Matthias SCHMUCK, PhD candidate (corresponding author)
rsrch@bid-intelligence.com
Alexandru Ioan Cuza University of Iasi, Iasi, Romania

Radu Mircea GEORGESCU, PhD
mirceag@uiac.ro
Alexandru Ioan Cuza University of Iasi, Iasi, Romania

Enablers of Effective Internal Data Governance: A Performance-Oriented Analysis

Abstract. *In the light of increasing data volumes and regulatory requirements, effective internal data governance (IDG) has become a critical driver of organisational data performance. This study investigates key drivers of effective IDG using a performance-oriented approach. Based on a mixed-methods design combining literature search, expert interviews and a cross-industry survey (145 observations from all around Europe), the study identifies seven dimensions that significantly influence IDG effectiveness in a conceptual model: strategy, guidelines, processes, organisation, controlling, communication and change management. We used partial least squares structural equation modelling technique (PLS-SEM) and after analysis of the results, the findings suggest that a clearly defined strategy aligned with business objectives, consistent internal communication, and proactive change management are essential to build organisational support. Formalised guidelines, clear processes and organisational structures clarify roles and responsibilities, while dedicated controlling mechanisms ensure transparency and enable continuous performance monitoring. The interplay of these factors contributes to the successful implementation and sustained operation of IDG. The study offers practical implications for organisations aiming to improve their DG capabilities in a structured and measurable way.*

Keywords: *data governance, key driver, performance measurement, success.*

JEL Classification: M14, M15, M21, C38, C39, O33.

Received: 6 July 2025	Revised: 28 November 2025	Accepted: 2 December 2025

1. Introduction

Data-driven transformation is a strategic shift where data becomes central to decision-making, innovation, and value creation. In dynamic markets, effective data use provides a key competitive edge. This shift requires not only technology, but also cultural change and organisational alignment (Wamba et al., 2017). Its success depends on data governance, leadership support, and alignment between data strategy and business goals (Mikalef et al., 2019). Promoting data literacy and analytical thinking is essential (Provost & Fawcett, 2013). Ultimately, it is a company-wide transformation and not just a tech initiative.

DOI: 10.24818/18423264/59.4.25.07

Data governance (DG), inconsistently defined in research and practice (Jagals et al., 2021), is understood here as a cross-functional framework for managing an organisation's entire data landscape. As part of IT and corporate governance, it ensures reliable, consistent, and sustainable data use. DG implementation spans strategic, tactical, and operational levels and involves both technical and organisational structures. Especially in the context of digitalisation, it must be anchored at the leadership level and supported by clear roles in a human-task-machine system.

This article reviews the literature (Section 2), introduces a conceptual framework and hypotheses (Section 3), outlines the research methodology (Section 4), presents empirical findings (Section 5), and concludes with key theoretical and practical implications (Section 6).

2. Literature review

DG is a complex construct that integrates both technical and organisational aspects, which mutually influence one another (Gluchowski, 2024, pp. 85f). Not all components of DG need to be fully implemented; what matters is a meaningful configuration and combination of the identified dimensions, tailored to the specific requirements of the organisational context. In this study, these dimensions are used as qualitative indicators to examine the interrelationships within a DG approach.

The academic literature offers a wide range of research on success dimensions related to DG, which are summarised in Table 1.

Table 1. Relevant publications for DG and their components

Research area	Reference(s)
DG in general	Alhassan et al. (2019a); Alhassan et al. (2019b); Black et al. (2023); Chandra et al. (2023); Mahanti (2018); Mahanti (2021); Cheong and Chang (2007); Rifaie et al. (2009)
Cloud DG	Al-Rhuite and Benkhelifa (2017); Al-Ruithe et al. (2019)
Data democratisation	Samarasinghe and Lokuge (2022)
Platform ecosystems	Lee et al. (2017)
Urban DG	Bozkurt et al. (2023)
Specific industry/sectors	Romero et al. (2019); Tsavatewa (2023)
Specialised research	Brous et al. (2017)

Source: Authors' elaboration.

The number of dimensions varies across the studies mentioned above, but they show a high degree of conceptual overlap. Therefore, the identified success dimensions were grouped based on common characteristics into the functional categories of "strategic," "tactical," and "operational," and further consolidated into key elements within these categories (Table 2). These key elements represent the latent variables in the IDG model, which will be seen in Section 3.

The IDG model development focused on the strategic and tactical levels of DG, because data management primarily implements the specifications defined by DG, making it more of an operational function. As such, it is not a considered part of DG in the narrower sense. This leads to seven independent variables. The IDG will be also surveyed directly as dependent variable.

Table 2. Data governance dimensions

DG Level	Key components	Variables
strategical	Strategy (ST)	1
tactical	Organisation (OR); Processes (PR); Guidelines (GL); Controlling (CO); Communication (CN); Change management (CM)	6
operational	Data Management	0

Source: Authors' elaboration.

3. Model specification and hypotheses

The core hypothesis is that an effective DG arises from identifying key tasks and integrating them meaningfully. The model begins at the strategic level (ST), which provides direction, defines data objectives, and shapes DG guidelines (GL). Without a clear strategy, DG risks misalignment with business goals. ST influences organisational structure (OR), processes (PR), communication (CN) and change management (CM), leading to Hypotheses H1-H5: ST positive influences GL (H1), OR (H2), PR (H3), CM (H4) and CN (H5). Then follows the tactical level.

