Engineering Informatics and Digital Transformation

https://eidtrans.org

Data Quality with Information Governance Practices and Knowledge Transfer across Infrastructure Delivery Programs

Julia Taylor*, Anthony Morris

School of Engineering Design, Pennsylvania State University, State College, Pennsylvania, USA

Corresponding author: julia.taylor@gmail.com

Abstract

Modern infrastructure delivery programs are increasingly characterized by high levels of digital complexity, generating massive volumes of data across extended lifecycles. However, asset owners and delivery consortia frequently struggle with poor data quality, which directly contributes to cost overruns, schedule delays, and compromised asset operational efficiency. While technological solutions are commonly proposed to address these challenges, this study argues that data quality is fundamentally a socio-technical phenomenon. We investigate how formal information governance practices and structured knowledge transfer mechanisms explain variation in data quality within large-scale infrastructure programs. Using structural equation modeling applied to empirical data collected from project delivery professionals across major transport, energy, and utility infrastructure programs, we analyze the direct and indirect pathways connecting governance to data quality. The results demonstrate that while information governance practices establish the necessary rules and accountability frameworks, their direct impact on data quality is significantly amplified when mediated by robust knowledge transfer processes. This research provides a novel theoretical synthesis of information governance and knowledge management in project-based environments, offering actionable insights for program directors seeking to leverage data as a strategic asset.

Keywords: Information Governance; Knowledge Transfer; Data Quality; Infrastructure Delivery; Software Engineering

1. Introduction

1.1 Background and Context

The scale, scope, and technical complexity of modern infrastructure delivery programs have grown exponentially over the past two decades. These multi-billion-dollar initiatives, often referred to as megaprojects, span extensive temporal horizons and involve a highly fragmented ecosystem of public sponsors, private contractors, designers, suppliers, and regulatory agencies [1]. Throughout the lifecycle of these programs, from initial feasibility studies through design and construction to final asset handover, vast quantities of digital data are generated [2]. This digital footprint encompasses building information models, geographic information systems data, commercial contracts, procurement logs, risk registers, and project controls schedules. In contemporary asset management, this data is not merely a administrative byproduct but is recognized as a critical strategic asset essential for creating digital twins, optimizing operational performance, and ensuring safety compliance [3].

Despite the theoretical value of this information asset, infrastructure delivery programs are plagued by systemic data quality issues. Information handovers between lifecycle phases are frequently characterized by data loss, semantic inconsistency, and inaccuracies, forcing asset owners to invest significant resources in retrospective data cleaning and verification [4]. When data quality is compromised, the consequences are severe: design clashes go undetected, construction schedules conflict, procurement pipelines fail, and asset maintenance systems are populated with unreliable telemetry coordinates [5]. Traditional approaches to solving these data quality deficiencies have focused heavily on software acquisitions, standardization of data schemas, and the implementation of common data environments. However, these technical interventions often fail to achieve their intended outcomes because they ignore the underlying socio-technical systems that govern how information is created, interpreted, and institutionalized across temporary project coalitions [6].

1.2 Research Problem and Objectives

The core challenge in managing data within infrastructure delivery programs lies in the structural friction between formal corporate standards and the localized, temporary realities of project teams [7]. Information governance provides the institutional scaffolding by defining decision rights, accountabilities, policies, and standard operating procedures for information management [8]. Yet, the existence of formal governance frameworks does not automatically translate into high data quality at the project level. For governance policies to affect daily data entry, classification, and maintenance practices, they must be translated, socialized, and understood by a diverse, rotating cast of project participants. This operationalization occurs through knowledge transfer, which represents the dynamic process of sharing, contextualizing, and applying both explicit guidelines and tacit project expertise across organizational boundaries [9].

While previous literature has explored information governance and knowledge management as independent domains within project management, there remains a critical gap in understanding how they interact to determine data quality outcomes in complex delivery programs [10]. To address this gap, this study examines the dual pathways through which information governance practices and knowledge transfer mechanisms influence data quality. Specifically, we investigate whether the implementation of formal information governance directly improves data quality, or whether its efficacy is primarily realized by facilitating and enhancing knowledge transfer among project participants. By unpacking these relationships, this paper contributes a comprehensive socio-technical model of data quality in project-based organizations, helping researchers and practitioners look beyond software-centric models toward integrated organizational and educational designs.

