



ESG RESEARCH INSIGHTS PAPER

2022 State of Data Governance and Empowerment

By Mike Leone, ESG Senior Analyst; and Keir Walker, Senior Market Research Analyst July 2022

This ESG Research Insights Paper was commissioned by Quest Software, Inc. and is distributed under license from TechTarget, Inc.



2

Contents

Executive Summary	3
Research Objectives	3
Key Findings	3
Introduction	4
Adopting and Defining Data Governance	4
Examining Data Governance Actuation	5
Roadblocks to Successful Data Governance Programs	7
Data Management Challenges and Bottlenecks: The Quest for Efficient Insights and Clarity	7
Massive Opportunity to Leverage More Data	9
Automation Affords Opportunities to Improve Data Governance	10
Data Analytics Challenges	10
The Recognized Value of DataOps	11
Heavily Used DataOps for Data Engineering Purposes	12
Conclusion	12
Research Methodology	13
Respondent Demographics	14

Executive Summary

In today's complex business world, there are few commodities as crucial to an organization's success as data. At every level of business, the ability to access, manage, analyze, and make decisions based on data is of paramount importance. And yet, with so much information, so many entry points, and vast reservoirs of data stored and often remaining dark (unknown or unused), the technology and processes that facilitate access, enhance visibility, automate workflows, and improve analytical output based on data are increasingly critical. Conceptually, managing data continues to evolve; where once it was treated as individual tasks or efforts with limited (or almost no) cohesive strategy, today methods such as DataOps streamline and consolidate fragmented data management strategies. DataOps is a driver to accelerate and strengthen an organization's data delivery capabilities to its consumers. DataOps and data governance are tightly interwoven and, as such, continue to grow in importance.

Research Objectives

To extend and expand upon Quest's prior State of Data Governance and Empowerment research and companion report, ESG conducted a research study of 220 business and IT professionals responsible for and/or familiar with data governance and empowerment strategies, investments, and operations at their organizations. All the organizations represented in the research have 1,000+ employees and annual revenues of \$100M+. Key aspects of the project included:

- Understanding how organizations define, adopt, and prioritize data governance.
- Uncovering the drivers and challenges of data governance implementations across key data lifecycle and data technology integration points.
- Providing YoY comparisons/data of fundamental data governance functions and trends (2021 versus 2022).
- Investigating the adoption levels and efficacy of DataOps and its relationship to an organization's data governance initiatives.

Key Findings

The key findings of this research can be summarized as follows:

- *Emphasis on data quality resonates amongst all aspects of data governance.* Concerns and efforts around data quality permeated throughout all levels of the research and were clearly established re-occurring themes throughout the findings. One such highlight underlying quality concern found that 92% of all respondents monitor databases as part of data governance. Yet a more precise illustration of data quality's value is highlighted by the fact that data quality was the top response in several survey areas: Data quality was the top response for how organizations define data governance (47%), it was the top challenge that organizations have to overcome to maximize the return on their data governance efforts (45%), and it was the top driver for organizations' data governance programs (41%).
- Data management is greatly challenged by a lack of visibility and compounded by data quality issues. A fundamental piece of data governance is the ability for organizations to effectively manage data. At its core, two of the overwhelming challenges to optimized data management are limited visibility and data quality concerns. This is well exemplified by the levels of dark (i.e., retained but unused, or unknown) data: 42% of all respondents indicated at least half of their data was "dark." Additionally, 46% of all respondents identified understanding the quality of source data as an impediment to efficiently maximizing the use of data in their value chain, which was the top response.

• *DataOps is translating into improved data accuracy and efficiencies*. DataOps is an agile, automated, and processoriented methodology used by data stakeholders to improve the quality, delivery, velocity, and management of data and analytics. While it is still somewhat new conceptually, the research data reflects that it is being adopted: 86% of respondents believe their organizations are engaged in a DataOps approach to managing data, 90% believe DataOps initiatives are improving data quality, and 87% believe DataOps initiatives have resulted in increased operational efficiencies.