GL, the foundation of DG, translates strategic goals into binding standards, ensuring consistent and compliant data handling. GL influences PR, controlling (CO), and the maturity of internal data governance (IDG), supporting H6-H8: GL positive influences PR (H6), CO (H7), and IDG (H8).

OR, the mediator of DG, determines how GL are implemented through roles, responsibilities, and process integration. It is shaped by ST, supported by CN, and enables smooth PR, contributing to IDG maturity (H9-H11): OR positive influences GL (H9), PR (H10), and IDG (H11).

Clearly defined, standardised PR, the operational core of DG, ensure data quality, efficiency, and compliance. They are guided by ST and GL, supported by CN and OR, leading to H12: PR positive influences IDG.

CO, the feedback mechanism within the DG, monitors compliance with GL, tracks data quality, and informs strategic adjustments. With support from CN, CO strengthens IDG, supporting H13 and H14: CO positive influences PR (H13) and IDG (H14).

CN, linking all DG dimensions, is key to communicating GL, fostering understanding and compliance, and enabling adjustments in OR and CO. It supports H15-H18: CN positive influences OR (H15), GL (H16), PR (H17) and CO (H18).

CM, the companion of DG, reflects adaptability to change, supported by ST and CN, and enables improvements in OR, GL, PR, and CO, leading to H19-H22: CM positive influences OR (H19), GL (H20), PR (H21), and CO (H22).

Based on this logic, the proposed research model and hypotheses are shown in Figure 1.

The configuration is a typical example of a complex, multi-level impact model, in which numerous latent constructs are interconnected. The aim of such models is not only to test theoretical assumptions, but also to explore whether and how the constructs influence one another and ultimately explain the target construct.

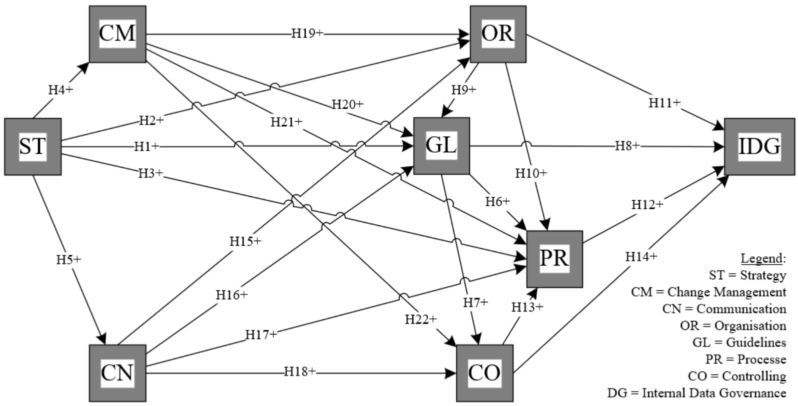

Figure 1. Proposed study model
Source: Authors' elaboration.

4. Methodology

This study employed a mixed-methods approach, integrating both qualitative and quantitative data collection techniques. The methodology was structured in three key phases: literature review, expert interviews, and a quantitative survey.

The first phase of the study involved an extensive review of the relevant literature in the field of DG. This review aimed to establish a comprehensive research framework and identify key success factors influencing DG across three distinct levels: strategic, tactical, and operational. By synthesising existing theoretical frameworks, industry reports, and case studies, the literature review provided a solid foundation for developing the study's research questions and hypotheses.

In the second phase, qualitative data was gathered through expert interviews with both academic scholars and practitioners with extensive experience in DG. These semi-structured interviews were designed to explore the nuances of DG practices in real-world settings. The insights gained from these interviews served two main purposes: refining the research framework identified in the literature review and enhancing the survey instrument. Specifically, the expert input helped identify key DG factors, operational challenges, and decision-making processes that would later inform the survey questions. A total of 15 expert interviews were conducted, each lasting approximately 45 minutes. These interviews were transcribed, and the data was analysed using thematic coding to identify recurring themes and patterns.

Building on the findings from the literature review and expert interviews, a quantitative online survey was developed. The survey aimed to empirically test the hypotheses derived from the research framework. It was designed using LimeSurvey and included 55 items distributed across eight variables, with each item measured on a six-point Likert scale. The survey was targeted at professionals with prior experience in DG, specifically individuals working in data-related roles who held decision-making responsibilities within their organisations. The selection criteria ensured that the participants were knowledgeable about DG and could provide meaningful responses. After pretesting with 10 participants, data collection took place between December 2024 to May 2025, yielding 273 complete responses. The final dataset included 145 responses from European participants, which were used for model evaluation.

5. Data analysis and results

Data analysis was conducted using Microsoft (MS) Excel and SmartPLS 4.1.1.2 with a two-phase PLS-SEM approach. First, the measurement model was assessed for reliability and validity; second, the structural model was analysed to test hypotheses (Hair et al., 2022). PLS was selected to explore psychometric properties and identify potential relationships (Fornell & Larcker, 1981).

5.1 Descriptive statistics

The demographic data of the participants were analysed descriptively to determine their general characteristics and composition. A comprehensive summary is shown in Table 3.