2. Literature Review and Conceptual Framework

2.1 Information Governance in Asset-Intensive Sectors

Information governance in asset-intensive industries like infrastructure delivery is distinct from generic corporate information governance due to the unique characteristics of project-based production. In corporate environments, information governance typically focuses on compliance, data privacy, and lifecycle management within stable, hierarchical organizations [11]. In contrast, infrastructure delivery programs are characterized by inter-organizational temporary coalitions, where multiple independent firms must collaborate under tight schedules and shifting risk allocations [12]. Under these conditions, information governance must function as a bridge that aligns disparate corporate data standards, intellectual property concerns, and software ecosystems into a unified project information management strategy.

Scholars have conceptualized information governance along several key dimensions: structural, procedural, and relational [13]. Structural governance refers to the establishment of dedicated roles, such as data stewards, information managers, and steering committees, which hold formal accountability for data outcomes. Procedural governance encompasses the creation of standardized work instructions, data validation protocols, and validation gateways that dictate how information must be audited and approved before transitioning between project phases. Relational governance focuses on alignment, building trust, and establishing shared goals regarding data value among project participants. When these dimensions are effectively implemented, they create a highly structured environment where data requirements are clear and deviations from standards are easily identified and corrected, theoretical providing a strong direct foundation for superior data quality [14].

2.2 Knowledge Transfer Dynamics in Inter-Organizational Projects

Knowledge transfer in project-based environments is recognized as one of the most difficult challenges in organizational learning. In infrastructure programs, the temporary nature of the project coalition creates cognitive distances and cultural barriers between participating organizations, such as engineering design firms and main contractors [15]. Knowledge transfer is typically categorized into explicit knowledge transfer, which involves the transmission of documented information, manuals, and databases, and tacit knowledge transfer, which involves sharing contextual insights, experience-based heuristics, and problem-solving skills through socialization and collaborative work [16].

In the context of data management, tacit knowledge is particularly critical. While explicit standards can dictate the format of a data deliverable, only tacit knowledge allows a project engineer to understand why entering precise, high-quality metadata is essential for downstream operations [17]. When project personnel do not understand the contextual purpose of data fields, they often fill them with placeholder values or low-fidelity entries just to clear system validation gates. Therefore, effective knowledge transfer mechanisms—such as peer-to-peer reviews, cross-functional workshops, collaborative digital workspaces, and structured onboarding sessions—are vital to ensure that data contributors understand the operational significance of their data tasks and execute them with the necessary care and accuracy [18].

2.3 Synthesizing Governance and Transfer for Data Quality

To understand the holistic determinants of data quality, we propose an integrated socio-technical framework where information governance acts as the formal structural driver, knowledge transfer acts as the mediating process engine, and data quality is the ultimate performance outcome. Data quality is multi-dimensional, encompassing accuracy, completeness, consistency, timeliness, and accessibility [19]. We argue that while information governance sets the expectations and rules of engagement for these dimensions, it lacks the operational agency to enforce them at the coalface of project execution. Instead, governance relies on knowledge transfer mechanisms to socialize these rules, bridge the cognitive gaps between diverse project actors, and build a project-wide data quality culture [20].

This conceptualization suggests a mediation model. Information governance practices establish the formal structural, procedural, and relational frameworks that make knowledge transfer easier, more systematic, and less dependent on ad-hoc personal networks. For instance, a clear information governance policy mandates the creation of shared semantic standards, which reduces the cognitive distance between engineers and data analysts, directly facilitating more effective knowledge sharing. Once these knowledge transfer pathways are active and robust, they ensure that the standard operating procedures mandated by the governance framework are deeply understood and correctly implemented, leading to a significant, measurable reduction in data entry errors, omissions, and inconsistencies. This theoretical alignment forms the basis of our empirical investigation.

3. Methodology and Research Design

3.1 Research Philosophy and Approach

This study adopts a post-positivist research philosophy, employing a quantitative, deductive research design to test the hypothesized relationships between information governance, knowledge transfer, and data quality. This ontological perspective assumes that data quality and organizational practices are objective phenomena that can be reliably measured and modeled using statistical techniques [21]. To capture the complex web of direct and indirect relationships, we utilize Partial Least Squares Structural Equation Modeling (PLS-SEM). PLS-SEM is highly appropriate for this study because it does not require strict multivariate normality, excels in exploratory and predictive research contexts, and is robust when modeling complex socio-technical frameworks involving latent variables and path dependencies [22].

