

Introduction

In an Al-obsessed world that depends on data to thrive, traditional data governance is starting to look like a relic of a bygone era. Designed for structured, human-only data consumption, these clunky and rigid models are as efficient as trying to navigate with a paper map in the era of GPS. As we enter a future where humans and Al use data together, stale governance practices put an organization's ability to compete at risk, rendering it unable to make the leap to a future where Al will dominate as the sole consumer of data.

Data governance was once a settled affair, firmly rooted in best practices centered around neatly organized, well-structured data. Then came generative AI. Suddenly, organizations found themselves scrambling to figure out how to govern this technology, leaning on newly formed AI councils to create the rules and police the road. Today, as the speed of AI innovation races ahead of its diffusion across the enterprise, data governance has emerged as a key roadblock to AI progress: 62 percent of organizations believe a lack of data governance is the main data challenge inhibiting AI initiatives.¹

To cope, organizations are turning to modern data governance practices, but they present their own frontier of challenges. Enterprises must wrangle with a taxonomy that's evolved past its old grids and charts, an ontology that now includes AI interpretations, and data quality issues that were inconceivable in simpler times. It's akin to managing a library where the books not only refuse to stay on the shelves but also rewrite themselves each night, and leaders are exhausted from pumping resources into bloated, slow-moving governance programs that yield scant returns. With leaders lacking confidence that their enterprise data is ready to make the leap from human-first to AI-only decisioning, organizations remain stuck on yellow, waiting for the green light to accelerate.

The ideal model integrates AI and data governance under a single governance umbrella. It enables complete transparency; creates enforceable policies and standards; eliminates duplicate data sets; and uses data, analytics, and AI use cases to deliver tangible value.

Imagine a data governance model that's as swift and flexible as the Al it serves—designed not only to keep up with the pace of innovation but also to harness it. In this brave new world, data governance isn't a gatekeeper; it's an enabler. The ideal model integrates Al and data governance under a single governance umbrella. It enables complete transparency; creates enforceable policies and standards; eliminates duplicate data sets; and uses data, analytics, and Al use cases to deliver tangible value. By leveraging modern tools and methods that promote trust, agility, and rigorous privacy standards, organizations can maximize data-driven value creation and turn bureaucracy into a catalyst for innovation.

01

JZ

03

04

05

06

¹ Murugan Anandarajan and Diana Jones, 2025 Outlook: Data Integrity Trends and Insights, Drexel University LeBow's Center for Applied Al and Business Analytics and Precisely, 2024.



Current state: Turning from structured to unstructured data management

As businesses increasingly rely on a combination of data, files, content, analytics, and Algenerated output, KPMG LLP has observed chief data officers shifting their focus from structured to unstructured data management. They are grappling with oversight of vast data sets, some of which contain personally identifiable information, reflecting the challenge of managing data governance while balancing regulatory and ethical considerations. At the same time, their technology teams, already tasked with managing the increase in automation and Al requests, find themselves scrambling to address challenges with transparency, explainability, drift, bias, hallucination, and access control.

Old-style governance prevents the seamless integration of diverse and expanding data sets. It also struggles with the dynamic and iterative nature of Al development cycles, where continuous data flows and feedback loops are crucial. Consider, for example, a financial reporting system that relies on structured data stored in a relational database. This system operates on a periodic cadence, where data is collected, validated, and processed in batch mode to produce consolidated financial statements. The data flows are linear and the processes are well-defined with specific checks and controls, making it a largely static and periodic operation.

Old-style governance prevents the seamless integration of diverse and expanding data sets. It also struggles with the dynamic and iterative nature of Al development cycles, where continuous data flows and feedback loops are crucial.

In the future, human consumers of data and insights will be able to create and interact with their data, reports, and analytics using natural language. This shift will enable faster, predictive decision-making and advance reporting into a strategic, value-creating, always-on capability.