Table 3. Sample descriptive statistics

Factor	Sample	Frequency	Share
Regions	North Europe	2	1%
	West Europe	45	31%
	Middle Europe	10	7%
	East Europe	14	10%
	South East Europe	48	33%
	South Europe	26	18%
Headcount	Small companies	31	21%
	Medium-sized companies	35	24%
	Large companies	79	54%
Industry	Aerospace Industry	2	1%
	Automotive	5	3%
	Education/University/Research	10	7%
	Financial Services/Banking/Insurance	21	14%
	Food/Agricultural Industry	2	1%
	Industry/Production/Chemistry	6	4%
	Information and Communication Technology	41	28%
	Mechanical /Plant Engineering	4	3%
	Media Industry	6	4%
	Medical Technology/Pharmaceuticals	3	2%
	Other	17	12%
	Public Sector/Authorities	5	3%
	Services/Consulting	14	10%
	Trade/Logistics	7	5%
	Utilities /Water/Energy/Waste	2	1%

Source: Authors' elaboration using Microsoft (MS) Excel.

The sample is largely concentrated in South East (33%) and West Europe (31%), followed by South Europe (18%). North Europe is the least represented (1%), while Middle and East Europe account for 7% and 10%, respectively. This reflects a regional focus on southern and western Europe.

Large companies dominate the sample with 54%, followed by medium-sized (24%) and small enterprises (21%), indicating a clear emphasis on larger organisations.

The most represented sector is Information and Communication Technology (28%), followed by Financial Services (14%) and Consulting (10%). Sectors like Aerospace, Agriculture, and Energy are minimally represented (1% each), highlighting a strong focus on tech and finance industries.

5.2 Measurement model assessment

In the second step, construct quality is assessed using predefined reflective indicators. This includes checks for indicator reliability, internal consistency, convergent, and discriminant validity. Bootstrapping with 5,000 resamples ("complete (slow)" mode) was applied. The results are shown in Table 4.

Table 4. Measurement model assessment

Construct	Indicator code	λ	α	CR		AVE
	Thresholds	> 0,7	α ≥ 0,6	rho_a ≥ 0,7	rho_c ≥ 0,7	≥ 0,5
ST	ST_1	0.856	0,955	0,956	0,961	0,711
	ST_2	0.834				
	ST_3	0.835				
	ST_4	0.884				
	ST_5	0.872				
	ST_6	0.873				
	ST_7	0.807				
	ST_8	0.862				
	ST_9	0.800				
	ST_10	0.804				
GL	GL_1	0.774	0,868	0,872	0,901	0,602
	GL_2	0.748				
	GL_3	0.701				
	GL_4	0.744				
	GL_5	0.793				
	GL_6	0.835				
PR	PR_1	0.808	0,895	0,899	0,917	0,613
	PR_2	0.732				
	PR_3	0.768				
	PR_4	0.817				
	PR_5	0.770				
	PR_7	0.759				
	PR_8	0.800				
OR	OR_1	0.827	0,931	0,932	0,944	0,708
	OR_2	0.847				
	OR_3	0.880				
	OR_4	0.850				

Construct	Indicator code Thresholds	λ > 0,7	α α ≥ 0,6	CR rho_a ≥ 0,7	 rho_c ≥ 0,7	AVE ≥ 0,5
	OR_5	0.835				
	OR_6	0.794				
	OR_7	0.786				
CO	CO_2	0.722	0,816	0,829	0,879	0,645
	CO_3	0.825				
	CO_5	0.831				
	CO_6	0.748				
CN	CN_1	0.817	0,843	0,860	0,888	0,613
	CN_2	0.835				
	CN_3	0.783				
	CN_4	0.742				
	CN_6	0.726				
CM	CM_1	0.770	0,895	0,913	0,919	0,653
	CM_2	0.796				
	CM_3	0.852				
	CM_4	0.858				
	CM_5	0.757				
	CM_6	0.811				
IDG	DG_1	0.922	0,870	0,886	0,921	0,795
	DG_2	0.927				
	DG_3	0.821				

Note: CM = Change Management | CN = Communication | CO = Controlling | IDG = Internal Data Governance | GL = Guidelines | OR = Organisation | PR = Processes | ST = Strategy | λ = Loadings | α = Cronbach's Alpha | CR = Composite reliability | AVE = Average Variance Extracted

Source: Authors' elaboration using SmartPLS (PLS-SEM algorithm, Bootstrapping).

Seven indicators (GL_7, PR_6, OR_8, OR_9, CO_1, CO_4, CN_5) were removed due to low factor loadings. The remaining indicators showed acceptable values, confirming indicator reliability. Cronbach's Alpha and composite reliability indicated strong internal consistency, while the AVE values confirmed the convergent validity.

Discriminant validity was supported by HTMT values below the 0.9 threshold (Table 5). A value of 0.90 is acceptable under a conservative criterion, while 0.85 is recommended for a more stringent assessment.

Table 5. Discriminant validity using HTMT

	CM	CN	CO	IDG	GL	OR	PR	ST
CM								
CN	0.705							
CO	0.494	0.557						
DG	0.588	0.689	0.793					
GL	0.535	0.595	0.810	0.854				
OR	0.561	0.598	0.808	0.804	0.884			
PR	0.476	0.588	0.760	0.779	0.863	0.692		
ST	0.609	0.659	0.773	0.830	0.871	0.844	0.811	

Note: CM = Change Management | CN = Communication | CO = Controlling | DG = Data Governance | GL = Guidelines | OR = Organization | PR = Processes | ST = Strategy

Source: Authors' elaboration using SmartPLS (PLS-SEM algorithm).

The combination of high reliability, significant indicator loadings, and confirmed convergent and discriminant validity provides a solid foundation for proceeding with the evaluation of the structural model, which outlines in the following section.