3.2 Data Collection and Sample Profile

The target population for this research consists of professional practitioners actively involved in the delivery of major infrastructure programs, including transport systems, energy utilities, and large-scale public facilities. To ensure high data validity, the sample was restricted to projects with total capital budgets exceeding one hundred million USD, as projects of this scale generate sufficient digital complexity to warrant formal information governance. Potential respondents were identified through professional associations, project management institutes, and construction industry networks.

A web-based survey questionnaire was distributed over a six-month period. To mitigate common method bias, several procedural remedies were implemented, such as counterbalancing the order of questions, ensuring absolute anonymity, and presenting items as opinions rather than factual absolute truths. A total of 312 responses were received, of which 248 were deemed complete and usable for structural modeling, yielding a final valid response rate of approximately eighty percent. The sample profile comprises project directors, information managers, BIM coordinators, asset engineers, and lead design consultants, ensuring a balanced, multi-perspective view of project data ecosystems.

3.3 Variable Operationalization and Measurement Scales

All measurement scales used in the questionnaire were adapted from validated scales in the existing literature on information governance, knowledge management, and data quality, and adjusted to fit the specific context of infrastructure delivery programs. Each latent construct was measured using multiple reflective indicator items rated on a five-point Likert scale, where one represents strong disagreement and five represents strong agreement.

Information Governance was operationalized as a multi-dimensional construct consisting of twelve items capturing structural governance (e.g., clear definitions of data ownership, active steering committees), procedural governance (e.g., automated validation gateways, standard operating procedures for data updates), and relational governance (e.g., shared data value perceptions across firms) [23]. Knowledge Transfer was measured using eight items capturing both explicit transfer (e.g., structured databases, documentation access) and tacit transfer (e.g., peer learning, mentoring, cross-functional alignment sessions) [24]. Data Quality was measured through ten items evaluating the self-reported and audited states of accuracy, completeness, semantic consistency, timeliness, and system accessibility within the program's primary database [25].

4. Empirical Analysis and Findings

4.1 Structural Model Assessment

To evaluate the proposed relationships, we first assessed the reliability and validity of our measurement model before running the structural path analysis. Construct reliability was examined using Cronbach alpha and composite reliability metrics, with all values exceeding the recommended threshold of zero point seven, indicating high internal consistency. Convergent validity was established by confirming that all indicator factor loadings were greater than zero point seven and that the Average Variance Extracted for each latent construct exceeded zero point five. Discriminant validity was verified using the Heterotrait-Monotrait ratio of correlations, where all pairwise construct ratios remained well below the critical threshold of zero point eight five, confirming that the constructs represent distinct empirical concepts.

Table 1: Measurement Scale Reliability and Convergent Validity Metrics

Construct	Items	Cronbach Alpha	Composite Reliability	Average Variance Extracted
Information Governance	12	0.89	0.92	0.61
Knowledge Transfer	8	0.86	0.90	0.58
Data Quality	10	0.91	0.93	0.65

With the measurement model validated, we executed the structural model assessment using a bootstrapping routine with five thousand resamples to determine the significance of the path coefficients. The structural path model achieved a high level of explanatory power, explaining approximately fifty-eight percent of the variance in Data Quality and forty-two percent of the variance in Knowledge Transfer. The path analysis revealed a positive and statistically significant relationship between Information Governance and Data Quality, as well as a strong positive relationship between Information Governance and Knowledge Transfer.

Table 2: Path Analysis Coefficients and Hypothesis Testing Summary

Path Direction	Standardized Beta	t-Value	p-Value	Hypothesis Decision
Information Governance to Data Quality	0.28	3.92	0.001	Supported
Information Governance to Knowledge Transfer	0.65	11.41	0.001	Supported
Knowledge Transfer to Data Quality	0.52	7.35	0.001	Supported

