray *et al.*, 2022). Public accountability through democratic oversight mechanisms aligns governance with societal interests rather than shareholder returns, ensuring that standardization serves broad public welfare (Akinsola, 2025). Recent initiatives demonstrate the government's potential when it exercises these capabilities effectively. Singapore's Smart Nation platform showcases how state-led coordination can integrate data across transportation, energy and urban planning sectors, creating synergies impossible through market coordination alone (Woods *et al.*, 2023). Existing research on data governance tends to focus on either technical interoperability, examining standards development and application programming interface (API) design as engineering challenges (Wahyuni *et al.*, 2025); corporate data management, analyzing internal governance and analytics capabilities; or privacy regulation, studying GDPR compliance and data protection frameworks. This is a comprehensive framework for government-coordinated, cross-sector data standardization that addresses the sociotechnical dimensions of ecosystem-wide data integrity (Hietala, 2024). While scholars recognize the need for interoperability (Noura *et al.*, 2019), few explore the government's role as active orchestrator rather than passive regulator responding to market failures. This gap has practical consequences. Policymakers lack clear frameworks for designing interventions that balance regulatory authority with innovation incentives (Renda and Pelkmans, 2023). Corporate leaders cannot anticipate what government-led standardization might require, leading to defensive postures rather than proactive engagement (Liveris, 2023). Civil society organizations struggle to advocate for specific policy mechanisms that would enhance data quality without stifling commercial activity (Mohyeddin, 2024). This paper fills this gap by proposing the data cold chain as a policy framework for government-led standardization, grounded in empirical evidence from practitioners navigating data deterioration across organizational boundaries. By examining both the problems that necessitate intervention and the mechanisms through which government can effectively coordinate solutions, this research provides actionable guidance for policymakers while advancing theoretical understanding of data governance in complex ecosystems.

3. Methodology

We used semistructured interviews to understand how data professionals experience fragmentation across organizational boundaries and what role they see for government intervention (Eppich *et al.*, 2019). The study deliberately included both public-sector officials and private-sector data leaders to capture tensions between regulatory mandates and corporate operational realities. This dual perspective proved essential for understanding not just where standardization is needed, but where it is feasible and how resistance might emerge. The interview approach allowed participants to share detailed narratives about specific instances of data deterioration, moving beyond abstract discussions of interoperability to concrete examples of value loss (Hicks *et al.*, 2021). By grounding the research in lived experience rather than hypothetical scenarios, we captured the practical challenges that any government-led framework must address. The flexibility of semistructured interviews also enabled us to explore unexpected themes that emerged during conversations, leading to insights that structured surveys would have missed (Charmaz, 2014). To capture diverse insights and maintain proximity to unorthodox paradigms, this study used a broad geographical and industrial sample of 49 experts from 27 countries (see Table 1) across sectors including high-tech, pharmaceuticals, public sector, manufacturing and banking (Bamberger, 2018).

Table 1. Role for manuscript table

Country	Role	Sector
Afghanistan	Data Scientist	Government consulting
Bahrain	Head Data Analytics	Public telco
Brazil	Data Scientist	Marketing and advertising
China	Quality Analyst	Social media
Dominican R.	Consultant and Data Analyst	Tech consulting
Dutch	Data Scientist	Tech consulting
Finland	VP, Head of Artificial Intelligence	Finance
France	Senior Data Scientist	Retail: cosmetics
Germany	Co-Founder and VP R&D	Software services
India	Chief Data Scientist	Retail
India	Data Scientist	Consulting
India	Data Scientist	Hard disk manufacturing
Ireland	CDO	Commodities
Italy	Innovation Officer – Data Science and AI	Government
Japan	Head of AI Standardisation	Public R&D
Kazakhistan	Middle Data Scientist	Public R&D
Mexico	Director of Analytics	Banking
Mexico	Data Scientist	Technology R&D
Morocco	Professor	Public education
Moroco	CEO and Founder	Pharma
Myanmar	Data Scientist	software
Nigeria	Cloud AI and HPC	Hardware
Russia	Data Scientist	Semiconductors
Russia	Senior Data Scientist	Banking
Russia	Head of Computer-Aided Design	Public R&D
Saudi Arabia	Machine Learning Engineer	Finance
Serbia	R&D Director	Public health
Singapore	Managing Director	Innovation consulting
South Africa	Artificial Intelligence Practice Lead	Banking
Spain	Senior Data Scientist	Tech
Spain	Director of the Applied AI Unit	Public technology NGO
Spain	Director, Language Technology	Computer software
Spain	VP Artificial Intelligence	Pharma coftware
Spain	Head of Data Science and Artificial Intelligence	Software services
Spain	Global Business Development Director	Software services
Spain	Co-founder and CEO	Pharmaceutical
Spain	Head of Business Analytics	Telecommunication
Spain	Director Enterprise AI	Software services
Spain	CTO and Director of Innovation	Computer software
Spain	CEO	Tech
Spain	Data Scientist	Information exchange
Switzerland	CEO	Technology NGO
Switzerland	AI expert	Pharma
UK	Opera Composer	Art
UK	EMEA & America IT Project Manager	Recruitment services
USA	President	Health care
USA	Principal Product Data	Retail
USA	AI Services Director	Telecommunication
USA	Technical Data Scientist	Software development

Recruited via LinkedIn with a 4% response rate, participants were selected based on their explicit involvement, active engagement and primary professional responsibility in machine learning (Hicks *et al.*, 2021; Wang *et al.*, 2019B). All experts possessed at least three years of professional experience with multiorganizational data projects, ensuring depth of perspective on practical challenges that transcends theoretical understanding (Ashok *et al.*, 2017). This study was not originally designed with an exclusive focus on government, and as a result, public-sector officials represent a minority of the participants. The participants included government data officials, but mainly involved corporate data executives, including data architects, analytics directors and information governance leads, who encounter the operational consequences of standardization and interoperability requirements in practice. This sampling approach ensured diverse perspectives on current pain points in cross-organizational data exchange, revealing both universal challenges and sector-specific issues. Questions were deliberately framed to elicit concrete examples rather than abstract opinions, grounding responses in lived experience (Ren, 2022). Interviews were conducted in English, recorded with permission and transcribed using automated tools with researcher verification for accuracy (Chen, 2012). The conversational format allowed us to probe unexpected responses, following interesting threads that emerged during the discussion. Several interviews led to follow-up conversations when initial discussions revealed particularly illuminating cases, enabling deeper exploration of complex dynamics. Following grounded theory methodology, we conducted iterative coding that allowed theory to emerge from data rather than imposing predetermined frameworks (Corbin and Strauss, 2015). Analytically, these findings were derived through a multistage qualitative coding process (Locke *et al.*, 2022) from first-order to higher-order codes, which identifies the axial coding stage at which structural parallels with the cold chain model became analytically legible (Table 2).