5.3 Structural model assessment and hypotheses testing

Step 1 - Collinearity analysis (Table 6): Before evaluating the structural model, collinearity was assessed. All VIF values were below the critical threshold of 5.0 (Hair et al., 2022). GL, OR, and ST showed values above 3.3 for some predictors, indicating moderate but acceptable collinearity.

Table 6. Collinearity (VIF values) in the structural model

	CM	CN	CO	GL	IDG	OR	PR	ST
CM			1.771	1.886		1.868	1.887	
CN			1.828	1.917		1.901	1.930	
CO					2.403		2.312	
GL			1.478		4.083		3.592	
IDG								
OR				2.829	3.207		3.771	
PR					2.641			
ST	1.000	1.000		3.169		1.759	3.897	

Note: CM = Change Management | CN = Communication | CO = Controlling | IDG = Internal Data Governance | GL = Guidelines | OR = Organisation | PR = Processes | ST = Strategy

Source: Authors' elaboration using SmartPLS (PLS-SEM algorithm).

Step 2 - Path coefficients analysis (Table 10): Analysis of the path coefficients shows that 14 out of 22 values are ≥ 0.1, indicating general relationships, with 11 of them ≥ 0.2, suggesting substantial effects. ST emerges as a key driver with strong influences on GL (0.429), OR (0.706), and PR (0.380), but also for CM (0.588) and CN (0.5979. GL acts as a central mediator between ST and IDG by activating other resources. Notably, GL has a strong direct effect on CO (0.592), highlighting CO's role in implementing and monitoring GL. There is also a strong effect to PR (0.464) underline, that well-defined GL significantly support the development and implementation of structured, efficient, and compliant data-related PR. This relationship highlights the importance of formalized GL in shaping operational routines and ensuring that data handling within the organisation follows defined quality, security, and compliance requirements. OR supports the activation of GL (0.448), enhancing its integration into business processes. Overall, the relationships are theoretically sound and consistent with the proposed model.

Table 7. Total effects

	CM	CN	CO	GL	IDG	OR	PR	ST
CM			0.080	0.023	0.033	0.078	-0.028	
CN			0.170	0.065	0.098	0.076	0.146	
CO					0.204		0.193	
GL			0.592		0.458		0.579	

	CM	CN	CO	GL	IDG	OR	PR	ST
IDG								
OR			0.265	0.448	0.439		0.052	
PR					0.241			
ST	0.588	0.597	0.590	0.797	0.676	0.797	0.735	

Note: CM = Change Management | CN = Communication | CO = Controlling | IDG = Internal Data Governance | GL = Guidelines | OR = Organisation | PR = Processes | ST = Strategy

Source: Authors' elaboration using SmartPLS (PLS-SEM algorithm).

Step 3 - Total effects analysis (Table 7): The strongest total effects on IDG come from GL and OR (both 0.797), making them key drivers that should be prioritised. PR (0.735) also exerts a strong, primarily indirect influence. GL (0.676) remains an important mediator within the network. CM, CN, and CO show slightly lower effects (0.588 to 0.597), but still have a substantial impact, indicating a broad distribution of contributing factors to IDG. All this points to strong structural linkages within the model.

Step 4 - Model fit analysis: To assess the model's explanatory power, the coefficient of determination R^2 (and adjusted R^2) is used. An R^2 of 0.650 for IDG was evaluated and indicates substantial explanatory strength (Hair et al., 2022). The model explains 65.0% of the variance in IDG through direct effects of strategy, organisation, guidelines, processes, controlling, communication, and change management, as well as their mediating roles. The evaluated adjusted R^2 of 0.640 deviates only slightly, confirming the model's robustness.

Step 5 - Predictive power analysis of the model: To assess predictive power, the PLSpredict procedure was applied in SmartPLS using default settings (10 folds, 10 repetitions) to calculate a first metric, Q^2 (Table 8). For the target construct IDG (indicators DG_1 to DG_3), all Q^2_predict values were clearly above zero, indicating superior predictive performance of the PLS-SEM over indicator means. This also holds when compared to a linear model (LM), as PLS-SEM showed lower RMSE and MAE values across all three indicators.

Table 8. PLSpredict MV summary

	Q²predict	PLS-SEM		LM		IA	
	> 0	RMSE	MAE	RMSE	MAE	RMSE	MAE
DG_1	0.517	0.845	0.673	0.867	0.684	1.215	0.992
DG_2	0.493	0.858	0.702	0.880	0.713	1.205	1.001
DG_3	0.331	1.039	0.881	1.100	0.926	1.271	1.024

Note: RMSE = Root Mean Square Error | MAE = Mean Absolute Error | LM = linear model | IA = Indicator Average

Source: Authors' elaboration using SmartPLS (PLSpredict/CVPAT).

Another metric used is **CVPAT LV**, whose results (Table 9) provide additional evidence of the model's predictive strength.

Table 9. CVPAT LV summary

	PLS-SEM vs. IA			PLS-SEM vs. LM		
	ALD	t value	p value	ALD	t value	p value
CM	-0.278	3.976	0.000	-0.278	3.976	0.000
CN	-0.255	4.027	0.000	-0.255	4.027	0.000

	PLS-SEM vs. IA			PLS-SEM vs. LM		
CO	-0.539	4.706	0.000	-0.539	4.706	0.000
GL	-0.608	6.338	0.000	-0.608	6.338	0.000
IDG	-0.671	7.176	0.000	-0.671	7.176	0.000
OR	-0.925	6.192	0.000	-0.925	6.192	0.000
PR	-0.592	5.236	0.000	-0.592	5.236	0.000
TOTAL	-0.563	7.421	0.000	-0.563	7.421	0.000

Note: IA = Indicator Average | LM = linear model | ALD = Average loss difference | CM = Change Management | CN = Communication | CO = Controlling | IDG = Internal Data Governance | GL = Guidelines | OR = Organisation | PR = Processes | ST = Strategy

Source: Authors' elaboration using SmartPLS (PLSpredict/CVPAT).

