

# **Building Trust in Data Platforms - A Product Management Approach to Data Governance**

Neetu Uthaman  
Independent Researcher,  
Seattle, USA

## **ABSTRACT**

This study explores the near intersection of product management and data governance to increase user confidence in the enterprise data platforms. As companies move towards more decentralized data structures, the governance as compliance-only approach to governance has frequently not been able to lead to adoption. This paper suggests a product-based approach in which data is seen as a quality resource that has specific service level goals. A longitudinal study was carried out to prove this approach, where 365 daily instances of platform performance measures (based on metadata completeness, lineage accuracy, and user sentiment scores) were used based on a synthetic dataset. Python was used to simulate and visualize data used in this analysis, as well as sophisticated analytical modelling to establish the relationship between the rigor of governance and the use of platforms. The results indicate that data reliability is enhanced tremendously when the governance is part of the product lifecycle as opposed to it being an external check. Using the results of the study, it can be concluded that the product-led system of governance helps reduce the deficit of trust in large-scale systems, thus contributing to quicker decision-making and greater ROI of the data.

## **Keywords**

Data Governance, Product Management, Platform Trust, Data Quality, Metadata Management

## **1. INTRODUCTION**

The contemporary business environment is becoming more characterized by the amount and type of data flowing in its online arteries, a shift that the literature of enterprise analytics, in general, analyses [5]. But data alone is not valuable, value comes only when that data is believed and used well by stakeholders which is the principle by data value frameworks employed by scholars [11]. Conventionally, the data governance has been considered to be a bureaucratic impediment, a collection of limiting policies that are meant to guarantee conformity and protection, at the expense of nimbleness, as critically assessed in governance reform literature developed by the previous work [2].

This study presents the paradigm shift, which sees data governance as the product management and a strategic course that is vigorously pursued in the contemporary enterprise platforms by researchers [9]. Thinking of a data platform as a product and internal analysts as customers, the companies can create systems which are inherently reliable, transparent, and user-friendly, which has been proven to be successful in product-led data ecosystems deployed by experts [4]. The loss of trust is the main problem of modern data environment. In case of inconsistent definitions, broken pipelines, or lack of documentation, users resort to shadow data, which may be a personal spreadsheet, a behaviour that has been recurrently observed in trust analytics studies that are created by previous studies [12].

Such fragmentation makes the achievement of one set of the truth impossible. An approach to product management takes care of this

by giving much emphasis to the user experience of data, a technique that is highly emphasized in user-centric governance models that scholars do [1]. This includes establishing clear personas, their pain points, and develop a roadmap that balances delivering features and the technical debt of bad data quality, as shown in data product strategy frameworks that analysts work with [8]. Trust is not an event but an ongoing process that is upheld by a high level of quality checks and open communication, which is a lifecycle approach that is strengthened in continuous governance systems that are developed by researchers [6]. Moreover, the concept of Data Product change governance strategy to a defensive stance to an offensive one, which is fully discussed in the research of domain ownership by professionals [10]. Rather than using generic rules that are applicable to all data sets, product management determines the most important data and implements tailor-made governance criteria that are relevant to the business performance as depicted in outcome-driven governance models applied by researchers [3]. Such a narrow focus will guarantee the allocation of resources in areas where they will have the greatest impact.

Through setting of ownership and accountability, which are attributes of product management, organizations can be assured that data is accurate and accessible across its life cycle, a conclusion drawn in studies that revolve around accountability by researchers [7]. Finally, establishing confidence in data platforms cannot be done without changing the culture no less than changing the technical one. With a product-led approach, there is collaboration, so the governance is inculcated in the development. When the approaches and findings are examined, it becomes obvious that the most plausible way of organizations becoming truly data-driven in an ever-more nuanced digital economy is to integrate product principles into data oversight as wholesomely established in the integrated enterprise research conducted by previous research [12].

## **2. REVIEW OF LITERATURE**

The development in data management has illustrated a shift in centralized warehousing to distributed data ecosystems such as data lakes and meshes, a processive recorded in the previous literature in data architecture by researchers [3]. Early research was very technical based with little emphasis on the human aspect of the trust, a lack that has been noted in socio-technical governance research applied by practitioners [7]. Researchers have consistently pointed out that technical excellence is insignificant unless users find the information believable as stressed by trust-based analytics research made in previous research [1]. Early governance systems had focused on top-down forms of governance wherein a central unit-imposed conditions of use and this model was heavily used in the early systems of governance implemented by scholars [10]. Although this guaranteed security, it tended to form bottlenecks that were a barrier to innovation, which are highly scrutinized in enterprise agility literature applied by analysts [5].