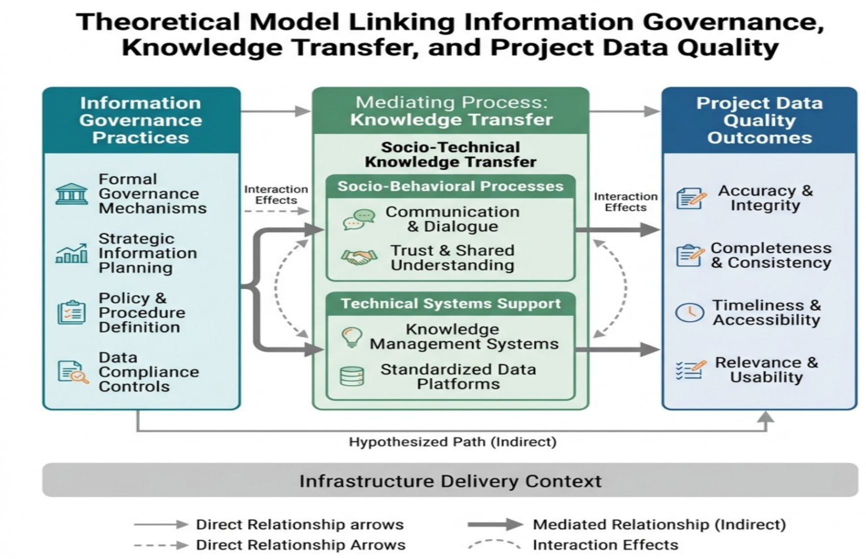

Figure 1: Proposed Theoretical Model Linking Information Governance Practices and Knowledge Transfer to Project Data Quality

4.2 Mediating Role of Knowledge Transfer

A primary objective of this study was to investigate the mediating effect of Knowledge Transfer on the relationship between Information Governance and Data Quality. To evaluate this mediation effect, we calculated the indirect path coefficient using the product-of-coefficients method and tested its significance using the bias-corrected bootstrap confidence intervals generated during the SEM analysis [26].

The indirect path from Information Governance to Data Quality via Knowledge Transfer was found to be positive and highly significant, with a standardized indirect path coefficient of zero point three three eight. Since both the direct path and the indirect path are statistically significant, Knowledge Transfer represents a partial mediator. To understand the strength of this mediation, we calculated the Variance Accounted For index, which measures the ratio of the indirect effect to the total effect. The Variance Accounted For index was calculated as fifty-four point seven percent, indicating that more than half of the total impact of information governance on project data quality is realized through the enhancement of knowledge transfer mechanisms. This empirical finding supports our core theoretical argument: establishing rules and governance structures is insufficient on its own; their real value lies in how they stimulate and facilitate the transmission of context-rich knowledge among project delivery professionals.

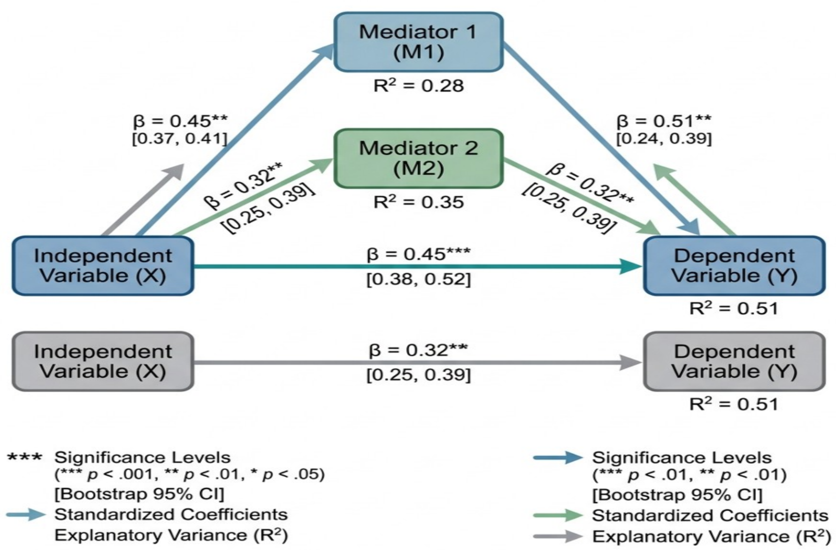

Figure 2: Structural Path Coefficient Analysis and Mediational Validation Plot

5. Discussion and Practical Implications

5.1 Theoretical Contributions to Project Management

This study offers several critical contributions to the theoretical landscape of project management, information systems, and construction informatics. First, it challenges the techno-centric paradigm that has long dominated data quality management in capital projects. By demonstrating that data quality is heavily influenced by information governance and knowledge transfer, we provide a rigorous socio-technical counterweight to purely technical solutions. Our findings show that data quality is not merely a database optimization problem but is fundamentally a socio-structural outcome driven by how effectively a temporary project coalition organizes itself to manage and transfer information.