PLS-SEM compared to Indicator Average (IA) means, the negative values indicate that PLS-SEM consistently shows lower average loss, confirming superior prediction. The differences are statistically significant (see t- and p-values). The same applies when comparing PLS-SEM to the linear model (LM), with negative loss differences and significant results (> 1.645).

Overall, the structural model demonstrates strong predictive quality, making it suitable for identifying which factors contribute simultaneously to effective data governance. The following section presents the relationships between constructs based on bootstrap significance testing.

5.4 Evaluating the single variables

The results of the testing hypotheses are shown in Table 10.

Table 10. Results of hypothese testing

Paths	Hypotheses		β	t-value [1]		Result	f^2	Effect(s)[2]
ST → GL	H1	+	0.429	4.995	***	confirmed	0.200	medium
ST → OR	H2	+	0.706	10.504	***	confirmed	0.801	strong
ST → PR	H3	+	0.380	3.283	***	confirmed	0.113	weak
ST → CM	H4	+	0.588	11.070	***	confirmed	0.528	strong
ST → CN	H5	+	0.597	11.780	***	confirmed	0.555	strong
GL → PR	H6	+	0.464	4.388	***	confirmed	0.183	medium
GL → CO	H7	+	0.592	8.764	***	confirmed	0.479	strong
GL → IDG	H8	+	0.225	2.124	**	confirmed	0.035	weak
OR → IDG	H9	+	0.283	2.643	***	confirmed	0.071	medium
OR → PR	H10	+	-0.208	2.073	**	confirmed	0.035	weak
OR → GL	H11	+	0.448	5.813	***	confirmed	0.245	weak
PR → IDG	H12	+	0.241	2.145	**	confirmed	0.063	weak
CO → PR	H13	+	0.193	2.794	***	confirmed	0.049	weak
CO → IDG	H14	+	0.158	2.219	**	confirmed	0.029	weak
CN → OR	H15	+	0.076	1.066	n.s.	rejected	0.009	weak
CN → GL	H16	+	0.031	0.518	n.s.	rejected	0.002	weak
CN → PR	H17	+	0.099	1.589	n.s.	rejected	0.015	weak
CN → CO	H18	+	0.132	1.714	*	confirmed	0.019	weak
CM → OR	H19	+	0.078	1.069	n.s.	rejected	0.009	weak

Paths	Hypotheses		β	t-value [1]		Result	f²	Effect(s)[2]
CM → GL	H20	+	-0.012	0.195	n.s.	rejected	0.000	weak
CM → PR	H21	+	-0.038	0.633	n.s.	rejected	0.002	weak
CM → CO	H22	+	0.067	0.983	n.s.	rejected	0.005	weak

Note: CM = Change Management | CN = Communication | CO = Controlling | DG = Data Governance | GL = Guidelines | OR = Organisation | PR = Processes | ST = Strategy | β = path coefficient | [1] Significance level: n.s. (not significant) | * (p < 0,1 / 10%) → t ≥ 1,645 | ** (p < 0,05 / 5%) → t ≥ 1,960 | *** (p < 0,01 / 1%) → t ≥ 2,576 | [2] f-Square: 0,02 (weak effect); 0,15 (medium effect); 0,35 (strong effect)

Source: Authors' elaboration using SmartPLS (PLS-SEM algorithm, Bootstrapping).

ST: In the model, positive relationships were hypothesised between ST and GL (H1), OR (H2), PR (H3), CM (H4), and CN (H5). All were empirically confirmed and are statistically highly significant ($p < 0.01$). ST shows a medium effect on GL (H1), a weak effect on PR (H3), and strong effects on OR (H2), CM (H4), and CN (H5), supported by the f^2 values.

GL: As hypothesised, positive relationships were confirmed between GL and PR (H6), CO (H7), and IDG (H8), and partly statistically highly significant ($p < 0.01$). According to the f^2 values, the effects on PR are medium, on IDG are weak to moderate, while the effect on CO is strong.

OR: Positive relationships were hypothesised between OR and GL (H9), PR (H10), and IDG (H11). These were all confirmed. Significant effects were found for H9 and H11, with a weak effect on GL, a medium effect on IDG and a weak effect on PR.

PR: The hypothesised positive relationship between PR and IDG (H12) with significant level was confirmed with weak significance, though the effect is weak.

CO: The hypothesised positive relationships between CO and both PR (H13) and IDG (H14) were confirmed with statistical significance and weak effect.

CN: Of the hypothesised positive effects of CN on OR (H15), Guidelines (H16), PR (H17), and CO (H18), only the link to CO was empirically confirmed, with high significance and a strong effect.

CM: The hypothesised positive effects of CM on OR (H19), Guidelines (H20), PR (H21), and CO (H22) could not be empirically confirmed.

5.5 Mediating effects in the network

The model includes multiple mediations, revealing indirect relationships between key constructs. Indirect effects are first calculated using the PLS-SEM algorithm, then tested for significance via bootstrapping. Results are shown in Table 11.