Today's forward-looking data leaders are evolving their governance models. They are accommodating systems that are increasingly real-time, flexible, and intelligent, and pulling from diverse sources and processing information continuously using modular pipelines. They are preparing for a future where data will flow more freely through APIs and AI will be integrated with enterprise applications using Model Context Protocol. AI agents will interact with each other via Agent-to-Agent, a direct communications channel. Automation will be used to verify data and its business context to predict and identify issues early on. Known issues will be fixed automatically, while new or unknown problems will trigger real-time alerts to data stewards and owners, providing them with recommended solutions.

In the future, human consumers of data and insights will be able to create and interact with their data, reports, and analytics using natural language. This shift will enable faster, predictive decision-making and advance reporting into a strategic, value-creating, always-on capability.

01

02

03

04

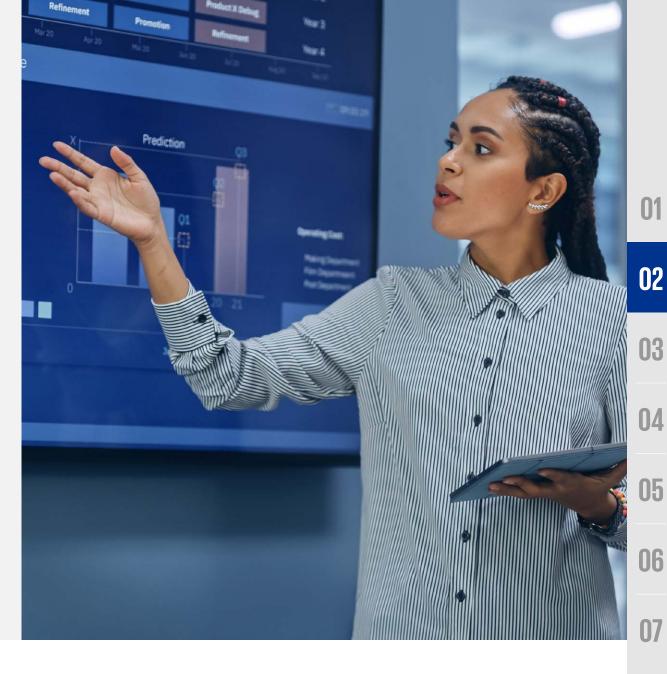
05

06

Common triggers for governance modernization

Balancing diverse data governance needs may be the overarching challenge, but the motivations for modernization vary. The common thread is the need to evolve and adapt faster to match the pace of AI, but the starting point for each organization may be different. The starting points may include:

- **Regulatory compliance:** Companies prioritize this to avoid penalties, ensure audit readiness, and demonstrate transparent data practices, building trust with regulators and customers.
- Reliable and ethical AI: Ensuring AI systems are fair, explainable, and secure mitigates reputational and operational risks, supporting stakeholder confidence and responsible innovation.
- **Use case integration:** Embedding governance in AI and analytics development enables faster innovation while maintaining compliance and protecting sensitive information.
- **Data integrity and transparency:** Investments in high-quality, well-governed data reduce decision-making errors, improve efficiency, and ensure sensitive data is handled ethically.
- **Data product lifecycle:** Managing data as a product ensures consistency and quality, facilitating collaboration between business and technical teams and speeding up data initiatives.
- **Business metadata:** Emphasizing this makes metadata more meaningful to business users and Al systems, enhancing insight generation and domain-specific governance.
- **Technical metadata:** This allows companies to track data flow, identify inefficiencies, resolve issues promptly, and validate systems of record.







On the horizon: Humans and Al govern data together

Unifying data and AI governance on the same roadmap enables organizations to accelerate innovation, reduce risk, and create consistent, transparent oversight across the information that powers AI and the AI systems themselves. In this next paradigm, humans and AI will oversee data governance together. AI will be built and trained to perform prescriptive and specialized tasks, such as reviewing emails for orders and automating purchase and fulfillment, with far greater efficiency. Completing new business requests will become faster, smarter, and more reliable, whether for insights, automation, or innovation.

Imagine Gregory, a data and AI governance manager, collaborating with digital employees—Sarah, Alex, Emma, and Oliver—to optimize operations. Sarah identifies the data needed to support a request, Alex analyzes the data to identify data quality rules that need to be applied, Emma identifies and implements the most efficient way to virtually integrate the data, and Oliver applies predictive logic to the data while providing transparency and evidencing controls and trust-related policies. These digital employees enhance productivity, allowing Gregory to focus on deep work. Gregory's oversight remains crucial, validating AI actions. From a business perspective, this synergy accelerates decision-making, reduces costs, and improves customer support.