First-order codes were generated directly from interview language, such as “incompatible assumptions,” “unclear reuse rules” and “deadline-driven shortcuts.” Line-by-line open coding identified recurring themes of data degradation and coordination failure, producing hundreds of observations later organized into higher-order categories (Charmaz, 2014). Related codes were then clustered into second-order themes reflecting shared mechanisms, while axial coding linked them to distinct deterioration processes and possible government interventions (Jean *et al.*, 2016). Although grounded theory guided coding, the data cold chain framework emerged through abductive reasoning, iterating between empirical patterns and regulatory analogies (Timmermans and Tavory, 2012). The four deterioration categories arose directly from interview data describing breakdowns, context loss, incompatible formats and progressive degradation across institutions. During selective coding, these patterns showed strong parallels with cold chain regulation in perishable goods: deterioration during transfer, prevention through continuous monitoring and limits of voluntary compliance. Integrating these insights produced four higher-order analytical categories centered on a government-orchestrated data cold chain, later validated through theoretical sampling and participant feedback. We revisited earlier interviews to confirm analytical fit and used later interviews to explicitly test the cold chain metaphor and its policy relevance. A recurring pattern showed that governance and public intervention are critical to resolving coordination failures, motivating a government-centered interpretation. Comparative analysis across sectors (e.g. health care and logistics) and contexts showed consistent underlying dynamics despite contextual variation. Findings indicate that data deterioration is not merely technical. But a policy barrier, as boundary crossings without a cold chain cause context loss and administrative friction. Overall, we identified three types of cases: cases where respondents expressed strong confidence in voluntary coordination and questioned the need for government intervention. These were primarily found among participants from

Table 2. Data structure: from first-order codes to higher-order categories

Higher-order category	Second-order theme	First-order codes (representative)	Source excerpt (paraphrased)
Fragmented standards and absent regulation	Inherited incompatible assumptions	"incompatible assumptions"; "locally optimized standards"; "no agreed deadlines or standards"	Analysts "inherit implicit assumptions embedded in data sets created elsewhere" (#34)
	Voluntary coordination failure	"working around missing standards"; "basic categorie differ across units"	Even product labels like "South" interpreted differently across business units (#49)
Underutilization due to legal ambiguity	Regulatory uncertainty blocking reuse	"privacy limits reuse"; "unclear anonymization"; "fragmented data sets"	Privacy agreements "strictly limit what data can be exposed or reused, even when analytical value is clear" (#31)
	Institutional risk aversion	"organizational boundaries prevent reuse"; "data left outdated and unmanaged"	Customers ask whether existing data can update contact info; organizational barriers prevent it (#41)
Human error as governance symptom	Assumptions replacing standards	"unclear provenance"; "unspoken assumptions"; "communication failures"	Errors stem from "misunderstanding about how certain quantities were meant to be counted" (#8)
	Bias risk from weak specification	"poorly specified practices"; "gender or skin-tone bias"	Weak standards introduce bias "that organizations are highly motivated to avoid but poorly equipped to prevent" (#33)
Time pressure and absent lifecycle regulation	Short-termism and siloed practices	"deadline-driven shortcuts"; "narrow problem definitions"; "proof-of-concept pressure"	Development teams "frequently focus on narrow, short-term objectives" under performance pressure (#35)
	Reuse sacrificed for immediacy	"rigid standards and expectations"; "little room for durable infrastructure"	Pressure to deliver proofs-of-concept "leaves little room for building durable data infrastructure" (#40)

market-liberal regulatory cultures (notably North American and some Southeast Asian contexts). We explain analytically why these accounts do not falsify the framework, but rather clarify the conditions under which voluntary mechanisms can partially succeed, and the limits of those conditions at the ecosystem scale. Cases where respondents consistently described government-led standardization efforts that had failed or created new coordination problems, for example, where mandated schemas were too rigid to accommodate legitimate domain variation. These cases led us to qualify the framework by emphasizing that the cold chain variables must be minimal and interoperability-enabling, not comprehensive or prescriptive. Finally, cases where time pressure and delivery demands were described positively, as drivers of pragmatic simplification rather than as sources of deterioration. These boundary cases prompted us to refine the claim in the discussion, distinguishing between productive time pressure that drives useful iteration and deadline-driven shortcuts that undermine data lifecycle coherence. To provide technical grounding, this study identifies four cold chain variables that must remain constant across transfers: provenance completeness, schema stability, temporal precision and integrity verification through automated checksums or cryptographic hashes. This abductive framework is both empirically grounded and theoretically generative. The cold chain is not merely a metaphor, but an analytical device revealing the regulatory gap between data protection, which governments regulate, and data integrity preservation, which they largely do not. It emerged through sustained engagement with empirical patterns, informed by adjacent regulatory domains, and later validated through participant feedback (Timmermans and Tavory, 2012).

4. Findings

The findings show that data deterioration is driven less by technical limits or individual incompetence than by the absence of uniform, enforceable regulation across the full public-sector data lifecycle. Interviews revealed that data loses consistency, meaning and policy relevance as it moves across agencies, projects and institutional boundaries. These recurring losses fall into four interrelated categories, each linked to a specific regulatory gap.

