Governance Trends for Data Leaders

CHALLENGES, CONSIDERATIONS AND
RECOMMENDATIONS FOR AI, MDM, AND DATA
GOVERNANCE MATURITY EFFORTS

phData | Fivetran

Contents

One of the most fascinating aspects of data governance today is how rapidly it is evolving, especially in the face of emerging technologies like AI. Data is no longer just about storage and access - it's about creating value, ensuring compliance, and managing the ever-growing complexity of how it's used.

To better understand where organizations are heading, we surveyed a diverse group of senior leaders, from VPs of Analytics to CTOs, representing industries such as IT, healthcare, and finance. Their insights reveal how data governance is being shaped in real industry settings.

In this trendbook, we reveal and summarize responses from these leaders to explore the top trends and challenges in data governance. These leaders can help provide us with a glimpse into the future. With the growing impact of disruptive technologies like generative AI on data management, the road ahead is one that will require adaptability, innovation, and a forward-thinking approach. Read on to see how leading organizations are preparing to meet the demands of this dynamic new era.

Impact of Data Governance on GenAI

Data governance plays a critical role in the successful implementation of Generative AI (GenAI) and large language models (LLM), with **86.7% of respondents rating it as highly impactful**. It serves as a vital protective measure, ensuring proper data access while managing risks like data breaches and unauthorized use. Strong data governance also lays the foundation for better model performance, cost efficiency, and improved data quality, which directly contributes to regulatory compliance and more secure AI systems.

Quotes

"Data Governance is the vehicle to run global data standards and strategies to leverage advantages of Gen AI and LLMs" - Chief Data and Analytics Officer, Information Technology Industry

"Data governance is a key to Gen AI and LLM applications, as the basis of data quality, metadata management, and lineage are foundational components for the right implementation of AI applications." – VP or Head of Architecture, Healthcare Industry

"It's extremely important because many of the Gen AI and LLM applications take an unstructured data approach, meaning many of the tools require you to give the tools full access to your data in an unrestricted way and let it crawl and parse it completely." - Chief Technology Officer, Information Technology Industry

"The impact on data governance due to GenAI/LLM is that these technologies can spot trends much faster than humans or other applications. They can identify where risks are and what to avoid." - Chief Information Officer, Legal Industry

"Rigorous Data Governance processes help organizations achieve two things with respect to Gen AI and LLMs. The first is accelerated discovery - strong governance procedures make quality data easier to find, which accelerates Gen AI and LLM Model Development. The second is security. Rigorous governance procedures ensure only the correct people with the correct access to data are able to leverage that data in an LLM. This ensures Data is not leaked outside its boundary of access." - Hugo Lu, Founder, Orchestra

Challenges and Considerations

Balancing data access and protection is essential as GenAI tools require broad access while still adhering to governance policies. Some respondents feel current governance practices may not fully align with the capabilities of LLMs and GenAI, and there's a tension between the need for speed in implementing AI solutions and the time needed for proper governance. Organizations also need a better understanding of how LLMs are trained, especially with external vendors or public cloud environments. In sectors like legal services, safeguarding client data from being used in public apps or external training models is critical.

> ## Quotes
>
> *"Data governance is going to play a large role in what data can go into an LLM. Right now, many external vendors have strict language around which content can go into an LLM."* - VP of Analytics, Finance Industry
>
> *"It will be increasingly important for organizations to understand how LLMs are trained – whether on the company's own data or paired with otwhers. Additionally, how your data might be used across vendor LLMs when used in public cloud scenarios will be important to ensure that valuable IP or PII does not leave your platform."* - Chief Technology Officer, Finance Industry

Recommendations for Data Leaders

1. Align Data and AI Governance: Ensure that data governance strategies are aligned with AI governance to support successful GenAI implementations.

2. Implement Robust Data Controls: Focus on establishing standardized data controls to improve data quality, enhance model performance, and ensure compliance.

3. Educate on GenAI Capabilities: Invest in understanding the full capabilities and potential risks of GenAI and LLM applications.

4. Balance Innovation and Protection: Develop strategies that allow for rapid innovation with GenAI while maintaining strong data protection measures.

5. Monitor and Adapt: Continuously assess the impact of GenAI on data governance practices and be prepared to adapt policies as technologies evolve.