Unifying data and AI governance on the same roadmap enables organizations to accelerate innovation, reduce risk, and create consistent, transparent oversight across the information that powers AI and the AI systems themselves.

From an oversight perspective, the human and AI model seamlessly integrates mature data governance practices, robust metadata, and advanced AI capabilities. Consider a mid-sized retail company looking to leverage AI and data governance systems to connect the dots across their operations. The chief executive officer (CEO) sees an opportunity to harness data systematically, aligning various departments toward unified goals and improving performance in areas like customer retention.

To uncover insights into customer churn and its impact on the broader business plan, the company uses a hybrid data governance framework that enables instant discovery of relevant data. When strategic planners request insights on churn, the system immediately accesses a centralized data catalog, ensuring data quality and lineage, and granting secure, role-based access to key stakeholders across marketing, sales, and customer service.

With relevant data in hand, seamless data integration is crucial. The system efficiently gathers data from multiple sources, such as CRM records, supply chain metrics, and sales performance figures, using predefined pipelines and metadata suggestions. This helps ensure all departments have a unified view of the information they need for cohesive decision-making.

Operationally, that assurance shows up in how data is classified and routed for access. To make this model work at scale, the company applies dual data classification by privacy sensitivity and business use. Crucially, this classification is shifted left, closer to the point of data collection, rather than being applied downstream at the point of consumption. By labeling each dataset early for protection and tagging it to its intended use cases and audiences, the organization embeds governance from the start.

01

02

03

04

05

06

17

© 2025 KPMG Advisory, a Belgian BV/SRL and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

This classification becomes the routing layer for access and reuse: RBAC grants least-privilege access by role, while ABAC evaluates attributes (purpose, project, region, sensitivity) to fine-tune decisions at query time. Handling rules such as labeling, retention, and sharing are enforced end-to-end and remain auditable against policy and law. All data activities are tracked and logged, Al models are monitored for bias, and compliance with regulations like GDPR is automatically enforced. This governance helps ensure transparency and accountability, building trust among stakeholders and aligning with corporate ethics.

Al-driven analysis now powers proactive strategy development. Automated insights and interactive dashboards identify key drivers of churn and show their potential impact on overall business performance, while predictive models simulate scenarios to guide strategic decisions. These insights are delivered in natural language summaries, keeping the leadership team informed and aligned. Intelligent processes automate tasks such as demand forecasting and inventory management, integrating effortlessly with existing systems. This allows the company to align operations with customer insights and market trends while reducing delays and costs.

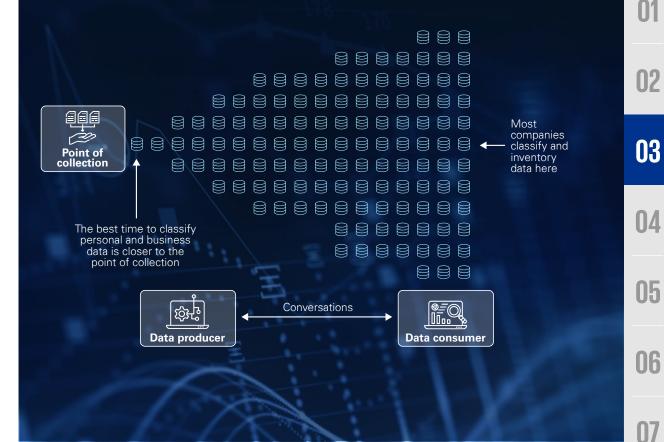
Throughout this transformation, social intelligence and feedback loops enrich the planning process. Teams provide real-time feedback on the Al-generated strategies, enhancing model accuracy and aligning insights with business needs. User-generated feedback refines the planning models, making future recommendations more precise and actionable. Simultaneously, built-in governance and compliance ensure ethical and secure data handling. All data activities are tracked and logged, Al models are monitored for bias, and compliance with regulations like GDPR is automatically enforced. This governance ensures transparency and accountability, building trust among stakeholders and aligning with corporate ethics.