4.1 Fragmented standards and the absence of binding regulation

Interviewees consistently described upstream fragmentation in data definitions, formats and validation rules as the primary source of downstream data failure. Private-sector data architects and public-sector officials described the problem in structurally identical terms, but diverged sharply in their attribution of responsibility, corporate participants tending to externalize responsibility to regulators, public officials tending to attribute fragmentation to corporate competitive resistance. While individual data sets may be internally "clean," they frequently become unusable once integrated with data produced under different institutional assumptions. One respondent explained that, *from a research perspective, analysts are often forced to inherit implicit assumptions embedded in data sets created elsewhere* (#34). In practice, this means accepting predefined ranges, categories or units without knowing how or why they were chosen. Crucially, respondents emphasized that *poor-quality data is rarely the result of negligence*. Instead, it is *often produced by incompatible standards that are either imposed or tolerated by different public bodies*. As one interviewee working on national AI standardization explained, *their role focused almost entirely on defining core concepts for government use* (#43). The challenge was not advanced analytics, but the absence of shared foundational definitions across organizations. Several participants consistently described *the disproportionate effort required to reconcile data sets that are nominally interoperable*. One expert described *industrial vision systems that fail as soon as environmental conditions change, such as lighting or camera position* (#48). While this example came from industry, respondents repeatedly drew parallels with government data systems, where *locally optimized standards prove fragile when exposed to real-world variation*. The consequences of this fragmentation are often severe. One participant *recounted multiple projects that failed outright because teams developed models without first agreeing on deadlines, standards or integration requirements* (#34). When these teams eventually attempted to align their outputs, the results were unusable. The interview data demonstrate a consistent pattern: voluntary coordination is insufficient. Without binding regulation enforcing shared schemas, metadata standards and validation protocols, organizations optimize locally while eroding collective data integrity. A Chief Data Officer from a highly data-driven firm described *how even basic product categories, such as geographic labels like "South," were interpreted differently across units, requiring an ongoing, resource-intensive standardization effort* (#49). Viewed through a data cold-chain framework, this pattern is consistent with a system in which deterioration is not an incidental outcome, but a structurally determined one, analogous to an architecture where refrigeration is rendered functionally optional by design.

4.2 Underutilization of data driven by legal and institutional ambiguity

A second, closely related pattern concerns the systematic underuse of available public data. Interviewees consistently described *situations in which valuable data sets remain partially exploited, or not used at all, due to regulatory uncertainty, privacy concerns or institutional risk aversion*. Importantly, this underutilization is not driven by a lack of analytical capability, but by unclear and inconsistently interpreted governance frameworks. It is important here to distinguish between participants from heavily regulated sectors (health care and finance) who described legal ambiguity as a persistent operational problem, and those from less regulated sectors (logistics and retail analytics) who more often described it as manageable through internal legal counsel. One respondent working in content moderation explained that *privacy agreements strictly limit what data can be exposed or reused, even when analytical value is clear* (#31). Navigating these constraints was described as a central part of the job, yet the boundaries themselves were often ambiguous. The interview data demonstrate how respondents consistently described how different agencies interpret the same legal frameworks in divergent ways. This leads to inconsistent anonymization practices and fragmented data sets that cannot be recombined. One participant illustrated *the frustration from a service perspective: customers frequently ask whether existing data, such as e-mail content or signatures, can be used to verify or update contact information* (#41). While technically feasible and potentially valuable, organizational boundaries prevent systematic reuse, leaving large volumes of data outdated and unmanaged. As a result, analysts are often forced to work with incomplete representations of reality, even when more comprehensive data exists elsewhere in the public sector.

4.3 Human error as a consequence of weak data governance

Human error is frequently cited as a primary cause of data failure. However, interview evidence suggests it is more accurately understood as a downstream symptom of weak governance and unclear standards. Respondents consistently described *environments in which missing documentation, unclear provenance and inconsistent definitions force professionals to rely on assumptions and informal workarounds*. One interviewee emphasized that *data work requires not only technical skills, but also a deep understanding of how data is prepared, structured and constrained* (#1). At the same time, they noted that rigid predefined options often fail to reflect real-world complexity, creating a tension between formal categories and practical insight. Another participant provided a particularly revealing account of how errors emerge. They described *being handed results that appeared incorrect, only to discover that the issue was not a technical bug but a misunderstanding about how certain quantities were meant to be counted* (#8). In their experience, such misalignments stemmed from communication failures and unspoken assumptions rather than faulty algorithms. When governance structures are weak, this reliance on human judgment becomes a liability rather than a safeguard. Several respondents stressed *that assumptions silently replacing standardized rules introduce significant risks, including bias*. One participant *warned that poorly specified data practices can lead to gender or skin-tone bias, outcomes that organizations are highly motivated to avoid but poorly equipped to prevent without clear standards* (#33). These findings invite a normative inference that human judgment should be eliminated from data processes. On the contrary, expert interpretation is often essential. However, governments must regulate the conditions under which judgment operates.

4.4 Time pressure and the absence of lifecycle-oriented regulation

The final category concerns time pressure and its interaction with regulatory misalignment. Interviewees consistently *reported operating under intense temporal constraints driven by political cycles, funding requirements and crisis-oriented decision-making*. While time pressure is intrinsic to public administration, its negative effects are amplified when data standards are inconsistent or absent. These negative effects of time pressure on data quality were described most acutely by participants in public-sector and quasi-governmental roles, consistent with the political cycle dynamics described conceptually, but not previously illustrated comparatively. Several respondents described *strategies for coping with deadline-driven work, emphasizing the need for flexibility and personal time management* (#40, #47). These strategies help individuals manage stress, but do little to address systemic fragmentation. In practice, they often reinforce short-term optimization at the expense of long-term coherence. One participant *observed that development teams frequently focus on narrow, short-term objectives, especially in environments driven by performance metrics or sales pressure* (#35). Under such conditions, collaboration declines and data practices become increasingly siloed. Others explicitly *linked time pressure to reduced learning and innovation, while process improvements often aim to shorten timelines and reduce manual errors* (#15), *respondents warned that rapid delivery frequently locks in overly limited problem definitions* (#1). A participant *captured this tension by describing the pressure to deliver proof-of-concept results at the end of research projects, often under rigid standards and expectations* (#40). The need to demonstrate immediate outputs leaves little room for building a durable data infrastructure. Overall, the findings suggest that organizations frequently demand evidence-based decision-making without investing in the regulatory foundations required to sustain reliable data over time. Treating data as a short-term project resource rather than long-term public infrastructure accelerates value deterioration. Lifecycle-oriented regulation would lower the cost of reuse and enable faster responses without sacrificing reliability. In this sense, regulation does not inhibit agility, it institutionalizes it. Across all four categories, interviewee accounts converge on a single conclusion: data deterioration is fundamentally a regulatory failure. Fragmented standards, underutilized data, error-prone interpretation and time-driven shortcuts all originate from the absence of universally enforced data rules. The data cold chain framework makes visible why local competence and technical excellence cannot compensate for systemic incoherence. These findings reinforce the central argument of the paper: governments must take responsibility not only for regulating how data is used, but for regulating how data is structured, transmitted and preserved across its entire lifecycle.