Introduction

Organizations rely on their data more than ever before. From IT and business operations, through strategic decisions, to sales and marketing pipelines, the need for technology professionals to maximize value from their organization's data continues to elevate in importance. Data governance encapsulates the policies and procedures necessary to ensure an organization's data is accurate and the strategies needed to deliver increased visibility, transparency, and control, while reducing risk across the entire data estate. Organizations are giving data governance more attention and increasingly prioritizing it in their IT planning. The following research has identified the drivers and challenges facing data governance practitioners, as well as the tools and trends that can help alleviate challenges, side-step obstacles, identify where data governance tools can assist, and show how a holistic DataOps approach can also help organizations better achieve their data governance goals.

As initially alluded to, the singular common thread unveiled throughout the research, and thus throughout this report, is data quality. Achieving data quality is both a driver and a challenge experienced by sizable respondents. The downstream effects of data quality intrinsically have ramifications felt throughout data governance efforts. However, before diving deep into how data quality touches all points of data governance, it's best to also establish exactly what data governance is, what is driving it, and what obstacles organizations experience to finding success with data governance.

Adopting and Defining Data Governance

The previously offered working definition of data governance stands true: data governance encapsulates the policies and procedures necessary to ensure an organization's data is accurate and that the organization has the strategies needed to deliver increased visibility, transparency, and control, while reducing risk across the entire data estate. This definition is provided, however, in a general context. As is so often the case, what we see and believe to be true is largely contextual, being highly dependent on our own point of view and biases. This should resonate with respondents, who share different points of view as to what data governance personally means to each one of them.

Figure 1. Defining Data Governance



Source: ESG, a division of TechTarget, Inc

When ESG asked the question "How does your organization define data governance?" of IT decision makers (ITDM) and line-of-business (LoB) personas, the two exhibited differences in their perspectives on data governance (see Figure 1). IT decision makers' top response was ensuring that sensitive data is classified and meets legal/regulatory requirements, at 52%. This stands to reason, since ITDMs have consistently raised their regulatory concerns over the years due to environmental changes (such as GDPR) and threat-mitigation-emphasizing controls (such as access control). When viewed by all personas, the top answer shared by 47% of all respondents, identified data quality as a defining characteristic of what data governance means. The largest delta between the personas defining data governance was ITDMs being 66% more likely than LoB respondents to identify data governance as clearly conveying data in terms of relatable business context.

Examining Data Governance Actuation

It should come as no surprise that those most heavily engaged in the data governance space recognize data quality as one of the top drivers of their data governance programs (41% of all respondents, see Figure 2). Improving data security (37%), improving analytics (30%), increasing customer trust/satisfaction (26%), and supporting better decision-making (25%) were the next highest ranked drivers of data governance, *all* of which are fundamentally predicated on sound data quality. Without accuracy of or confidence in one's data, improving all subsequent functions becomes virtually impossible.

Figure 2. Data Governance Top Drivers

What are the top drivers of your organization's data governance program (i.e., what are the top outcomes it hopes to achieve)? (Percent of respondents, N=220, three responses accepted)



Source: ESG, a division of TechTarget, Inc

What we find, however, when the responses are viewed in terms of the IT decision makers and line-of-business personas is that there is an almost inverse data quality and data security response rate:

- Improve data quality: IT 45% versus LoB 36%.
- Improve data security: IT 35% versus LoB 41%.

This highlights the difference in perspectives and priorities between the two respondent groups, illustrating that ITDMs tend to place greater significance on data quality, whereas LoB users place greater significance on security. Regardless of which aspect is given priority, data governance solutions with capabilities that feature automation and integration and data quality tools (such as data profiling, quality assessments, and scoring) can help to augment organizations' data quality and, ultimately, its many dependencies.

Roadblocks to Successful Data Governance Programs

If this research has uncovered anything, it is the catch-22 dilemma presented by quality. While there is an overwhelming respondent preference for data quality as a driver of data governance, it is also considered one of the obstacles to achieving data governance success. Data quality/accuracy (45% of all respondents) was the most-cited challenge to achieving ROI on data governance (see Figure 3). Skills shortage/gap (40%) and operationalizing data governance (35%) round out the next top responses. All three responses accentuate the need for automation to make operationalization of data governance more practical and to make data quality assessment, monitoring, improvement, and enhanced visibility of data fitness more broadly available across organizations.