Table 11. Total indirect effects

Paths	TIE	t-value [1]		Paths	TIE	t-value [1]	
CM → CO	0.013	0.323	n.s.	GL → PR	0.114	2.526	**
CM → GL	0.035	0.975	n.s.	OR → CO	0.265	4.806	***
CM → IDG	0.033	0.649	n.s.	OR → IDG	0.155	2.248	**
CM → PR	0.010	0.238	n.s.	OR → PR	0.260	3.850	***

Paths	TIE	t-value [1]		Paths	TIE	t-value [1]	
CN → CO	0.038	1.038	n.s.	ST → CO	0.590	13.196	***
CN → GL	0.034	1.026	n.s.	ST → GL	0.368	5.278	***
CN → IDG	0.098	2.336	**	ST → IDG	0.676	15.565	***
CN → PR	0.047	1.261	n.s.	ST → OR	0.091	2.015	**
CO → IDG	0.047	1.580	n.s.	ST → PR	0.355	3.610	***
GL → IDG	0.233	3.200	***				

Note: CM = Change Management | CN = Communication | CO = Controlling | DG = Data Governance | GL = Guidelines | OR = Organisation | PR = Processes | ST = Strategy | TIE = Total indirect effects | [1] Significance level: n.s. (not significant) | * (p < 0,1 / 10%) → t ≥ 1,645 | ** (p < 0,05 / 5%) → t ≥ 1,960 | *** (p < 0,01 / 1%) → t ≥ 2,576

Source: Authors' elaboration using SmartPLS (Bootstrapping).

ST emerges as the key driver in the network, showing the strongest total indirect effects on most target variables. In contrast, CM plays a minor role with weak indirect effects, likely serving a supportive, possibly moderating or stabilising function. GL acts as a mediator, channelling various effects. CN and OR show moderate to strong indirect effects, suggesting that they function as amplifiers within the network.

Figure 2 shows the estimated structure model in PLS SEM.

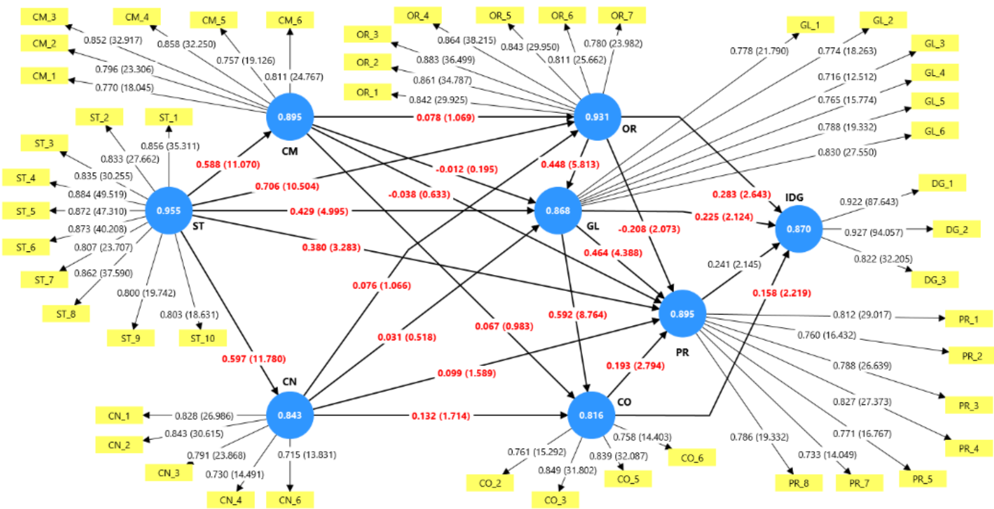

Figure 2. Estimated structural model

Source: Authors' elaboration with SmartPLS (Bootstrapping).

6. Discussion

This study has contributed to the development of a comprehensive measurement instrument for IDG. The proposed model incorporates six dimensions – ST (driver), GL (foundation), PR (core), OR (enabler), CO (feedback), CN (link), and CM (supporting process) – with IDG positioned as the target dimension. Empirical validation of the model confirmed its factor structure and demonstrated strong validity and reliability. These findings support the conceptualisation of IDG as a multi-dimensional construct with interconnected roles that contribute to its overall effectiveness and sustainability within organisations.

The core hypothesis suggests that high-quality IDG results when key tasks are identified and integrated effectively, supporting sustainable implementation. This aligns with research highlighting the importance of strategic and operational integration for governance success (Robey et al., 2008). The model's validation confirms its structure and reliability, reinforcing the idea that effective governance relies on the systematic organisation of tasks and information (Mert & Pattberg., 2015).

ST was identified as the primary driver of the IDG construct, influencing other dimensions like CM, OR, GL, PR, and CN. The strong effects of ST, with path coefficients ranging from 0.380 to 0.706, indicate that changes in ST can have system-wide impacts, underscoring its central role in governance models (Kogut & Zander, 1992). ST influence on OR and CN highlights the importance of adaptable, responsive governance systems (Pfeffer, 1982), supporting the view that strategic drivers are crucial to organisational success.

GL acts as a foundation interface between strategy and operations, translating strategic goals into actionable tasks, consistent with Kaplan and Norton's (1996) work on strategy execution. OR and CN further amplify and distribute the strategic framework, emphasising the need for effective communication and resource allocation in governance (Vatne & Taylor, 2000). CO serves as a feedback loop, enabling continuous improvement, while CM plays a moderating role, ensuring stability within the governance system (Gulati & Puranam, 2009).