5. Discussion

This study reframes data deterioration as a systemic governance failure rather than a technical or behavioral shortcoming. While prior literature has predominantly addressed data quality through coding practices, internal governance models or post hoc documentation mechanisms (Miller and Mork, 2013; Gebru *et al.*, 2021), our findings demonstrate that such approaches are insufficient when data traverses organizational and sectoral boundaries. The four categories of deterioration were present across all sectors and regulatory contexts, but the relative weight and primary mechanism differed systematically by actor group and institutional setting. The empirical evidence shows that even highly competent data professionals are constrained by fragmented regulatory environments that allow incompatible standards, inconsistent interpretations of legal requirements and uncoordinated lifecycle management. The data cold chain framework advances the literature by shifting the *locus* of responsibility from individual organizations to public institutions (Ghahramani, 2015).

Analogous to food safety regulation, where governments mandate continuous temperature monitoring regardless of handler expertise, this study argues that governments must define and enforce a limited set of universal data integrity variables that remain stable across the entire data lifecycle. Central public authorities, such as a National Data Office, a central IT authority or sector-specific regulators (e.g. health or financial monitors) would define and enforce a limited set of universal data integrity variables that remain stable across the entire data lifecycle (Toscani, 2025a). These variables are not intended to replace domain-specific standards, but to provide a minimal, binding foundation upon which interoperability becomes structurally feasible. Importantly, the findings challenge the assumption that improved communication, collaboration or professional training alone can resolve data fragmentation. While sociotechnical scholarship correctly emphasizes the role of human judgment and epistemological diversity (Markus and Topi, 2015; Pachidi *et al.*, 2014), our results indicate that human discretion becomes a source of deterioration precisely when regulatory anchors are absent. Without enforceable upstream standards, downstream collaboration merely redistributes ambiguity rather than eliminating it. The research also suggests a critical regulatory imbalance: governments have developed sophisticated regimes for data protection and privacy, yet have largely neglected data usability, interoperability and lifecycle coherence (Xi, 2024). This asymmetry produces a paradox in which data is legally protected, but analytically impoverished. The study thus extends debates on digital governance by arguing that protecting citizens' rights and preserving data value are not competing objectives, but complementary ones that require coordinated regulatory design (Toscani, 2025b). To ensure the rapid adoption of these standards, governments should look beyond traditional legislation. A powerful nonregulatory lever is the modification of public procurement rules: by mandating that the government only procures digital services from vendors who adhere to "data cold chain" standards, the public sector can use its market power to drive industry-wide compliance (Pemmasani and Abd Nasaruddin, 2022). Ultimately, the data cold chain should be understood as enabling infrastructure rather than a regulatory burden (Fan *et al.*, 2014). A lifecycle-oriented regulatory framework lowers coordination costs, accelerates decision-making under uncertainty and enables scalable evidence-based policymaking that serves both business innovation and the public good (Bidgood *et al.*, 1997; Walravens, 2010). Finally, the findings contribute to policy discourse by demonstrating that regulation does not inherently reduce agility or innovation (Kazantsev *et al.*, 2023). On the contrary, the absence of uniform standards increases time pressure, encourages short-term workarounds and undermines reuse. The data cold chain concept would also help address the data integration challenge that was identified years ago (Dong and Srivastava, 2013), as illustrated by the example of Palantir, which viewed AI largely as a layer built upon the prior necessity of data integration. Although generative AI has now reduced part of this constraint, a solution implemented at the source, like the data cold chain, would still be more logical and effective (Klein, 2026). A lifecycle-oriented regulatory framework lowers coordination costs, accelerates decision-making under uncertainty and enables scalable evidence-based policymaking. In this sense, the data cold chain should be understood as enabling infrastructure rather than a regulatory burden.

6. Limitations

This study has several limitations that should be acknowledged. First, while the research proposes the data cold chain as a regulatory framework, it does not specify a definitive set of universal data variables that governments should mandate. The identification of such variables is intentionally left open, as their precise configuration is likely to vary across sectors, risk profiles and institutional contexts. As a result, the framework is conceptual rather than prescriptive at the technical level. Second, the study relies on qualitative interviews with data professionals, which capture perceptions, experiences and institutional constraints but do not directly measure data quality outcomes before and after regulatory intervention. While this approach is appropriate for theory building and policy framing, it limits the ability to make causal claims about the quantitative impact of standardization on performance or innovation. Third, although the sample is geographically diverse, regulatory cultures differ significantly across jurisdictions. The findings may therefore reflect a stronger alignment with regulatory systems where government intervention is institutionally accepted, such as the EU or parts of Asia, compared to more market-driven governance models. This limits the immediate generalizability of implementation pathways, even if the underlying mechanisms of data deterioration are broadly applicable. Finally, the cold chain analogy, while analytically powerful, necessarily simplifies the multidimensional nature of data quality. Unlike temperature, data integrity cannot be reduced to a single scalar variable. The value of the analogy lies in regulatory logic rather than literal equivalence, and misuse of the metaphor could lead to overly rigid or technocratic interpretations if applied uncritically.

7. Future research

Future research can extend this work in several directions. First, empirical studies are needed to identify and test candidate "cold chain variables" for data, such as provenance completeness, schema stability, temporal resolution or transformation traceability. Comparative sectoral studies, across health care, finance, logistics and climate governance, would help determine which variables are truly universal and which must remain context-specific. Second, quantitative and quasi-experimental research could evaluate the effects of government-mandated data standards on interoperability, policy effectiveness and innovation outcomes. Natural experiments arising from regulatory changes, such as the introduction of mandatory reporting schemas or cross-agency data platforms, offer promising opportunities for causal analysis. Third, future work should examine institutional design questions related to enforcement, liability and compliance. How responsibility for data deterioration should be allocated across data producers, intermediaries and users remains underexplored. Insights from financial reporting regulation, food safety and environmental monitoring could inform viable enforcement models for data governance. Finally, interdisciplinary research is needed to explore the political economy of data standardization. Resistance from incumbents, concerns about surveillance and geopolitical competition over digital infrastructure all shape governments' capacity to impose uniform standards. Understanding how regulatory legitimacy is constructed, and contested, in data ecosystems will be essential for translating the data cold chain framework into practice.

References

Adepoju, A.H., Austin-Gabriel, Hamza, O. and Collins, A. (2022), "Advancing monitoring and alert systems: a proactive approach to improving reliability in complex data ecosystems", *IRE Journals*, Vol. 5 No. 11, pp. 281-282.

Akinsola, K. (2025), "Legal oversight of state-owned enterprises: governance structures, compliance, and public accountability", *Compliance, and Public Accountability* (accessed 7 February 2025).