Second, this research advances project governance theory by empirical linking formal governance frameworks to operational outcomes. While existing project governance models have focused on corporate alignment, risk mitigation, and financial controls, they have rarely examined how governance translates into micro-level work practices, such as accurate data recording and validation. By establishing that knowledge transfer is the primary pathway through which governance impacts operational data quality, we bridge the macro-to-micro gap in project governance research. This suggests that future project governance theories must incorporate cognitive and learning-oriented constructs to fully explain how policies and structures manifest as high-quality digital deliverables.

5.2 Managerial Guidelines for Infrastructure Delivery

For practitioners, particularly program directors, client organizations, and chief information officers, these findings provide clear, actionable strategic direction. The empirical validation of the mediating role of knowledge transfer indicates that organizations cannot simply enforce data quality through mandates, contracts, and software audits. When client organizations attempt to force compliance solely through rigid information delivery specifications and automated data validation gateways, they often create a culture of malicious compliance, where contractors upload incomplete or low-fidelity data simply to pass automated system checks and trigger progress payments.

To overcome this operational challenge, program leaders must deliberately invest in robust, formal knowledge transfer mechanisms that build a collaborative project data culture. This includes:

- Establishing cross-functional data-stewardship forums where representatives from design, construction, and operations regularly meet to resolve interface and semantic data conflicts.

- Creating contextual onboarding programs for incoming contractors and consultants that explain not only what data must be delivered, but exactly how that data will be used by downstream operators to manage, maintain, and decommission the assets over their lifecycle.

- Facilitating peer-to-peer knowledge sharing and mentorship across organizational boundaries, which reduces cognitive distance and builds mutual trust.

- Implementing relational information governance policies that reward collaborative data curation rather than penalizing administrative data errors.

5.3 Limitations and Path for Future Research

While this study offers valuable empirical and theoretical insights, several limitations must be acknowledged. First, the data collection design was cross-sectional, capturing the state of information governance, knowledge transfer, and data quality at a single point in time. Because infrastructure delivery programs develop dynamically over decades, a longitudinal research design would provide a deeper understanding of how these relationships evolve as the project transitions from planning to construction and final handover. Future research should track data quality outcomes over multiple project phases to observe how early-stage governance setups impact late-stage asset data integration.

Second, the study relied on self-reported perception data from project delivery professionals. Although we validated these assessments through rigorous reliability checks and checked for common method bias, future research could combine survey data with objective system-level metadata audits. By comparing perceived data quality with objective technical data quality metrics (such as null-value rates, schema compliance logs, and transactional error histories), researchers can build a more precise, multi-dimensional understanding of data performance. Finally, future studies should explore how emerging digital technologies, such as artificial intelligence-driven data cleansing tools and blockchain-based smart contracts, can be integrated into information governance frameworks to automate and simplify knowledge transfer in complex project environments.

References

1. MacLean, D., & Titah, R. (2022). A systematic literature review of empirical research on the impacts of e-government: A public value perspective. Public Administration Review, 82(1), 23–38.

2. Heilig, L.; Schwarze, S.; Voß, S. An Analysis of Digital Transformation in the History and Future of Modern Ports. 2017. Available online: https://scholarspace.manoa.hawaii.edu/items/56dffdc6-c8db-42de-9c4e-8585cce8515c (accessed on 12 January 2025).

3. Fan, Y.; Guoyong, W.; Riaz, N.; Radlińska, K. Technical efficiency and farm size in the context of sustainable agriculture. Agric. Econ. 2024, 70, 446–456.

4. Matarazzo, M.; Penco, L.; Profumo, G.; Quaglia, R. Digital transformation and customer value creation in Made in Italy SMEs: A dynamic capabilities perspective. J. Bus. Res. 2021, 123, 642–656.

5. Sys, C.; Vanelslander, T. Future maritime supply networks: Key issues in and solutions. In Maritime Supply Chains; Elsevier: Amsterdam, The Netherlands, 2020; pp. 261–282.

6. Sarkar, B.D.; Shankar, R.; Kar, A.K. Port logistic issues and challenges in the Industry 4.0 era for emerging economies: An India perspective. Benchmarking Int. J. 2023, 30, 50–74.