7. Conclusions, limitations, and future studies

Globalisation and digitalisation have fundamentally transformed the way organisations operate. Data, seen as the "digital footprints" of established business processes and activities, are playing an increasingly important role in process optimisation, informed decision-making, innovation, and value creation, evolving into a true strategic asset. This transformation increasingly depends on well-defined data policies and processes, active support from all levels of leadership, and strong alignment between data strategy and overall business objectives.

From a theoretical perspective, this study contributes to the existing literature on DG by developing a conceptual framework that identifies key success factors and integrates them into a causal model for measuring IDG. The effectiveness of IDG is influenced by the following factors: ST (driver), GL (foundation), PR (core), OR (enabler), CO (feedback), CN (link), and CM (supporting process).

Furthermore, this study introduces the first measurement model for internal data governance (IDG). In doing so, it addresses a current gap in the literature: Although numerous existing models describe the structure of data governance, they do not capture the mutual dependencies and interrelationships between its components, nor do they offer ways to quantify them.

One of the key practical implications of this study relates to strategic considerations. In developing the overall model, ST was identified as a central driver, meaning that any intervention in, or even the presence of, ST is likely to significantly influence internal data governance (IDG). GL channels these effects and serves as a functional interface between the strategic and operational architecture. OR and CN act as amplifiers and distributors of the strategic framework, which is reflected in the GL. CO functions as a feedback mechanism, while CN also acts as a connecting link, both contributing to the overall effectiveness of IDG. CM was found to have a moderating or stabilising effect on the system.

Overall, this study provides valuable insights for practitioners seeking to enhance their data management performance. It offers specific strategies and tactics that can help achieve better outcomes in internal data governance. In particular, we believe that the study's findings may foster a more positive perception of data governance and help to overcome existing negative connotations.

The study is subject to certain limitations in terms of scope. Moreover, as a quantitative study, the use of a questionnaire constrains the depth and richness of the qualitative insights.

This study is based on cross-sectional, cross-industry data. For future research, it would be valuable to expand the geographical scope to include companies outside Europe, particularly in North America and Asia, because of amount of publications. There is also a need to investigate industry-specific success factors in more depth. In addition, future studies should consider longitudinal designs to better capture temporal dynamics and causal relationships.

DG should not be viewed in isolation, but rather as part of broader governance frameworks such as corporate governance, IT governance, or AI governance. Further research could explore the boundaries, overlaps, and integration of these governance forms, especially regarding their internal implementation and impact on performance.

Soft factors such as corporate culture, employee acceptance, and readiness for change were only partially addressed through constructs within the success model. Future studies should examine more closely how organisational culture, leadership, and data literacy at various hierarchical levels influence the effectiveness of data governance.

There is also a continued need for research into the measurability and operationalisation of data governance initiatives. Lastly, the role of data governance within the context of ongoing technological change warrants further investigation.

References

[1] Abraham, R, Schneider, J, vom Brocke, J (2019), *Data governance: A conceptual framework, structured review, and research agenda. International Journal of Information Management,* 49, 424-438, https://doi.org/10.1016/j.ijinfomgt.2019.07.008.

[2] Alhassan, I., Sammon, D., Daly, M. (2019a), *Critical Success Factors for Data Governance: A Theory Building Approach. Information Systems Management*, 36(2), 98-110, https://doi.org/10.1080/10580530.2019.1589670.

[3] Alhassan, I., Sammon, D., Daly, M. (2019b), *Critical success factors for data governance: a telecommunications case study. Journal of Decision Systems*, 28(1), 41-61, https://doi.org/10.1080/12460125.2019.1633226.

[4] Al-Ruithe, M., Benkhelifa, E. (2017), *Analysis and Classification of Barriers and Critical Success Factors for Implementing a Cloud Data Governance Strategy. In Proceedings of the 8th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2017), Procedia Computer Science,* 113, 223-232, https://doi.org/10.1016/j.procs.2017.08.352.

[5] Al-Ruithe, M., Benkhelifa, E., Hameed, K. (2019), *A systematic literature review of data governance and cloud data governance. Personal and Ubiquitous Computing*, 23(5-6), 839–859, https://doi.org/10.1007/s00779-017-1104-3.

[6] Vatne, E., Taylor, M. (2000), *The Networked Firm in a Global World. Small Firms in New Environments*. 1st Edition. London, UK: Routledge, https://doi.org/10.4324/9781315182957.

[7] Black, S., Davern, M., Maynard, SB., Nasser, H. (2023), *Data governance and the secondary use of data: The board influence. Information and Organization*, 33(2), 1-23, https://doi.org/10.1016/j.infoandorg.2023.100447.

[8] Bozkurt, Y, Rossmann, A, Konanahalli, A, Pervez, Z. (2023), *Toward Urban Data Governance: Status-Quo, Challenges, and Success Factors. IEEE Access*, 11, 85656-85677, https://doi.org/10.1109/ACCESS.2023.3302835.

[9] Brous, P, Janssen, M., Schraven, D., Spiegeler, J., Can Duzgun, B. (2017a), *Factors Influencing Adoption of IoT for Data-driven Decision Making in Asset Management Organizations. In Conference: IoTBDS, 2nd International Conference on IoT, Big Data and Security*, https://doi.org/10.5220/0006296300700079.