Armstrong, C.S., Barth, M.E., Jagolinzer, A.D. and Riedl, E.J. (2010), "Market reaction to the adoption of IFRS in Europe", *The Accounting Review*, Vol. 85 No. 1, pp. 31-61.

Ashok, A., Brison, M. and LeTallec, Y. (2017), "Improving cold chain systems: challenges and solutions", *Vaccine*, Vol. 35 No. 17, pp. 2217-2223.

Bamberger, P.A. (2018), "Clarifying what we are about and where we are going", *Academy of Management Discoveries*, Vol. 4 No. 1, pp. 1-10.

Bhatt, T., Gooch, M., Dent, B. and Sylvia, G. (2017), “Implementing interoperability in the seafood industry: learning from experiences in other sectors”, *Journal of Food Science*, Vol. 82 No. S1, pp. A22-A44.

Bidgood, W.D., Jr, Horii, S.C., Prior, F.W. and Van Syckle, D.E. (1997), “Understanding and using DICOM, the data interchange standard for biomedical imaging”, *Journal of the American Medical Informatics Association*, Vol. 4 No. 3, pp. 199-212.

Borgogno, O. and Colangelo, G. (2019), “Data sharing and interoperability: fostering innovation and competition through APIs”, *Computer Law and Security Review*, Vol. 35 No. 5, p. 105314.

Calzati, S. and van Loenen, B. (2023), “A fourth way to the digital transformation: the data republic as a fair data ecosystem”, *Data and Policy*, Vol. 5, p. e21.

Chan, B.T., Veillard, J.H., Cowling, K., Klazinga, N.S., Brown, A.D. and Leatherman, S. (2019), “Stewardship of quality of care in health systems: core functions, common pitfalls, and potential solutions”, *Public Administration and Development*, Vol. 39 No. 1, pp. 34-46.

Charmaz, K. (2014), *Constructing Grounded Theory*, (2nd edition). Sage Publications Inc.

Chen, J. (2012), “Business intelligence and analytics: from big data to big impact”, *MIS Quarterly*, Vol. 36 No. 4, pp. 1165-1188.

Corbin, J., and Strauss, A. (2015), *Basics of Qualitative Research. Techniques and Procedures for Developing Grounded Theory*, (4th edition). Sage Publications Inc.

Dong, X.L., and Srivastava, D. (2013), “Big data integration”, *2013 IEEE 29th International Conference on Data Engineering (ICDE), IEEE*, pp. 1245-1248.

Eppich, W.J., Gormley, G.J., and Teunissen, P.W. (2019), “In-depth interviews”, *Healthcare Simulation Research: A Practical Guide*, Springer International Publishing, Cham, pp. 85-91.

Fan, J., Han, F. and Liu, H. (2014), “Challenges of big data analysis”, *National Science Review*, Vol. 1 No. 2, pp. 293-314.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J.W., Wallach, H., Iii, H.D. and Crawford, K. (2021), “Datasheets for datasets”, *Communications of the ACM*, Vol. 64 No. 12, pp. 86-92.

Ghahramani, Z. (2015), “Probabilistic machine learning and artificial intelligence”, *Nature*, Vol. 521 No. 7553, p. 452.

Gray, B., Purdy, J. and Ansari, S. (2022), “Confronting power asymmetries in partnerships to address grand challenges”, *Organization Theory*, Vol. 3 No. 2, p. 26317877221098765.

Heinrich, B., Hristova, D., Klier, M., Schiller, A. and Szubartowicz, M. (2018), “Requirements for data quality metrics”, *Journal of Data and Information Quality (JDIQ)*, Vol. 9 No. 2, pp. 1-32.

Hernández, C., López, M., García, J. and Vander Peterson, C. (2024), “Optimizing collaborative intelligence systems for end-to-end cybersecurity monitoring in global supply chain networks”.

Hicks, N., Millar, R.J., Girling, L.M., Cummins, P.A. and Yamashita, T. (2021), “Conducting virtual qualitative interviews with international key informants: insights from a research project”, *Grantee Submission*, Vol. 26 No. 9, pp. 2857-2871.

Hietala, H. (2024), “Collective ambidexterity in the public sector: collaborating towards innovativeness and efficiency in digital service ecosystems”.

Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B. and Ermon, S. (2016), “Combining satellite imagery and machine learning to predict poverty”, *Science*, Vol. 353 No. 6301, pp. 790-794.

Jussen, I., Möller, F., Schweihoff, J., Gieß, A., Giussani, G. and Otto, B. (2024), “Issues in inter-organizational data sharing: findings from practice and research challenges”, *Data and Knowledge Engineering*, Vol. 150, p. 102280.

Kazantsev, N., Islam, N., Zwiegelaar, J., Brown, A. and Maull, R. (2023), “Data sharing for business model innovation in platform ecosystems: from private data to public good”, *Technological Forecasting and Social Change*, Vol. 192, p. 122515.

Klein, E. (2026), (Host). "Why are Palantir and OpenAI scared of Alex Bores? [audio podcast episode]. In The Ezra Klein Show.", The New York Times, available at: www.nytimes.com/2026/04/21/opinion/ezra-klein-podcast-alex-bores.html

Liveris, A. (2023), *Leading through Disruption: A Changemaker's Guide to Twenty-First Century Leadership*, HarperCollins Leadership.

Locke, K., Feldman, M. and Golden-Biddle, K. (2022), "Coding practices and iterativity: beyond templates for analyzing qualitative data", *Organizational Research Methods*, Vol. 25 No. 2, pp. 262-284.

Markus, M.L. and Topi, H. (2015), "Big data, big decisions for science, society, and business: report on a research agenda setting workshop", ACM Technical Report.

Miller, H.G. and Mork, P. (2013), "From data to decisions: a value chain for big data", *IT Professional*, Vol. 15 No. 1, pp. 57-59.

Mohyeddin, Z. (2024), "The role of civil society in democratic governance", *Journal of Political Stability Archive*, Vol. 2 No. 3, pp. 61-74.

Nadkarni, S. and Prügl, R. (2021), "Digital transformation: a review, synthesis, and opportunities for future research", *Management Review Quarterly*, Vol. 71 No. 2, pp. 233-341.

Noura, M., Atiquzzaman, M. and Gaedke, M. (2019), "Interoperability in internet of things: taxonomies and open challenges", *Mobile Networks and Applications*, Vol. 24 No. 3, pp. 796-809.

Pachidi, S., Berends, H., Faraj, S., Huysman, M.H. and Weerd, I. (2014), "What happens when analytics lands in the organization? Studying epistemologies in clash", *Academy of Management Proceedings*, Philadelphia, PA.