By embedding these governance activities into their integrated business planning process, the company not only uncovers and addresses customer churn but also fosters a data-driven, cohesive approach that synchronizes business functions, enabling it to achieve the CEO's vision of an agile and forward-thinking enterprise.

Shift-left approach to data classification

Data size is smallest at the point of collection and grows through its lifecycle through joins, copies, and analysis.

Rather than working in silos, data producers and data consumers should work together to classify, label, document lineage, and implement data quality earlier in the data lifecycle.



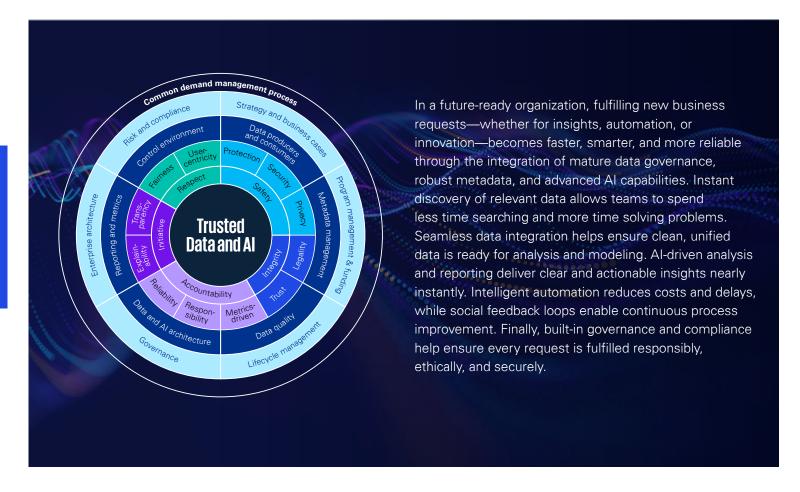
(E)

Beyond the horizon: Autonomous and connected Data + Al governance

In the future, Al-first organizations will be architected around intelligent systems that autonomously manage core functions like customer service, HR, finance, and supply chain, while humans shift into roles focused on supervising Al behavior, shaping strategic direction, and ensuring ethical and compliant decision-making across the enterprise.

Al-first organizations will be architected around intelligent systems that autonomously manage core functions like customer service, HR, finance, and supply chain, while humans shift into roles focused on supervising Al behavior, shaping strategic direction, and ensuring ethical and compliant decision-making across the enterprise.

Consider a fully autonomous supply chain. Al, in the form of a digital employee, logs decisions and explains reasoning in human-readable formats. When a digital employee decides to switch suppliers due to a forecasted shortage of raw materials, it provides detailed logs explaining the decision based on market trends, supplier reliability, and cost-benefit analysis. This transparency ensures that all decisions are traceable and understandable by human stakeholders.



01

02

)3

04

05

06



Al will have a digital chain of custody tracking training data, controlled tests, and the scoring used to validate every iteration of Al-enabled processes. This chain of custody creates accountability, allowing employees to audit Al against defined policies and standards.

Guardrails define what AI can and cannot do without human approval. Critical decisions, such as entering new supplier contracts or making significant changes to production schedules, require human oversight, ensuring that AI autonomy is balanced with human judgment. Model instructions (behavioral guidelines, task-specific rules, domain constraints, compliance and safety boundaries, etc.) are managed centrally and reconsumed every time AI is triggered, allowing for immediate updates. For instance, if a new regulation impacts how data privacy must be handled, then the instructions governing AI are updated centrally and applied across all operations seamlessly.

Every data point carries metadata about its origin, quality, and usage rights, enabling traceability and trust. For example, when forecasting demand, the AI system uses data tagged with metadata indicating its source, ensuring that decisions are based on verified and high-quality information. Governance responsibilities are distributed across AI teams, with centralized oversight ensuring consistency and compliance. Each team managing a segment of the supply chain has governance protocols tailored to their specific operations, while a central team ensures overall compliance and consistency.