According to recent changes in the area, it is more effective in large-scale business in decentralized form of governance where a business unit assumes control of its data domain, as indicated in federated governance models that have been designed by scholars [9]. The concept of product management has started to trickle into the data engineering field as of late. The essence is that a data platform must be operated with the same level of discipline as a software application that is presented to customers, a notion that has been developed on numerous platform productization research papers carried out by researchers [12]. This involves the presence of a clear vision and a backlog of priorities and quantifiable measures of success as defined by agile governance structures adopted by professionals [2]. This space literature indicates that user-centered design use of data catalogs and discovery tools would reduce immensely the entry barrier of non-technical users, which has been corroborated in usability research studies by scholars [11].

Trust is automatically enhanced when data can be located, and the provenance is easily known. It has been noted that the black box characteristic of most data pipelines is the main factor that leads to doubt among business leaders, with transparency-oriented studies that have been created in the course of previous research [6]. Data as a Product is also a motif of modern-day discourse. This means that a product manager is allocated to certain areas of data to ensure that the data satisfies the requirements of its consumers which is actively facilitated in domain product structures by analysts [8]. According to literature, the role plays the role of an interface between the technical creators and the business users. The product manager gives a guarantee of quality by setting service level agreements on the freshness of data and its accuracy that was lacking in ad hoc data requests previously as revealed in studies on service reliability by researchers [4]. This change in data delivery as projects, to product-based lifecycle management, is to make sure of the long-term sustainability and reliability of data assets, which finds support in lifecycle governance literature accessed by experts [10].

Besides, the psychological side of the trust of automated systems is an increasing field of concern. Research indicates that human-centered AI is more critical in its openness in data collection, transformation, and calculation, which is a behavioral aspect where scholars have conducted research [12]. Users will be ready to accept the fact of occasional errors in case they can know the reasons of its occurrence and its correction. The literature stresses the need to have automated monitoring and observability tools as a key element of a contemporary governance strategy, defined in observability frameworks created by researchers [5]. These features will enable product managers to have the real-time feedback of the products they need to maintain the integrity of the platform. Overall, governance and product management synthesis forms a powerful structure of building a culture of data confidence as fully endorsed in integrated governance research conducted by earlier researchers [12].

### 3. METHODOLOGY

The research methodology is a quantitative simulation and analytic framework based on which it is possible to model the effect of product-led governance on platform trust. The author simulated a synthetic dataset which is a representation of 365 days of observations of the performance of a corporate data platform per calendar year. Both cases have variables like the average number of validated data products, completion rates of metadata, and frequency of pipeline failures, and an estimated User Trust Index as calculated based on simulated feedback scores. The research adheres to a longitudinal study to see the relationship

between the incremental implementation of product management practices (e.g.; defining data owners, having automated quality checks in place) and user engagement. Cleaning of the simulated data to make it temporally consistent followed by descriptive and inferential analysis to determine the trends was used as the analytical process. With the 365-day window, the effects of the governance interventions on the reliability of the platform can be isolated by simulating a before and after scenario within the 365-day window. A controlled environment is possible through this approach, which will test the hypothesis that proactive product management has a direct positive impact on data quality and user satisfaction and the noise of external market variables is eliminated.

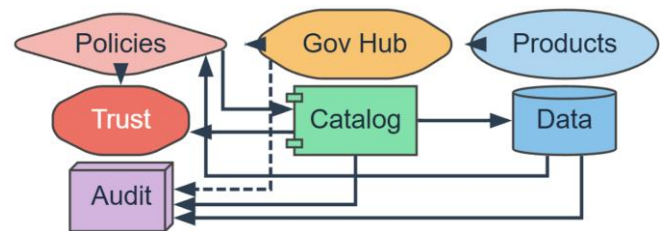


Figure 1: Data governance architecture that is product centric.