Pemmasani, P.K. and Abd Nasaruddin, M.A. (2022), "Strengthening public sector data governance: risk management strategies for government organizations", *International Journal of Modern Computing*, Vol. 5 No. 1, pp. 108-118.

Reimann, M. (2003), "Liability for defective products at the beginning of the twenty-first century: emergence of a worldwide standard", *The American Journal of Comparative Law*, Vol. 51 No. 4, p. 751.

Ren, S. (2022), "Optimization of enterprise financial management and decision-making systems based on big data", *Journal of Mathematics*, Vol. 2022 No. 1, pp. 1-11.

Renda, A., and Pelkmans, J. (2023), "EU regulation: hindering or stimulating innovation?", *Handbook of Innovation and Regulation*, Edward Elgar Publishing, pp. 263-293.

Shao, X., Ahmad, M. and Javed, F. (2024), "Firm-level digitalization for sustainability performance: evidence from Ningbo City of China", *Sustainability*, Vol. 16 No. 20, p. 8881.

Singh, A.K., Subramanian, N., Pawar, K.S. and Bai, R. (2018), "Cold chain configuration design: location-allocation decision-making using coordination, value deterioration, and big data approximation", *Annals of Operations Research*, Vol. 270 Nos 1-2, pp. 433-457.

Skopik, F., Settanni, G. and Fiedler, R. (2016), "A problem shared is a problem halved: a survey on the dimensions of collective cyber defense through security information sharing", *Computers and Security*, Vol. 60, pp. 154-176.

Studer, S., Welford, R., and Hills, P. (2005), "Drivers and barriers to engaging small and medium-sized companies in voluntary environmental initiatives", *The Centre of Urban Planning and Environmental Management*, The University of Hong Kong: Hong Kong.

Tanzi, A., and Pitea, C. (2009), "Non-compliance mechanisms: lessons learned and the way forward", *Non-Compliance Procedures and Mechanisms and the Effectiveness of International Environmental Agreements*, TMC Asser Press, pp. 569-580.

Thakur-Weigold, B. and Miroudot, S. (2024), "Supply chain myths in the resilience and deglobalization narrative: consequences for policy", *Journal of International Business Policy*, Vol. 7 No. 1, pp. 99-111.

Timmermans, S. and Tavory, I. (2012), "Theory construction in qualitative research: from grounded theory to abductive analysis", *Sociological Theory*, Vol. 30 No. 3, pp. 167-186.

Toscani, G. (2025a), "From code to context: the impact of organizational problem framing on machine learning success", *Journal of Decision Systems*, Vol. 34 No. 1.

Toscani, G. (2025b), *Augmented prAIority to Enhance Human Judgement through Data and AI*, Routeldge, London.

Toscani, G. (2026a), "Future-proofing expertise: AI as a power-up, not a takeover", *Futures*, Vol. 175

Toscani, G. (2026b), *The Unknown-Known Trap: Navigating Invisible Expertise Boundaries in Machine Learning*, AI and Society.

Toscani, G. and Prendergast, G. (2022), "The role of reciprocity and reputation in service relationships with arts organisations", *Journal of Services Marketing*, Vol. 36 No. 6, pp. 800-881.

Van de Walle, S. (2016), "When public services fail: a research agenda on public service failure", *Journal of Service Management*, Vol. 27 No. 5, pp. 831-846.

Wahyuni, H., Mutiarin, D., Ismail, N.S.A. and Febriani, W. (2025), "Inter-agency data interoperability in local government for sustainability in data governance", *SSRN 5185989*.

Walravens, H. (2010), *ISBN-Internationale Standard-Buchnummer*, Simon Bibliothekswissen.

Wang, R.Y., Storey, V.C. and Firth, C.P. (1995), "A framework for analysis of data quality research", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 7 No. 4, pp. 623-640.

Wang, D., Weisz, J.D., Muller, M., Ram, P., Geyer, W., Dugan, C., Tausczik, Y., Samulowitz, H. and Gray, A. (2019), "Human-AI collaboration in data science: exploring data scientists' perceptions", *Proceedings of the ACM on Human-Computer Interaction*, Vol. 3 No. CSCW, pp. 1-24.

Woods, O., Bunnell, T. and Kong, L. (2023), "The state-led platformisation of financial services: frictionless ecosystems and an expansive logic of "smartness" in Singapore", *Geoforum*, Vol. 146, p. 103849.

Xi, W. (2024), "Regulatory changes and compliance challenges", *In Strategic Financial Management: A Managerial Approach*, Emerald Publishing Limited, pp. 119-134.

Zhao, H., Liu, S., Tian, C., Yan, G. and Wang, D. (2018), "An overview of current status of cold chain in China", *International Journal of Refrigeration*, Vol. 88, pp. 483-495.

Zurstrassen, D. (2025), "Learning from the past? EU industrial policy challenges", *Governance*, Vol. 38 No. 4, p. e70068.

Further reading

Abbasi, A., Sarker, S. and Chiang, R.H. (2016), "Big data research in information systems: toward an inclusive research agenda", *Journal of the Association for Information Systems*, Vol. 17 No. 2, p. 3.

Ahsen, M.E., Ayvaci, M.U.S. and Raghunathan, S. (2019), "When algorithmic predictions use human-generated data: a bias-aware classification algorithm for breast cancer diagnosis", *Information Systems Research*, Vol. 30 No. 1, pp. 97-116.

Al-Dmour, H., Saad, N., Basheer Amin, E., Al-Dmour, R. and Al-Dmour, A. (2023), "The influence of the practices of big data analytics applications on bank performance: field study", *VINE Journal of Information and Knowledge Management Systems*, Vol. 53 No. 1, pp. 119-141.

Allen, R.T. and Choudhury, P. (2022), "Algorithm-augmented performance and domain experience", *Organization Science*, Vol. 33, pp. 149-169.

Archibald, M.M., Ambagtsheer, R.C., Casey, M.G. and Lawless, M. (2019), "Using Zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants", *International Journal of Qualitative Methods*, Vol. 18, p. 1609406919874596.

Badia-Melis, R., Mc Carthy, U., Ruiz-Garcia, L., Garcia-Hierro, J. and Villalba, J.R. (2018), “New trends in cold chain monitoring applications – a review”, *Food Control*, Vol. 86, pp. 170-182.