Governance is embedded into every data interaction, automatically tagging, classifying, and enforcing policies as data is created or accessed. Known patterns are seamlessly governed whereas new patterns are flagged for reinforced learning.

Al systems monitor data flows continuously, flagging potential compliance issues and remediating them before they escalate. For instance, if data from a supplier indicates a potential breach of environmental regulations, then the Al system flags this in real time and initiates corrective actions. Governance is embedded into every data interaction, automatically tagging, classifying, and enforcing policies as data is created or accessed. Known patterns are seamlessly governed whereas new patterns are flagged for reinforced learning.

Al assesses and enhances data quality in real time, using ML-driven anomaly detection, autocleansing, and feedback loops. This ensures that the data used for inventory management and demand forecasting is always of the highest quality, reducing errors and inefficiencies. Governance oversight remains human-involved, but Al continuously analyzes evolving Alenabled processes and data usage to recommend policy and standard updates. For example, if Al identifies a new risk in the supply chain, then it can suggest updates to governance policies and standards to mitigate this risk.

02

03

04

05

06

The integrated Data + Al governance model

The following framework outlines how integrated Data + Al Governance manifests across core governance dimensions—demonstrating how organizations can operate efficiently, ethically, and transparently by aligning autonomous Al capabilities with human oversight, policy enforcement, and strategic control. To support the shift toward Al-first business models, governance must evolve in tandem across both data and Al domains.

Al discipline			
Focus area	Traditional governance	Human + Al governance	Auto
Transparency	Basic model documentation	Explainable AI integrated into workflows	Real-time, user-facing transparency with traceable decisions
Accountability	Human oversight and audit logs	Shared accountability between humans and Al agents	Al systems self-report and justify decisions with audit trails
Ethics boards	Periodic reviews by ethics committees	Embedded ethics checks in Al development lifecycle	Continuous, autonomous ethical reasoning and escalation
Autonomy limits	Hard-coded constraints	Dynamic guardrails based on context and risk	Self-regulating autonomy with override mechanisms
Instruction management	Manual prompt and rule updates	Al agents learn from feedback and adapt instructions	Autonomous instruction evolution with humanaligned goals
Data discipline			
Data governance	Manual tracking and metadata	Automated lineage tracking with AI tagging	Self-maintaining, wide-ranging data traceability
Federated governance	Centralized control with limited delegation	Domain-level governance with AI policy enforcement	Fully distributed, Al-coordinated governance across domains
Real-time compliance	Periodic audits and checks	Continuous monitoring with Al alerts	Autonomous compliance enforcement and reporting
Ambient data governance	Governance as a separate process	Embedded governance in data workflows	Invisible, always-on governance integrated into all data activities
Data quality	Manual checks and rules	Al-driven anomaly detection and correction	Self-healing data systems with predictive quality assurance
Policy and standards	Static documentation	Al-assisted policy enforcement and updates	Dynamic, Al-evolving standards aligned with business goals



Key recommendations

As Al reshapes the landscape, navigating the data highway requires more than following the old governance roadmap; it demands a rethinking of the journey itself. The integrated Data + Al governance framework isn't meant to deliver full autonomy but can instead equip businesses for a future of data governance maturity, where every organization charts its own course based on its starting position and strategic priorities.

Using a maturity lens, businesses can assess where they are on their journey—in manual, augmented, or autonomous governance. Rather than prescribing a single approach, this framework adapts to the varied needs of each business. helping them progress toward increasingly autonomous governance models. Recognizing that governance should precede individual Al initiatives helps ensure consistent value estimation and prioritization throughout the enterprise. Ultimately, governance becomes a centralized capability able to support multiple efforts as organizations expand their Al strategies.

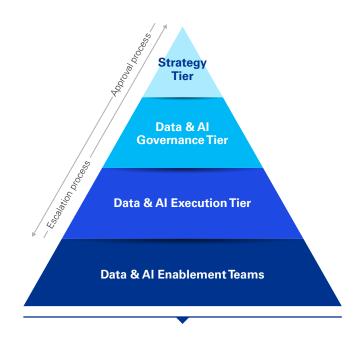
Think of these "big rock" priorities as the directional signposts that help enable Al's full potential, while ensuring governance and compliance are in the front seat for the ride.