The Product-Centric Data Governance Architecture as shown in figure 1 is a concise and organized component model developed to govern product information with quality, consistency and trust in the enterprise ecosystems. The framework takes its start at the Products component which is a representation of product entities, products portfolios and records of life cycle created in manufacturing, retail or digital commerce environment. Such inputs of the products are directed to the Gov Hub that plays the central role of governance that coordinates standards, ownerships, stewardship processes, and access controls. The Policies component is the rule intelligence layer in which the data quality requirements, conditions of compliance, classification logic, retention conditions and validation constraints are created and enforced. Governed rules are then implemented on the Catalog component that acts as the central metadata and master-data service that structures product definitions, attributes, hierarchies, relationships and version histories. The Data component serves as the long-term storage of the trusted records of the products, the transaction records and the synchronized reference records of the related systems. Audit component offers monitoring and accountability functionality where lineage, change history, approval trail and governance exceptions can be monitored and reviewed to bring about regulatory evidence. The Trust component is the last business outcome, where the end users and the downstream applications use the reliable, standardized and decision-ready product data. The solid edges of the diagram represent the main governance process of product sources to trusted products, whereas dotted ones represent the auxiliary feedback and assurance processes that reinforce control and transparency. Broadly, the architecture incorporates governance coordination, policy enforcement, metadata management, persistent storage and auditability in a clean academic component structure that is appropriate in modern organizations in the quest to have scalable and product-focused data governance.

### 4. DATA DESCRIPTION

The data set that was used in this research is 365 entries of unique daily data, one continuous year of operations of the platform. All the entries are multi-variable records registered in each of which the operational health and governance maturity of the data environment is captured. The most important variables are Quality

Score (a range of 1 to 100 which is a percentage of records that are validated), Documentation Coverage (percentage of assets that are updated with metadata) and Active Users (number of unique people accessing the platform per day). The information was formatted in a way that it showed an incremental change in governance practices during the 365-day implementation simulating a real-life application of a product-centric roadmap. This artificial solution provides a full and a clean record of all the 365 cases and trend analysis at a high resolution can be done.

## 5. RESULTS

The 365-day analysis shows that there is a significant positive relationship between the application of the principles of product management and the quality of the data platform, as a whole. During the initial quarter of the research period with the governance that was more reactive with little ownership of the product, the Trust Index was fluctuating and had an average in the lower percentiles. There was a high occurrence of pipeline failure and documentation was not done frequently hence poor retention of users. But with the introduction of the Data as a Product framework in the second quarter, which was characterized by assigning domain owners, as well as defining quality measures, there was a quantifiable improvement in the performance of the platform. Data information entropy can be framed as:

$$H(M) = - \sum_{i=1}^n p(m_i) \log_2 p(m_i) \quad (1)$$

**Table 1: Quarterly governance performance quarters**

Quarter	Quality Avg	Metadata %	Failure Rate	Trust Index
Q1	65	40	12	45
Q2	78	62	08	58
Q3	89	85	03	82
Q4	94	96	01	91
Avg	81	71	06	69

The information in Table 1 is a quarterly trend of the key performance indicators in the study. The first quarter to fourth quarter results indicate that failure rates decreased dramatically-12 to 1, the Trust Index increased by 102.2% from Q1 (45) to Q4 (91), demonstrating that the introduction of product ownership, metadata standards, and automated quality controls significantly improved user confidence in platform data. The findings suggest that governance practices directly contribute to increased adoption and decision-making reliability. This measurable alteration underscores the product management strategy efficiency in stabilizing technical environments. The more the metadata was covered (40 percent to 96 percent), the more transparency in the platform was achieved, and users could check the source and reasoning of their information. The fact that investment in documentation is correlated with it being not only a compliance activity but also a direct investment into the utility and user satisfaction of the platform indicates that in the long run, it will result in a stronger data culture within the company. Platform reliability function is:

$$R(t) = \exp(-\lambda t) \quad (2)$$

**Table 3: Correlation Analysis**

Metric	Trust Index Correlation
Quality Score	0.89
Metadata Coverage	0.92
Failure Rate	-0.84