Figure 3. Obstacles in Achieving Data Governance ROI



What challenges does your organization have to overcome to maximize the return on its data governance efforts? (Percent of respondents, N=220, multiple responses accepted)

Source: ESG, a division of TechTarget, Inc

Data Management Challenges and Bottlenecks: The Quest for Efficient Insights and Clarity

Speaking more broadly to data empowerment, the survey also questioned respondents about data management challenges. The data shows that data management challenges may be primarily viewed as obstacles around efficient insights and clarity. With respect to insights, when respondents were asked about their top challenges related to data management, the demand for and challenge around data quality resonated again (31%), as did challenges managing insights or analyses (23%) and issues integrating enterprise-wide solutions (28%).

Data clarity obstacles related more to an access and governance perspective. These challenges are rooted around cataloging for data discovery (26%), proper access control (24%), and visibility (23%)—all of which provide clarity around data management. These segments represent the top 6 responses. Those solutions with capabilities that leverage metadata to automate data profiling and streamline data quality assessment to enhance data quality scoring, correlation visuals, and other data quality measures can help alleviate such challenges. And it is in this manner that solutions that effectively also support data observability can help minimize the bottlenecks felt as a result of data management challenges.

Data observability is an organization's ability to fully understand the health of the data in its systems, enabling data trust and an evolving component of data governance. 46% of all respondents identified understanding the quality of source data as an impediment to efficiently maximizing the use of data in their value chains (see Figure 4).

Figure 4. Data Management Value Chain Bottlenecks



What are the most serious bottlenecks in your organization's data value chain? (Percent of respondents, N=220, multiple responses accepted)

Source: ESG, a division of TechTarget, Inc.

Further challenges around data visibility insights (42%), or lack thereof, often lead to downstream bottlenecks and obscure data observability, accuracy, and effectiveness, as well—all of which are impediments to data observability.

As we've seen, the proliferation of increased data brings the challenge of maintaining the necessary control systems needed to optimize data management by key stakeholders. This is the essence of data empowerment. Vital to its success is the ability to provide data intelligence, expedite the ease of finding data, and understand and use relevant enterprise data assets for faster, smarter decisions. In 2021, the top response found by ESG as to what "data empowerment meant to

respondents" was data analytics (31%). This, however, changed to data management (34%) in 2022, and data analytics dropped substantially in 2022 to17%. This shift underscores the notion that organizations are realizing the importance of first being able to manage data effectively before focusing on its extractions and analysis. The second highest response in both 2021 (30%) and 2002 (31%) was data quality.

Massive Opportunity to Leverage More Data

Challenges and bottlenecks can lead to a serious underusage, or even a lack of awareness, of existing data. This is perhaps made most evident in Figure 5, which focuses on organizations' aggregation of "dark data" (i.e., data that is retained but unused and/or unknown). The estimated mean average of dark data amassed by respondents is 47%. Meanwhile, 42% of all respondents indicated that at least half of their data was dark. More concerning still, a staggering 21% of respondents further indicated that *no less* than 71% of all their data was identified as dark data. Clearly organizations are struggling to uncover and utilize their data at scale, with many in the industry believing the prevalence of dark data within organizations is even higher than respondents' estimates in this research.

Figure 5. Dark Data Details

To the best of your knowledge, what percentage of your organization's data do you believe is "dark data" (digital information that is collected and stored but not being used for business purposes)? (Percent of respondents, N=220)



Source: ESG, a division of TechTarget, Inc.

With such potentially vast amounts of dark data reported by organizations, the good news is that concerted efforts to identify and explain data presents organizations with opportunities to be embraced, such as subsequent cost savings (by reduction of extraneous spending on unused data), increased storage capacity, and increased analytics/compute ability. Solutions featuring capabilities like automated metadata harvesting and data cataloging, data lineage and impact analysis, and integrated asset discovery and social collaboration tools for all data stakeholders can help reduce dark data and, therefore, fuel data governance programs.