Belcore, E., Angeli, S., Colucci, E., Musci, M.A. and Aicardi, I. (2021), “Precision agriculture workflow, from data collection to data management using FOSS tools: an application in northern Italy vineyard”, *ISPRS International Journal of Geo-Information*, Vol. 10 No. 4, p. 236.

Bindseil, U., Panetta, F. and Terol, I. (2021), “Central bank digital currency: functional scope, pricing and controls”, ECB Occasional Paper, (2021/286).

Cai, Y. and Lattu, A. (2022), “Triple helix or quadruple helix: which model of innovation to choose for empirical studies?”, *Minerva*, Vol. 60 No. 2, pp. 257-280.

Cheng, Y., Li, D., Guo, Z., Jiang, B., Geng, J., Bai, W., … Xiong, Y. (2020), “Accelerating end-to-end deep learning workflow with codesign of data preprocessing and scheduling”, *IEEE Transactions on Parallel and Distributed Systems*, Vol. 32 No. 7, pp. 1802-1814.

Cock, P.J., Fields, C.J., Goto, N., Heuer, M.L. and Rice, P.M. (2010), “The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants”, *Nucleic Acids Research*, Vol. 38 No. 6, pp. 1767-1771.

Collins, C., Dennehy, D., Conboy, K. and Mikalef, P. (2021), “Artificial intelligence in information systems research: a systematic literature review and research agenda”, *International Journal of Information Management*, Vol. 60, p. 102383.

Davenport, T.H. and Kirby, J. (2016), *Only Humans Need Apply: Winners and Losers in the Age of Smart Machines*, Harper Collins, New York, NY.

Davenport, T.H. and Ronanki, R. (2018), “Artificial intelligence for the real world”, *Harvard Business Review*.

Draucker, C.B., Martsolf, D.S., Ross, R. and Rusk, T.B. (2007), “Theoretical sampling and category development in grounded theory”, *Qualitative Health Research*, Vol. 17 No. 8, pp. 1137-1148.

Englander, H., Jones, A., Krawczyk, N., Patten, A., Roberts, T., Korthuis, P.T. and McNeely, J. (2022), “A taxonomy of hospital-based addiction care models: a scoping review and key informant interviews”, *Journal of General Internal Medicine*, Vol. 37 No. 11, pp. 2821-2833.

Ennen, E. and Richter, A. (2010), “The whole is more than the sum of its parts—or is it? A review of the empirical literature on complementarities in organizations”, *Journal of Management*, Vol. 36 No. 1, pp. 207-233.

Erevelles, S., Fukawa, N. and Swayne, L. (2016), “Big data consumer analytics and the transformation of marketing”, *Journal of Business Research*, Vol. 69 No. 2, pp. 897-904.

Farrukh, C. and Holgado, M. (2020), “Integrating sustainable value thinking into technology forecasting: a configurable toolset for early stage technology assessment”, *Technological Forecasting and Social Change*, Vol. 158, p. 120171.

Fielding, R., Nottingham, M. and Reschke, J. (Eds), (2022), “RFC 9110: HTTP Semantics”.

Fu, S., Liu, J., Chu, X. and Hu, Y. (2016), “Toward a standard interface for cloud providers: the container as the narrow waist”, *IEEE Internet Computing*, Vol. 20 No. 2, pp. 66-71.

Fu, J., Zhang, Y., Wang, Y., Zhang, H., Liu, J., Tang, J., … Zhu, F. (2022), “Optimization of metabolomic data processing using NOREVA”, *Nature Protocols*, Vol. 17 No. 1, pp. 129-151.

Gilli, A. and Lanata. (2020), *Strategies for Leading in the Age of Artificial Intelligence*, NATO Defense College.

Greve, H.R. (2003), “A behavioral theory of R&D expenditures and innovations: evidence from shipbuilding”, *Academy of Management Journal*, Vol. 46 No. 6, pp. 685-702.

Günther, W.A., Mehrizi, M.H.R., Huysman, M. and Feldberg, F. (2017), “Debating big data: a literature review on realizing value from big data”, *The Journal of Strategic Information Systems*, Vol. 26 No. 3, pp. 191-209.

Gupta, B.B., Gaurav, A. and Panigrahi, P.K. (2023), "Analysis of security and privacy issues of information management of big data in B2B based healthcare systems", *Journal of Business Research*, Vol. 162, p. 113859.

Hill, K. (2012), *How Target Figured Out a Teen Girl was Pregnant before Her Father Did*, Forbes, Inc, Vol. 7, pp. 4-1.

Huang, L. (2018), "The role of investor gut feel in managing complexity and extreme risk", *Academy of Management Journal*, Vol. 61 No. 5, pp. 1821-1847.

Iansiti, M. and Lakhani, K.R. (2020), *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*, (Harvard Business School Press, Boston).

James, S.J. and James, C.J.F.R.I. (2010), "The food cold-chain and climate change", *Food Research International*, Vol. 43 No. 7, pp. 1944-1956.

Kahneman, D. and Tversky, A. (2000), *Choices, Values, and Frames*, Cambridge University Press.

Kahneman, D. and Tversky, A. (2013), "Prospect theory: an analysis of decision under risk", *Handbook of the Fundamentals of Financial Decision Making: Part I*, pp. 99-127.

Kane, G.C., Johnson, J. and Majchrzak, A. (2014), "Emergent life cycle: the tension between knowledge change and knowledge retention in open online coproduction communities", *Management Science*, Vol. 60 No. 12, pp. 3026-3048.

Kaplan, E.D. and Hegarty, C. (Eds) (2017), "Understanding GPS/GNSS: principles and applications", *Artech House*.

Keerthana, G. (2025), "A survey on big data classification", *Data and Knowledge Engineering*, Vol. 156, p. 102408.

Kitchin, R. and Moore-Cherry, N. (2021), "Fragmented governance, the urban data ecosystem and smart city-regions: the case of metropolitan Boston", *Regional Studies*, Vol. 55 No. 12, pp. 1913-1923.

Kush, R.D., Warzel, D., Kush, M.A., Sherman, A., Navarro, E.A., Fitzmartin, R., … Hudson, L. (2020), "FAIR data sharing: the roles of common data elements and harmonization", *Journal of Biomedical Informatics*, Vol. 107, p. 103421.

Laureiro-Martínez, D. and Brusoni, S. (2018), "Cognitive flexibility and adaptive decision-making: evidence from a laboratory study of expert decision makers", *Strategic Management Journal*, Vol. 39 No. 4, pp. 1031-1058.