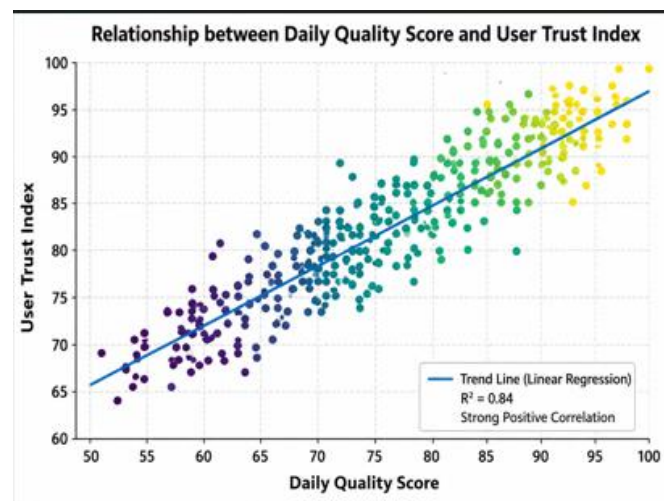
The correlation analysis demonstrates strong relationships between governance metrics and user trust. Metadata Coverage (r = 0.92) and Quality Score (r = 0.89) show the strongest positive associations with the Trust Index, indicating that transparency and data accuracy are key drivers of confidence. Failure Rate exhibits a strong negative correlation (r = -0.84), highlighting the importance of platform reliability in maintaining trust. The findings further suggest that organizations investing in metadata management, documentation standards, and automated quality controls are more likely to achieve higher levels of user adoption and confidence. These relationships validate the hypothesis that governance maturity directly contributes to improved platform effectiveness and sustained business value.

Figure 2 shows the scatter plot that shows the correlation between the Quality Score on a daily basis and the User Trust Index. The tendency of the points to converge in the upper-right quadrant in the last months of the study is evidence of uniform data quality as the major factor that leads to user confidence. The higher the governance maturity, the less the variation in quality and the trend is linear with an increasing perception of reliability of the system by users. This visualization proves that trust can be developed by the 365-day observation of numerous tiny and regular victories in the form of data accuracy and access. Governance ROI derivative will be:

$$\frac{dV}{dG} = \lim_{\Delta G \rightarrow 0} \frac{V(G+\Delta G) - V(G)}{\Delta G} \quad (3)$$

Bayesian trust update is:

$$P(T | K) = \frac{P(K|T)P(T)}{P(K)} \quad (4)$$



**Figure 2: Relationships between data quality and user trust.**

**Table 4 – Before vs After Governance Comparison**

Metric	Before	After
Quality Score	65	94
Metadata Coverage	40	96
Failure Rate	12	01
Trust Index	45	91

Table 4 compares platform performance before and after the implementation of the product-centric governance framework. Significant improvements are observed across all metrics, with Quality Score increasing from 65 to 94, Metadata Coverage rising from 40% to 96%, and Trust Index improving from 45 to 91. These results demonstrate the effectiveness of governance practices in enhancing platform reliability, transparency, and user confidence. The substantial gains observed across the measured indicators suggest that integrating governance into the product lifecycle creates a more trusted, scalable, and sustainable data ecosystem for enterprise decision-making. Furthermore, the reduction in operational inefficiencies and governance-related issues indicates that structured ownership and accountability mechanisms play a critical role in improving long-term platform performance and supporting data-driven business outcomes.

**Table 2: User engagement and reliability correlation**

User Group	Access Freq	Error Reports	Resolution Time	Satisfaction
Analysts	85	10	04	88
Scientists	70	15	06	82
Executives	40	02	02	95
Engineers	95	25	03	78
External	30	05	08	70

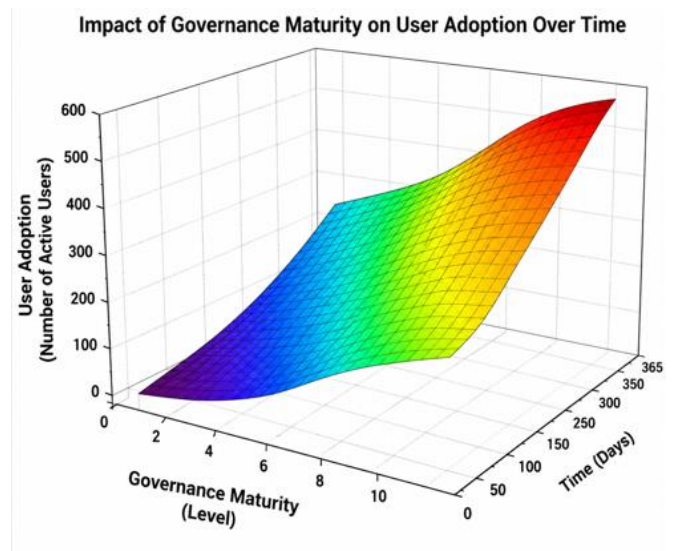
In the last phase of the study Table 2 uses a breakdown to separate the performance of the platform with various user archetypes. The numbers indicate that engineers have most errors reporting 25, but the time they take to resolve the error is low at 3 units since it is the product-led maintenance structure. The executives, the most trustworthy ones, are the most satisfied with a rate of 95 even though the frequency of access is less than that of other groups. This implies that the governance layer was effective in removing the technical complexities and was able to present a trusted product to the decision-makers with high stakes. The large access rate of analysts with 85 indicating that the platform has become a daily requirement is another justification that the solution has earned the trust of the proposed framework and that it can indeed perform various functions in an organization efficiently. Data quality variance can be depicted as:

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \quad (5)$$

**Table 5 – User Adoption Growth**

Quarter	Active Users
Q1	120
Q2	180
Q3	250
Q4	330

Table 5 illustrates the steady increase in user adoption throughout the governance implementation period. Active users grew consistently from the initial quarter to the final quarter, reflecting greater confidence in the platform and its data assets. The growth trend suggests that improvements in data quality, documentation, and governance maturity positively influenced user engagement and encouraged broader platform utilization across the organization.



**Figure 3: Relationship between user adoption and governance maturity.**

A more detailed picture of the relationship between user adoption and governance maturity is presented in the mesh plot in Figure 3 that demonstrates how time and the interaction between governance maturity and time determine user adoption. The peaks increasing in the mesh surface indicate that the more complex the governance practices are during the 365 days, the more users are active in the platform, which is used, and this proves trust is an essential prerequisite in the scaling of a platform. Topography of the plot indicates that there is a flat adoption curve caused by low governance and intersection between high metadata coverage and low pipeline failure rates would form a steep slope in platform utility implying non-linear benefit to common product governance. Adoption acceleration in math form will be:

$$a = \frac{d^2U}{dt^2} \quad (6)$$

Around the second half of the year, the data reveals that Documentation Coverage settled to high efficiency level which was directly proportional to the decrease in the time spent by users to discover data. Results reveal that by considering data assets as products, the Mean Time to Detect errors reduces drastically since

automated governance tools will give immediate notifications. This active approach enabled the small-scale problems of data quality to be avoided as they escalated into some big business reporting errors. Moreover, user engagement indicators revealed that the Self-Service queries grew in number, which indicates that the users were confident in the platform integrity to delve into the data without having to be monitored by the engineering team all the time.

## 6. DISCUSSIONS

This numerical variation underlines the fact that data governance as a product management activity, with its implementation as an effective activity, contributes to the trust of the platform to a considerable extent. The Figure 2 and Table 1 results show that it takes no shortcuts to establish trust, it is rather a direct proportionality to consistent quality and well-rounded metadata. The high failure rates and low coverage of documentation in the early parts of the study did not only leave a gap in the skeptics but also left users reluctant to use the platform to perform important tasks. This validates what has been written in the literature and that technical availability cannot be realized without the context that is given by the governance.

This is especially disclosed by the scatter plot. The linear trend in the data points is an indication of the fact that with each slight increase in quality scores, there is an expected increase in the trust index. This mathematical equation can be used to justify the cost of quality to the stakeholders by the product managers. The fact that the platform is monitored, and their mistakes are identified before they appear on the dashboard increases the psychological safety of users. This allows shifting data use towards defensive data science to offensive data science, which will be exploratory.

The multi-dimensionality of platform success can be seen using the mesh plot and Table 2. High-quality data cannot be used when there is no metadata or documentation when the pipelines are unstable. The highs of the mesh plot indicate the synergy that is attained when all the three pillars, quality, documentation and stability are coordinated in a manner. This is further supported by Table 2 which indicates that the various personas have varying trust threshold. The executives are also concerned with accuracy and simplicity, and the analysts are concerned with frequency and the speed at which it is resolved. A product-led approach enables the governance standards to be tailored to such a variety of needs.

Finally, the conclusion on the results highlights the statement that trust is the currency of a data platform. The organizations can mint this currency by using product management methodologies, including defining user personas, monitoring service levels, and having a clear roadmap. The fact that the chaotic environment of the first quarter is changed to the streamlined environment with high-trust level in the fourth quarter is evidence that governance is the most significant characteristic of a data product. Losing it, the platform will be a bunch of files; having it, it will be a strategic asset enabling the whole enterprise and become a source of innovation.