Put the C-suite in the driver's seat

Ensuring the C-suite is deeply involved is vital for organizational change. Unlike traditional governance models where the C-suite's involvement can be reduced. placing data quality and risk management at the heart of Al success makes this a strategic, C-suite-based decision. An integrated governance model optimizes customer experience, operations, and research. Elevating governance to the C-suite mitigates risks and maximizes returns.

Key actions:

- Design and activate a Data + Al Governance Committee and charter that aligns with business strategy. Organizations should establish a Steering Committee to oversee Data + Al, including key stakeholders from privacy, data, and legal, and create foundational guidelines for responsible Al use, ensuring alignment with the organization's technology infrastructure and risk tolerance.
- Identify critical business processes that support business strategy goals. These business processes should then be decomposed to identify opportunities for technology enablement and automation. This approach helps to keep use case investments squarely focused on what will drive the greatest measurable value.



An interconnected governance structure enhances visibility and accountability, promoting ethical, transparent, and compliant use of data and Al. Topdown leadership involvement helps ensure alignment of data and AI use cases to enterprise strategy and cross-functional business goals.

 $\mathsf{N}4$

05

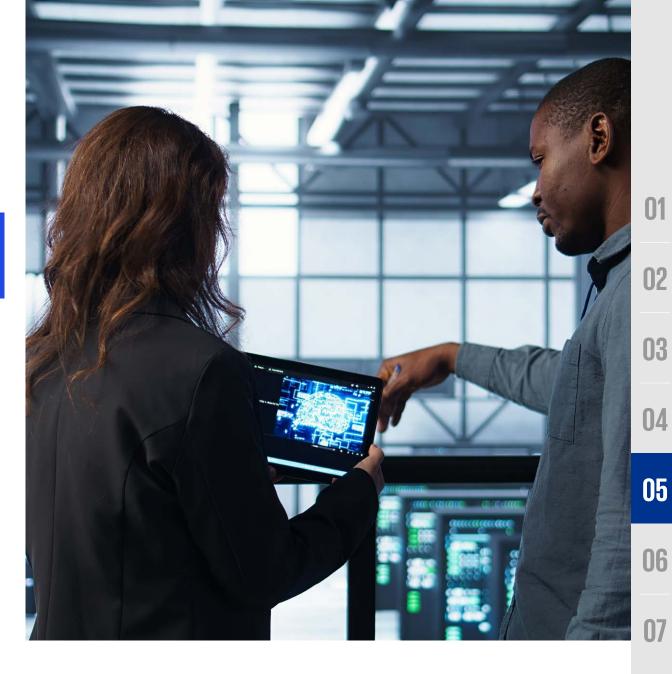
2 Enable federated governance

Federated governance balances centralized oversight with decentralized execution. Initially, a more centralized governance structure may be necessary to establish standards and protocols. Over time, this may evolve into a federated model where decentralized units maintain their autonomy while adhering to a cohesive governance framework, promoting greater flexibility and responsiveness across the organization.

Federated governance balances centralized oversight with decentralized execution. Initially, a more centralized governance structure may be necessary to establish standards and protocols.

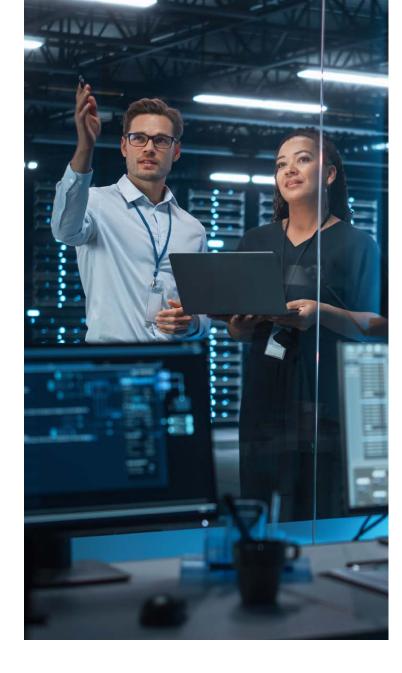
Key actions:

- Define policies and standards across the Data + Al governance dimensions.
 Establishing clear policies and standards ensures that data management and Al practices are consistent and aligned with organizational goals. This helps in maintaining compliance, reducing risks, and fostering trust among stakeholders.
- Design detailed role descriptions that align personnel with accountabilities. A well-defined RACI chart clarifies roles and responsibilities, ensuring that everyone knows their specific duties and accountabilities. This improves collaboration, accountability, and efficiency within the organization.
- Design a demand management process that incorporates data, reporting, analytics, Al, and automation. Developing a structured demand management process helps in prioritizing and managing data and Al-related requests effectively. It ensures that resources are allocated efficiently and organizational needs are met in a timely manner, enhancing operational effectiveness.



© 2025 KPMG Advisory, a Belgian BV/SRL and a member firm of the KPMG global organization of independent member firms affiliated

with KPMG International Limited, a private English company limited by guarantee. All rights reserved.



Recognize—and elevate—the role of metadata

Organizations are swimming in an ocean, rather than a lake, of data. Implementing and driving the adoption of a data catalog and marketplace will streamline the discovery and usage of data assets while allowing for consistent and appropriate usage of domains and data products. Robust metadata also helps ensure that data is secured and controlled through alignment with role-based access control (RBAC) and that sufficient attribute level lineage is available to support compliance and regulatory requirements.

Key actions:

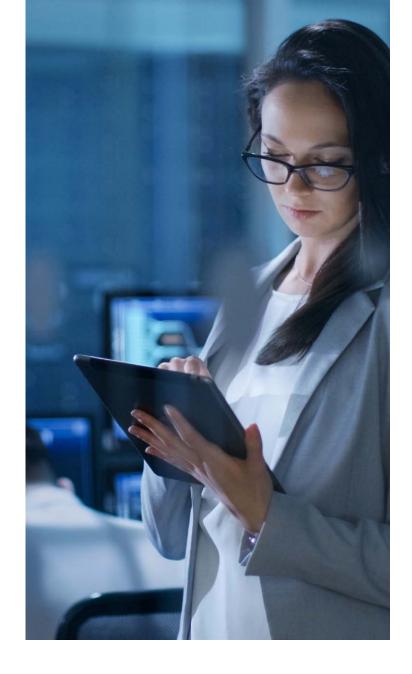
- Select data catalog. Focus on selecting tooling that addresses the greatest number of both functional and nonfunctional requirements and is best aligned with your architecture. This tool will become the center of your data universe by providing one-stop insights into data sources, field definitions, technical transformation logic, usage, and other information needed to enable your organization.
- Perform lineage harvesting from right-to-left (key reports/models back to source). Trace data origins and transformations
 backward from the end-use applications like reports and models to the original data sources. This approach ensures
 that organizations can fully understand where data resides, where it is used, and how it is transformed. It supports
 compliance and regulatory requirements by providing a clear audit trail and helps in maintaining robust data governance
 practices through detailed attribute-level lineage.
- Onboard key platforms and systems into the data catalog. By onboarding these platforms and systems, organizations can centralize their metadata, creating a unified view of data assets. This integration facilitates better data governance, enhances the discoverability and usability of data, and ensures that policies and controls are consistently applied across all data assets.
- Assign system owners and data stewards (domain focus) to enrich metadata, system-of-record (SOR) analysis
 including master data management (MDM), and use-case data alignment. This involves detailed analysis of SOR
 and MDM to ensure accurate and consistent data representation across the organization. By aligning metadata with
 specific use cases, these roles help in tailoring data assets to meet the needs of different business functions, thereby
 improving data discoverability, governance, and compliance.



Successfully integrating data across decentralized systems is vital for creating a robust and versatile data-Al governance framework. This integration supports both human and Al agents, helping ensure that diverse data types from various sources are effectively managed. Serving both human and Al organizations can enhance data accessibility, usability, and decision-making processes, delivering a comprehensive approach to governance.