Automation Affords Opportunities to Improve Data Governance

As has already been alluded to, automation has significant benefits in assisting data governance programs. Increasing the accuracy and efficiency of IT and data governance teams and lessening organizational impacts of the current skills/staffing gap are premier benefits of engaging automation. Automated product capabilities to effectuate data governance include automated data cataloging, data profiling and data quality assessment, sensitive data classification, and automation that aids self-discovery of data assets wrapped in business context and governance. With only 37% of all respondents describing their level of DataOps as completely or mostly automated, there is still room for many organizations to improve their DataOps efficiencies, while at the same time minimizing many data governance challenges (see Figure 6).

Figure 6. Level of DataOps Automation





Source: ESG, a division of TechTarget, Inc

Data Analytics Challenges

Data governance must balance rigorous rules for usage with curation that promote strategic data usage to be successful. Whether the data is designated for IT or operations performance analysis, cost analysis, customer insights, or inputs necessary for ML and AI (just to name a few), if the data used for any such task is compromised, the analytical outcome likely will be compromised as well.

Respondents indicated that the top challenge associated with their data analytics proved to be, predictably, quality of data (46% of all respondents). While analytics accuracy is made challenging by potentially sub-par data quality, analytics productivity is hindered by limited visibility and management of data, such as end-user experiences (35%), extract and load

(ETL) speeds (34%), and effective cataloging (33%), all of which especially impair data analytics teams' ability to optimize productivity.

Our research revealed that data analytics solutions that increase self-service productivity (26%), as well as revenue growth with correlations to analytics-based decisions (26%), were designated as the top-ranking ROI benchmarks associated with viable analytics programs.

The Recognized Value of DataOps

DataOps is an agile, automated, and process-oriented methodology used by data stakeholders to improve the quality, delivery, velocity, and management of data and analytics. Our research unveiled that 90% of all respondents either strongly agreed or agreed that DataOps efforts are improving data quality, and 90% similarly strongly agreed or agreed that DataOps is alleviating enterprise-wide data challenges (see Figure 7).

Figure 7. Gauging DataOps Benefits' Benchmarks

Please rate your level of agreement with the benchmarks as it relates to your organization's DataOps efforts. (Percent of respondents, N=220)



Strongly agree
Agree

Source: ESG, a division of TechTarget, Inc

By addressing the fundamentals of DataOps efficiencies, organizations can expand data visibility across the enterprise, raise data literacy, more efficiently deliver and operate data pipelines, empower their business users, better delineate data stewardship, and allocate more contextual business meaning and governance to existing data. Current DataOps challenges shared by respondents included 86% of respondents indicating a need to better formalize data quality requirements, a difficulty to determine the proper level of data accessibility and availability based on role (59%), and lacking business context when accessing and/or analyzing business data (56%).

Heavily Used DataOps for Data Engineering Purposes

Equal Distribution of DataOps Tool Deployments

No single DataOps solution format dominated responses. To the contrary, there was generally a three-way split when it came to DataOps solutions deployed:

- 1. Blended variety (35%).
- 2. Do-it-yourself (33%).
- 3. Out-of-the-box (31%).

Data warehouse (45% of all respondents) was the most identified tool used by organizations for data engineering purposes (see Figure 8). Visibility into the vast amount of information housed in data warehouses is supported by resource optimization tools (44%) and catalog/metadata management (42%).

Solutions featuring metadata management tools and automation particularly elevate data asset visibility and processing. Cataloging of metadata makes finding assets faster, while metadata-driven automation expedites delivery of data to appropriate pipelines.