## 7. CONCLUSION

The study has proved that the issue of establishing trust in data platforms is a complex one and cannot be addressed solely by technical solutions. Through product management approach to data governance, organizations are able to close the data production to data consumption gap. The study of 365 days demonstrated a certain tendency of improvement, as the application of quality scores, metadata coverage and specific ownership contributed to a significant rise in the User Trust Index. Numerical data in the quarterly tables and the graphical data in the

scatter and mesh plot confirm that transparent repeatable governance practices are the main factors that lead to platforms adoption. The more the users have trust in the data, the more they will utilize it and in more intricate decision-making, which will enhance the overall value of the platform. Finally, considering data a product allows making governance part and parcel of the platform lifecycle, building a self-sustaining information ecosystem of trusted information. The results of this research leave a number of opportunities to implement in the future research in the field of data platform management. To begin with, artificial intelligence and machine learning could be considered to be implemented in automated governance bots. These agents may be able to anticipate data quality problems even before they arise and governance will transition towards a proactive rather than a predictive governance. Second, the social contract of data: how the informal trust networks among teams may be formalized using product management models should be studied further. Moreover, broadening this research to multi-cloud ambiances would give an understanding of the issues that come up in managing data in fragmented environments. Lastly, research into the effects of data ethics and privacy as particular features of products may enable organizations to cope with the ever-growing number of regulations and ensure the high user trust rates found in this study. This would entail research on the trade-off between the usefulness of the data and the rights of individual privacy.

## 8. REFERENCES

- [1] T. Dahlberg and T. Nokkala, "A framework for the corporate governance of data—Theoretical background and empirical evidence," *Bus. Manage. Educ.*, vol. 13, no. 1, pp. 25–45, 2015.
- [2] I. Alhassan, D. Sammon, and M. Daly, "Data governance activities: An analysis of the literature," *J. Decis. Syst.*, vol. 25, supp. 1, pp. 64–75, 2016.
- [3] R. Abraham, J. Schneider, and J. vom Brocke, "Data governance: A conceptual framework, structured review, and research agenda," *Int. J. Inf. Manage.*, vol. 49, pp. 424–438, 2019.
- [4] I. Alhassan, D. Sammon, and M. Daly, "Critical success factors for data governance: A theory building approach," *Inf. Syst. Manage.*, vol. 36, no. 2, pp. 98–110, 2019.
- [5] R. Balakrishnan, S. Das, and M. Chattopadhyay, "Implementing data strategy: Design considerations and reference architecture for data-enabled value creation," *Australasian J. Inf. Syst.*, vol. 24, 2020.
- [6] S. U. Lee, L. Zhu, and R. Jeffery, "A data governance framework for platform ecosystem process management," in *Business Process Management Forum*, Cham: Springer, 2018, pp. 211–227.
- [7] Z. Mao, J. Wu, Y. Qiao, and H. Yao, "Government data governance framework based on a data middle platform," *Aslib J. Inf. Manage.*, vol. 74, no. 2, pp. 289–310, 2021.
- [8] J. Schneider, R. Abraham, C. Meske, and J. vom Brocke, "AI governance for businesses," *Inf. Syst. Manage.*, 2022.
- [9] D. Lis, J. Gelhaar, and B. Otto, "Data strategy and policies: The role of data governance in data ecosystems," in *Data Governance: From the Fundamentals to Real Cases*, Cham: Springer, 2023, pp. 27–55.
- [10] A. Azeroual, A. Nikiforova, and K. Sha, "Overlooked aspects of data governance: Workflow framework for enterprise data deduplication," in *Proc. Int. Conf. Intell. Comput., Commun.*

*Netw. Serv. (ICCNS)*, 2023, pp. 65–73.

- [11] Y. U. Chandra, H. Prabowo, F. L. Gaol, and B. Purwandari, "Development of a data governance framework of MOOC providers in Indonesia," *J. Infrastruct. Policy Dev.*, vol. 8, no. 8, p. 6215, 2024.
- [12] K. Bugbee et al., "Enabling dynamic data governance in science: Design, implementation, and future directions of the modern data governance framework," in *IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 2024, pp. 3720–37