7. Tsiulin, S.; Reinau, K.H.; Hilmola, O.P.; Goryaev, N.; Karam, A. Blockchain-based applications in shipping and port management: A literature review towards defining key conceptual frameworks. Rev. Int. Bus. Strategy 2020, 30, 201–224.

8. Moi, L., & Cabiddu, F. (2021). Leading digital transformation through an Agile marketing capability: The case of Spotahome. Journal of Management & Governance, 25(4), 1145–1177.

9. Huang, H.; Wang, C.; Wang, L.; Yarovaya, L. Corporate digital transformation and idiosyncratic risk: Based on corporate governance perspective. Emerg. Mark. Rev. 2023, 56, 101045.

10. Xiong, X. Examining the influence of knowledge transfer and dynamic capabilities on enterprise digital transformation. PLoS ONE 2024, 19, e0311176.

11. Peter, M. K., Kraft, C., & Lindeque, J. (2020). Strategic action fields of digital transformation: An exploration of the strategic action fields of Swiss SMEs and large enterprises. Journal of Strategy and Management, 13(1), 160–180.

12. Yang, X.; Han, Q. Nonlinear effects of enterprise digital transformation on environmental, social and governance (ESG) performance: Evidence from China. Sustain. Account. Manag. Policy J. 2024, 15, 355–381.

13. European Commission. Directorate-General for Education, Youth, Sport and Culture, Key Competences for Lifelong Learning, Publications Office. 2019. Available online: https://data.europa.eu/doi/10.2766/569540 (accessed on 15 January 2026).

14. Albert, M.B.; Avery, D.; Narin, F.; McAllister, P. Direct validation of citation counts as indicators of industrially important patents. Res. Policy 1991, 20, 251–259.

15. Brunetti, F., Matt, D. T., Bonfanti, A., De Longhi, A., Pedrini, G., & Orzes, G. (2020). Digital transformation challenges: Strategies emerging from a multi-stakeholder approach. The TQM Journal, 32(4), 697–724.

16. Tsekouropoulos, G., Vasileiou, A., Hoxha, G., Theocharis, D., Theodoridou, E., & Grigoriadis, T. (2025). Leadership 4.0: Navigating the challenges of the digital transformation in healthcare and beyond. Administrative Sciences, 15(6), 194.

17. Qi, Y.; Chen, Q.; Yang, M.; Sun, Y. Ambidextrous knowledge accumulation, dynamic capability and manufacturing digital transformation in China. J. Knowl. Manag. 2024, 28, 2275–2305.

18. Veena, G.; Balamurugan, B.B.; Mariofanna, M.; Yovan, F.A. Blockchain, IoT, and AI Technologies for Supply Chain Management: Apply Emerging Technologies to Address and Improve Supply Chain Management; Apress Publishing: New York, NY, USA, 2024.

19. Sofrankova, B.; Sira, E.; Horvathova, J.; Mokrisova, M. Digital Skills and Digital Transformation Performance in the EU-27: A DESI-Based Nonparametric and Panel Data Study. Economies 2025, 13, 315.

20. Cai, S.; Chen, X.; Bose, I. Exploring the role of IT for environmental sustainability in China: An empirical analysis. Int. J. Prod. Econ. 2013, 146, 491–500.

21. Su, Z.; Liu, Y.; Fang, M.; Liu, Z.; Su, M. Couples traveling together and long-haul truckers' transport performance: A theory-based empirical test. Travel Behav. Soc. 2024, 37, 100833.

22. Mvubu, M.; Naude, M.J. Digital transformation at third-party logistics providers: Challenges and best practices. J. Transp. Supply Chain Manag. 2024, 18, 1023.

23. Coe, N.M. Missing links: Logistics, governance and upgrading in a shifting global economy. Rev. Int. Political Econ. 2017, 21, 224–256.

24. Wei, J.; Shen, Y. Impact and mechanism of digital transformation on performance in manufacturing firms. Innov. Green Dev. 2025, 4, 100205.

25. An, J. Maritime logistics and digital transformation with big data: Review and research trend. Marit. Bus. Rev. 2024, 9, 229–242.

26. Yunis, M., Tarhini, A., & Kassar, A. (2018). The role of ICT and innovation in enhancing organizational performance: The catalysing effect of corporate entrepreneurship. Journal of Business Research, 88, 344–356.