> ## Quotes
>
> *"Data Governance needs to be aligned to AI governance to ensure successful, safe, and secure implementations of Gen AI and LLM systems."* - Head of Architecture, Healthcare Industry
>
> *"It's important to ensure that the right and proper data is being input and used for GenAI and LLMs, as well as ensure that proper data management is in place to adhere to client rules regarding GenAI use. Data governance is the only way to ensure those requirements are met."* - Chief Information Officer, Legal Industry

Survey respondents noted **improved data quality** and **compliance and risk management** as the top two outcomes for organizations with a focus on more standardized data controls when working to implement GenAI and LLM applications.

Data Governance and Modern Data Management

AI and machine learning (AI/ML) applications emerged as the leading trend in data management, significantly shaping organizations' data platform strategies. As data volumes grow, scalable solutions like data mesh and data fabric architectures are becoming more widespread due to their flexibility alongside complex organizational structures.

Additionally, more organizations are continuing to migrate their data to cloud platforms to meet demands for scalability, security, and cost-efficiency, making cloud adoption another important trend in data management.

> Survey respondents indicated **faster time-to-insight, increased scalability and flexibility, and reduced bottlenecks and dependencies on centralized data teams** as the top three effects on the agility of data teams following the implementation of a data mesh architecture.

Quotes

"AI will have a huge impact over the next few years as more organizations embrace it." - Head of IT, Healthcare Industry

"GenAI and LLM will impact data platforms as they need a bigger amount of data to better train the models." - Chief Technology Officer, Finance Industry

"Data fabric architectures with OneLake will become the key foundational architecture for data platform strategy going forward as organizations will need to scale the growing volumes, types, and structure to support the needs of the business." - VP of Architecture, Healthcare Industry

"Moving data to the cloud is generally a move towards centralization, it doesn't have to mean moving towards a singular record of truth or MDM model, but moving data from many proprietary systems into a generalized cloud data solution at least lets individuals in the organization start the process of information to data transition." - Chief Technology Officer, Information Technology Industry

---

> Survey respondents specified **easier risk management** and **more data access to personnel** as the top two benefits organizations can expect from moving data into a cloud platform.

Challenges and Considerations

Implementing data management in the age of AI comes with challenges, such as ensuring proper security and compliance while also needing high-quality, comprehensive data for accurate results. Organizations must balance agility with strong controls and standards, especially when managing data across multiple cloud platforms.

Developing skills is also crucial, as teams need education on new architectures and technologies, along with fostering collaboration across different areas. Additionally, managing costs is a key consideration, particularly in cloud environments where infrastructure expenses can easily escalate if not controlled.

Recommendations for Data Leaders

1. Develop a clear roadmap for AI/ML integration in data management processes
2. Carefully plan cloud migration strategies, addressing security and governance concerns
3. Invest in team education and skill development for new data architectures
4. Implement robust cost management and monitoring for cloud data platforms
5. Focus on improving data quality and completeness to support AI/ML initiatives
6. Enhance metadata tagging and management to improve data discoverability and usability

Data Governance Maturity

For many organizations, data governance is still in the early stages, focusing on defining policies and processes. However, real success requires more than just a plan—it needs strong executive support, clear rules, and regular involvement from stakeholders.

Looking ahead, data governance will focus on real-time management, data privacy, ethics, and easier access to data for everyone. Self-service analytics tools and technologies like data lineage and provenance tracking will also play a key role in making data management more transparent.

Quotes

"I think many organizations practice data governance at a very high level, mostly at the policy and process definitional level. Many organizations that have been a part of that try to institute data governance end up slowing things down due to not thinking through the compliance and auditing pieces fully in order to ensure a lightweight impact is felt throughout the organization, data governance programs can crush productivity through restricting access to relevant information in a timely manner." - Chief Technology Officer, Information Technology Industry

"The organizations are mostly early in their journey with areas like data governance frameworks established; however, the execution of the frameworks, ownership, and understanding of data literacy, especially within the business, is fairly low maturity. The technology for metadata management, data quality management, etc., is fairly advanced. However, implementation in a large complex environment is difficult due to investment challenges and buy-in from the business." - VP of Architecture, Healthcare Industry

"In my experience, too many organizations think of this only in terms of access to data and data systems. Organizations that are leading in this area understand that governance goes well beyond the connection - to the storage, sharing, and guidance on the types of insights generated." - Chief Technology Officer, Finance Industry

"The process looks like leaders within data governance understanding the business they are helping to lead. How data can work for and against that organization -- for example, what can be easily aggregated up to remove PII versus what must be strictly controlled and monitored. Effective data governance is barely noticed in people's usage in the organization for the most part. And only generates friction when there is real reason for data protection/compliance." - Chief Technology Officer, Finance Industry

---

Survey respondents selected **data encryption** as the most observed practice organizations are currently using to maintain data security.