Key actions:

- Perform sensitive data scanning and classification across the data estate (structured and unstructured). This helps
 identify and protect sensitive information, ensuring compliance with privacy and security regulations, while mitigating
 risks associated with data breaches and unauthorized access. It also enables organizations to manage data more
 effectively by categorizing it according to sensitivity and usage, which enhances overall data governance and
 operational efficiency.
- Extend data taxonomy to include business classifications. This helps organizations systematically categorize and manage data in a way that aligns with business needs and objectives, including strategic initiatives, operational efficiency, and compliance requirements. It also enhances the organization's ability to retrieve, analyze, and leverage data for decision-making, ultimately driving business value.
- Perform business data scanning and classification. This allows organizations to systematically identify, categorize, and
 manage data based on its relevance and importance to business processes. This action enhances data quality and
 reliability while also making it easier to retrieve, analyze, and leverage for decision-making. It also supports compliance
 with regulatory requirements and improves overall data governance by ensuring that data is handled consistently and
 securely.
- Align sensitive and business data classifications with RBAC, which allows organizations to define access permissions
 based on the roles within the company, ensuring that employees, applications, and Al agents can only access data
 relevant to their position and responsibilities. Properly classifying data according to its sensitivity and business
 relevance ensures that the most critical information is protected and accessible only to those with a legitimate need,
 thereby reducing the risk of data breaches and misuse.



UI

02

03

04

05

06

How KPMG can help

Imagine the art of the possible

In an increasingly competitive landscape, harnessing the power of your data unlocks new business possibilities, decreases risk, improves efficiencies, and drives growth. However, to do so requires data that is relevant, accurate, and in compliance with applicable regulations. Collaborate with KPMG to help lead your data governance journey. We have the skills and tools to implement a framework that is guided by leading practices and tailored to your business needs.

You can win with Al

- KPMG named a "Leader" in Worldwide Data Modernization Services IDC MarketScape:
 Worldwide Data Modernization Services 2024 Vendor Assessment
- KPMG is ranked #1 for quality Al advice and implementation in the US



01

02

03

)4

05

06

Authors



Matteo Colombo *Principal, Advisory*

Matteo Colombo serves as the Global Leader for Digital Technologies at KPMG. In this pivotal role, he spearheads the growth, transformation, and innovation across the KPMG Digital Foundation and the global Centers of Excellence for Cloud, Data, and Al. Matteo plays an instrumental role in shaping the use of technology in the firm's collective strategy and ensuring its effective implementation in local markets. Additionally, he manages relationships with key strategic technology partners. Matteo's profound expertise encompasses both strategy and technology, with a particular emphasis on emerging trends that drive growth and innovation for clients and partners. He is a staunch advocate for the effective and ethical use of Al. With over twenty years of experience, Matteo has adeptly guided large enterprises in harnessing cutting-edge technologies to achieve significant, large-scale transformations. As a seasoned technologist, Matteo excels in advisory roles for large and strategic clients, particularly within the Technology, Telecommunications, and Capital Markets industries. In his current position, he is also responsible for spearheading Data Modernization efforts within the US market.



Garrett Flynn *Principal, Advisory*

Garrett Flynn is a seasoned data and technology leader at KPMG US, specializing in the practical integration of modern data technologies to enhance business value and operational efficiency. Garrett leads the Trusted Data offering with a forward-thinking approach and is dedicated to helping organizations evolve incrementally while ensuring they remain competitive in an ever-changing landscape. His key focus areas include AI, data governance, and hybrid data architecture optimization. He also has extensive experience with master data management, data strategies, data migrations for transformation programs, technical and business metadata, as well as the establishment of data catalogs and marketplaces. Furthermore, Garrett drives the integration of data with Trusted AI, privacy, security, and cyber to create a robust and resilient data environment for businesses that also supports value realization through the use of Al. His commitment to leveraging data as a strategic asset positions him as a vital resource for organizations looking to harness the power of AI, data, and analytics effectively.

Our contributors:

The authors thank Lisa Bigelow, Daniel Cohen, Whitney La Bounty, Mary Lewis, Michael Ogden, Rich Schulte, and Lara Volpe for their contributions to this paper.

01

02

03

 n_{4}

05

06