[10] Mert, A., Pattberg, P. H. (2015). *Public-private partnerships and the governance of ecosystem services. In J. A. Bouma, & P. van Beukering (Eds.), Ecosystem Services: From Concept to Practice*, 230-249. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781107477612.016.

[11] Chandra, Y.U., Prabowo, H., Gaol, F.L., Purwandari, B. (2023), *Control and data integrity are important factors of data governance technology. In 2023 10th International Conference on ICT for Smart Society (ICISS)*, *Bandung, Indonesia, 2023*, 1-6, https://doi.org/10.1109/ICISS59129.2023.10291563.

[12] Cheong, L.K., Chang, V. (2007), *The Need for Data Governance: A Case Study. In ACIS 2007 - Proceedings of the 18th Australasian Conference on Information System, Toowoomba*, *Australia, 5-7 Dec 2007*, 999-1008.

[13] Fornell, C., Larcker, D.F. (1981), *Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research*, 18(1), 39-50, https://doi.org/10.2307/3151312.

[14] Gluchowski, P. (2024), *Datenmanagement. In Gluchowski, P., Datenmanagement und Datenanalyse*. Wiesbaden, Germany: Springer Vieweg, 2024, https://doi.org/10.1007/978-3-658-44392-4_9.

[15] Gulati, R, Puranam, P. (2009), *Renewal Through Reorganization*: *The Value of Inconsistencies Between Formal and Informal Organization. Organization Science*, 20(2), 262-276.

[16] Hair, J.F., Hult, T.G.M., Ringle, C.M., Sarstedt, M., Richter, N.F., Hauff, S. (2022), *Partial Least Strukturgleichungsmodellierung. Eine anwendungsorientierte Einführung. 2., vollständig überarbeitete Auflage*. München: Franz Vahlen, Germany.

[17] Jagals, M., Karger, E., Ahlemann, F. (2021), *Already grown-up or still in puberty? A bibliometric review of 16 years of data governance research. Corporate Ownership & Control*, 19(1), 105-120, https://doi.org/10.22495/cocv19i1art9.

[18] Kaplan, R.S., Norton, D.P. (1996), *The Balanced Scorecard: Translating Strategy into Action*. *Brighton (Boston),* Massachusetts, USA: Harvard Business Review Press.

[19] Kogut, B., Zander, U. (1992), *Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. Organization Science*, 3(3), 383-397, https://doi.org/10.1287/orsc.3.3.383.

[20] Lee, S.U., Zhu, L., Jeffery, R., Group, A.P. (2017), *Data Governance for Platform Ecosystems: Critical Factors and the State of Practice. In Twenty First Pacific Asia Conference on Information Systems, Langkawi, Malaysia, 2017* (PACIS 2017 Proceedings), 89, https://arxiv.org/ftp/arxiv/papers/1705/1705.03509.pdf.

[21] Mahanti, R. (2018), *Data Governance Success. Growing and Sustaining Data Governance*. Singapore: Springer, https://doi.org/10.1007/978-981-16-5086-4.

[22] Mahanti, R. (2021) *Data Governance and Data Management. Contextualizing Data Governance Drivers, Technologies, and Tools*. Singapore: Springer Nature Singapore Pte Ltd., 2021.

[23] Mikalef, P., Krogstie, J., Pappas, I. O., Pavlou, P. (2019), *Investigating the effects of big data analytics capabilities on firm performance: The mediating role of dynamic capabilities. Information & Management*, 56(8), 103207.

[24] Pfeffer, J. (1982), *Organizations and Organization Theory*. Boston, USA: Pitman Publishing.

[25] Provost, F., Fawcett, T. (2013), *Data science and its relationship to big data and data-driven decision making. Big Data*, 1(1), 51-59.

[26] Rifaie, M., Alhajj, R., Ridley, M. (2009), *Data governance strategy: A key issue in building enterprise data warehouse. In Proceedings of the iiWAS2009 - 11th Int Conf Inf Integr Web-based Appl Serv., Kuala Lumpur, Malaysia, 2009*, 587-591, https://doi.org/10.1145/1806338.1806449.

[27] Robey, D., Boudreau, M-C., Rose, G.M. (2008), *The Influence of Information Technology on Organizational Processes. Organizational Science*, 19(4), 705-722, https://doi.org/10.1016/S0959-8022(99)00017-X.

[28] Romero, A., Gonzales, A., Carlos Raymundo, C. (2019), *Data Governance Reference Model under the Lean Methodology for the Implementation of Successful Initiatives in the Peruvian Microfinance Sector. In Proceedings of the 8th International Conference on Software and Information Engineering (ICSIE '19). Association for Computing Machinery*, *New York, NY, USA*, 227-231, https://doi.org/10.1145/3328833.3328859.

[29] Samarasinghe, S., Lokuge, S. (2022), *Exploring the Critical Success Factors for Data Democratization. In ACIS - Australasian Conference on Information Systems*, *4 Dec 2022*, *Melbourne, Australia*, https://doi.org/10.48550/arXiv.2212.03059.

[30] Tsavatewa, C.C. (2023), *Perceptions of Data Governance: Identifying Critical Success Factors in a University System's Implementation Effort. Issues in Information Systems*, 24(4), 390-411, https://doi.org/10.48009/4_iis_2023_130.

[31] Wamba, S.F., Akter, S., Edwards, A., Chopin, G., Gnanzou, D. (2017), *How 'big data' can make big impact: Findings from a systematic review and a research agenda. International Journal of Information Management*, 37(1), 1-17.