Figure 8. Data Engineering 'Tools of the Trade'

Does your organization employ any of the following tools primarily for data engineering purposes? (Percent of respondents, N=220, multiple responses accepted)



Source: ESG, a division of TechTarget, Inc

Conclusion

The data ecosystem continues to evolve and become more complex. Data is one of the crucial assets in any business and is vital to various end-users, throughout enterprise personas, from executives and line-of-business teams to IT decision makers and staff. When you consider the fact that many organizations are experiencing data quality issues, which are difficult to manage (from accessibility to visibility), and in many cases have significant amounts of data that is dark, the need for robust data governance solutions providing data landscape transparency united with business context and

guidance becomes resoundingly apparent. Suffice to say, if there were but one takeaway from this research, it is the emphatic need to ensure that data quality is not an insurmountable challenge. There are many avenues available to improving data quality under the data governance umbrella of solutions and tools.

In addressing the need for improved and efficient data quality, those solutions and tools featuring automation (i.e., automated data profiling, data quality assessment, ongoing data observability, and data remediation) greatly enhance data quality visibility for a group not limited to merely traditional IT roles, but to all data governance teams and lines of business. This enhanced data quality, therefore, benefits downstream analytical confidence and accuracy for key stakeholders' insights and business decisions.

Incorporating DataOps is the methodology for organizations to improve the quality, delivery, velocity, and management of data and shapes the value of analytics within an organization. By implementing DataOps-based tools, organizations can not only alleviate the pain and delay of accessing data for end-users, but also improve their ability to govern, provision, and manage the appropriate data to be accessed and analyzed, as necessary.

In short, successful data governance efforts are built on an array of tools that can strengthen the processes of data management, data stewardship, and the data quality and discoverability of the data itself. Equally impactful to a successful data governance program is the ability to empower its end-users.

The value in a robust data governance is enterprise-wide. With an emphasis on elevating the quality of data, benefits abound. The efficiency and visibility in which the data can be used operationally is improved. Data access, controls, and security are vaunted. A robust data governance also helps build the foundation for automation and greater analytical capabilities, all of which ensure more stakeholders in the organization are guided and empowered to confidently derive insights and take action.

Research Methodology

This study—fielded between April 28, 2022 and May 9, 2022—included business and IT professionals responsible for and/or familiar with their organizations' data governance strategies, investments, and operations. The respondents in the study were employed at organizations with 1,000+ employees and an annual revenue of \$100M+.

After applying data quality control best practices and screening the remaining completed responses (on several criteria) for data integrity, a final sample of 220 respondents remained. All respondents were provided an incentive to complete the survey in the form of cash awards and/or cash equivalents.

Figures 9 - 12 detail the demographics and firmographics of the respondent base.

Note: Totals in figures and tables throughout this report may not add up to 100% due to rounding.

Respondent Demographics

Figure 9. Respondents by Job Function

Which of the following best describes your current job function? (Percent of respondents, N=220, one response accepted)



Source: ESG, a division of TechTarget, Inc

Figure 10. Respondents by Number of Employees

How many total employees does your company have worldwide? (Percent of respondents, N=220)



Source: ESG, a division of TechTarget, Inc

Figure 11. Respondents by Industry



What is your company's primary industry? (Percent of respondents, N=220)

Source: ESG, a division of TechTarget, Inc

Figure 12. Respondents by Annual Revenue

What is your company's total annual revenue (\$US)? (Percent of respondents, N=220)



Source: ESG, a division of TechTarget, Inc

All product names, logos, brands, and trademarks are the property of their respective owners. Information contained in this publication has been obtained by sources TechTarget, Inc. considers to be reliable but is not warranted by TechTarget, Inc. This publication may contain opinions of TechTarget, Inc., which are subject to change. This publication may include forecasts, projections, and other predictive statements that represent TechTarget, Inc.'s assumptions and expectations in light of currently available information. These forecasts are based on industry trends and involve variables and uncertainties. Consequently, TechTarget, Inc. makes no warranty as to the accuracy of specific forecasts, projections or predictive statements contained herein.

This publication is copyrighted by TechTarget, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of TechTarget, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact Client Relations at <u>cr@esg-global.com</u>.



Enterprise Strategy Group is an integrated technology analysis, research, and strategy firm that provides market intelligence, actionable insight, and go-to-market content services to the global IT community.



www.esg-global.com



508.482.0188

© 2022 TechTarget, Inc. All Rights Reserved.