Lebovitz, S., Lifshitz-Assaf, H. and Levina, N. (2022), "To engage or not to engage with DATA for critical judgments: how professionals deal with opacity when using DATA for medical diagnosis", *Organization Science*, Vol. 33 No. 1, pp. 126-148.

Leiner, B.M., Cerf, V.G., Clark, D.D., Kahn, R.E., Kleinrock, L., Lynch, D.C., … Wolff, S. (2009), "A brief history of the internet", *ACM SIGCOMM Computer Communication Review*, Vol. 39 No. 5, pp. 22-31.

Levine, J. (1999), "Introduction to time and frequency metrology", *Review of Scientific Instruments*, Vol. 70 No. 6, pp. 2567-2596.

Likar, K. and Jevšnik, M. (2006), "Cold chain maintaining in food trade", *Food Control*, Vol. 17 No. 2, pp. 108-113.

Loy, G. (1985), "Musicians make a standard: the MIDI phenomenon", *Computer Music Journal*, Vol. 9 No. 4, pp. 8-26.

Maharana, K., Mondal, S. and Nemade, B. (2022), "A review: data pre-processing and data augmentation techniques", *Global Transitions Proceedings*, Vol. 3 No. 1, pp. 91-99.

McKinsey & Company (2017), "A CEO guide for avoiding the ten traps that derail digital transformations", Digital McKinsey.

Mercier, S., Villeneuve, S., Mondor, M. and Uysal, I. (2017), "Time–temperature management along the food cold chain: a review of recent developments", *Comprehensive Reviews in Food Science and Food Safety*, Vol. 16 No. 4, pp. 647-667.

Merriam-Webster (2022), "Accessed on 14th October 2022", available at: www.merriam-webster.com/dictionary/skill

Mittal, D., Mease, R., Kuner, T., Flor, H., Kuner, R. and Andoh, J. (2023), "Data management strategy for a collaborative research center", *GigaScience*, Vol. 12, p. giad049.

Murray, A., Rhymer, J. and Sirmon, D.G. (2020), "Humans and technology: forms of conjoined agency in organizations", *Academy of Management Review*, Vol. 46 No. 3, pp. 551-572.

Naeem, M., Jamal, T., Diaz-Martinez, J., Butt, S.A., Montesano, N., Tariq, M.I., … De-La-Hoz-Valdiris, E. (2022), "Trends and future perspective challenges in big data", *Advances in Intelligent Data Analysis and Applications*, Springer Singapore, pp. 309-325.

Namvar, M. and Cybulski, J. (2014), *BI-based organizations: a sensemaking perspective. In Proceedings of the Thirty-Fifth International Conference on Information Systems, Auckland, New Zealand, December 14–17*.

Obermeyer, Z. and Lee, T.H. (2017), "Lost in thought–the limits of the human mind and the future of medicine", *New England Journal of Medicine*, Vol. 377 No. 13, pp. 1209-1211.

Park, J.S., Jang, J. and Lee, E. (2018), "Theoretical and empirical studies on essence-based adaptive software engineering", *Information Technology and Management*, Vol. 19 No. 1, pp. 37-49.

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.F., Breazeal, C., … Wellman, M. (2019), "Machine behaviour", *Nature*, Vol. 568 No. 7753, pp. 477-486.

Remadi, A., El Hage, K., Hobeika, Y. and Bugiotti, F. (2024), "To prompt or not to prompt: navigating the use of large language models for integrating and modeling heterogeneous data", *Data and Knowledge Engineering*, Vol. 152, p. 102313.

Sayogo, D.S., Zhang, J., Luna-Reyes, L., Jarman, H., Tayi, G., Andersen, D.L., … Andersen, D.F. (2015), "Challenges and requirements for developing data architecture supporting integration of sustainable supply chains", *Information Technology and Management*, Vol. 16 No. 1, pp. 5-18.

Schweinsberg, M., Feldman, M., Staub, N., van den Akker, O.R., van Aert, R.C., Van Assen, M.A., … Schulte-Mecklenbeck, M. (2021), "Same data, different conclusions: radical dispersion in empirical results when independent analysts operationalize and test the same hypothesis", *Organizational Behavior and Human Decision Processes*, Vol. 165, pp. 228-249.

Sharma, G. and Bansal, P. (2020), "Cocreating rigorous and relevant knowledge", *Academy of Management Journal*, Vol. 63 No. 2, p. 386-410.

Snihur, Y. and Zott, C. (2020), "The genesis and metamorphosis of novelty imprints: how business model innovation emerges in young ventures", *Academy of Management Journal*, Vol. 63 No. 2, pp. 554-583.

Solarino, A.M. and Aguinis, H. (2021), "Challenges and best-practice recommendations for designing and conducting interviews with elite informants", *Journal of Management Studies*, Vol. 58 No. 3, pp. 649-646.

Spradley, J. (1979), "Asking descriptive questions", *The Ethnographic Interview*, Vol. 1, pp. 44-61.

Susskind, R. and Susskind, D. (2022), *The Future of the Professions: How Technology will Transform the Work of Human Experts*, Oxford University Press.

Thornberg, R. and Charmaz, K. (2014), "Grounded theory and theoretical coding", *The SAGE Handbook of Qualitative Data Analysis*, Vol. 5 No. 2014, pp. 153-169.

Toscani, G. and Yamaguchi, Y. (2026), NAVOZYME: EL MAR, LAS CADENAS Y UN BARCO SIN ANCLA", *Harvard Deusto Business Review*, Vol. 366.

Urquhart, C. and Fernández, W. (2013), "Using grounded theory method in information systems: the researcher as blank slate and other myths", *Journal of Information Technology*, Vol. 28 No. 3, pp. 224-236.

Van Laar, E., Van Deursen, A.J., Van Dijk, J.A. and De Haan, J. (2017), "The relation between 21st-century skills and digital skills: a systematic literature review", *Computers in Human Behavior*, Vol. 72, pp. 577-588.

Von Krogh, G. (2018), "Artificial intelligence in organizations: new opportunities for phenomenon-based theorizing", *Academy of Management Discoveries*, Vol. 4 No. 4.

Whang, S.E., Roh, Y., Song, H. and Lee, J.G. (2023), "Data collection and quality challenges in deep learning: a data-centric AI perspective", *The VLDB Journal*, Vol. 32 No. 4, pp. 791-813.

Witten, I.H., Bainbridge, D. and Nichols, D.M. (2009), *How to Build a Digital Library*, Morgan Kaufmann.

Corresponding author
Giulio Toscani can be contacted at: giulio.toscani@bsm.upf.edu