Challenges and Considerations

Organizations are still figuring out how to effectively implement data governance strategies, facing challenges like employee turnover, lack of compliance among business partners, and difficulty in aligning stakeholders.

There's mixed agreement on whether clear patterns exist for identifying and protecting sensitive PII data. While some patterns are in place, they aren't consistently applied, especially in organizations with lower regulatory demands. The complexity of data and the challenges of marking PII during intake add to the difficulty.

To ensure data quality, platforms need consistent, automated processes with continuous testing and validation. Systems should include alerts to flag any changes or anomalies that could affect data integrity.

Quotes

"Quality is a never-ending exercise of testing, validating, and alerting to changes or patterns that can cause your data to be impacted. Security can feel slightly similar, but with proper tooling (like a Vanta, SecureFrame, or Drata), most of your structural or architectural controls can be handled. Compliance is the final, and hardest, mile as ensuring the very access that drives value within your business does not make it too easy to leak data either from lightly controlled websites to oversharing of information." - Chief Technology Officer, Finance Industry

"Quality is a never-ending exercise of testing, validating, and alerting to changes or patterns that can cause your data to be impacted." - Chief Technology Officer, Finance Industry

"Investment in metadata and data lineage systems and a strong data governance and ownership framework agreed with key information asset owners within the business is needed." - Head of Architecture, Healthcare Industry

"Patterns for identifying and protecting sensitive PII data are well defined and architectured, especially within mature cloud platforms, as there has been investment in technologies for a long time on encryption capabilities, data classification, and discovery." - VP of Architecture, Healthcare Industry

"It's getting better, but it's still hard to identify and protect PII data across an organization because it needs to be identified and marked upon intake. The workflows and human compliance for this still have a ways to go." - Chief Information Officer, Legal Industry

"The source of the data is the starting point for everything. It must be a trusted source that has been vetted and randomly tested upon intake. Afterwards, periodic data audits must occur to ensure the accuracy of the data." - Chief Information Officer, Legal Industry

Recommendations for Data Leaders

1. Invest in metadata and data lineage systems
2. Foster a culture of data accountability
3. Focus on data privacy and ethical use
4. Automate security processes where possible
5. Continuously educate users on data security best practices
6. Regularly audit and assess data quality and security measures

Survey respondents marked the **adoption of advanced data governance technologies** as the most anticipated change to how organizations will handle the introduction of more data sources, data types, consumers, and downstream use cases in the coming years.

Quotes

"Data governance assures that proper policy, procedure, personnel, communications, and enforcement is in place to assess the data estate and implement appropriate measures to ensure that data is created, maintained, secured, and disposed according to regulatory, client obligations, and best practices over time." - Chief Information Officer, Legal Industry

"Using industry standard frameworks and models, creating a culture of data accountability, and ensuring data governance is part of an overall data strategy." - VP of IT, Legal Industry

"It requires a software development approach to the release and management of code pipelines. Having a CI/CD process with a good approach of automation allows for fast development with the appropriate governance to ensure compliance and security."
- Adam Fokken, Chief Technology Officer, phData

Conclusion

Data governance is constantly evolving, especially with the rise of generative AI technologies. As businesses adopt these innovations, one thing is certain: change will be a constant in managing data and AI/ML applications in the coming years. Business leaders must be ready to adapt and guide their teams with agility, embracing the ongoing shifts in technology and regulation. The key to success will be staying flexible, proactive, and responsive to these changes.

About phData

phData is a leading data and AI consulting company specializing in AI and data applications, from conception to production, that drive real-world business value. Our global delivery team partners with the world's top brands to execute data initiatives in artificial intelligence, data engineering, applications, analytics, and managed services for cloud platforms.

About Fivetran

Fivetran is the global leader in data movement, helping customers use their data to power everything from AI applications and ML models to predictive analytics and operational workloads. Global brands, including Autodesk, Condé Nast, and Morgan Stanley, trust Fivetran to move their most valuable data to fuel analytics, drive operational efficiencies, and power innovation.

CONTACT PHDATA TODAY