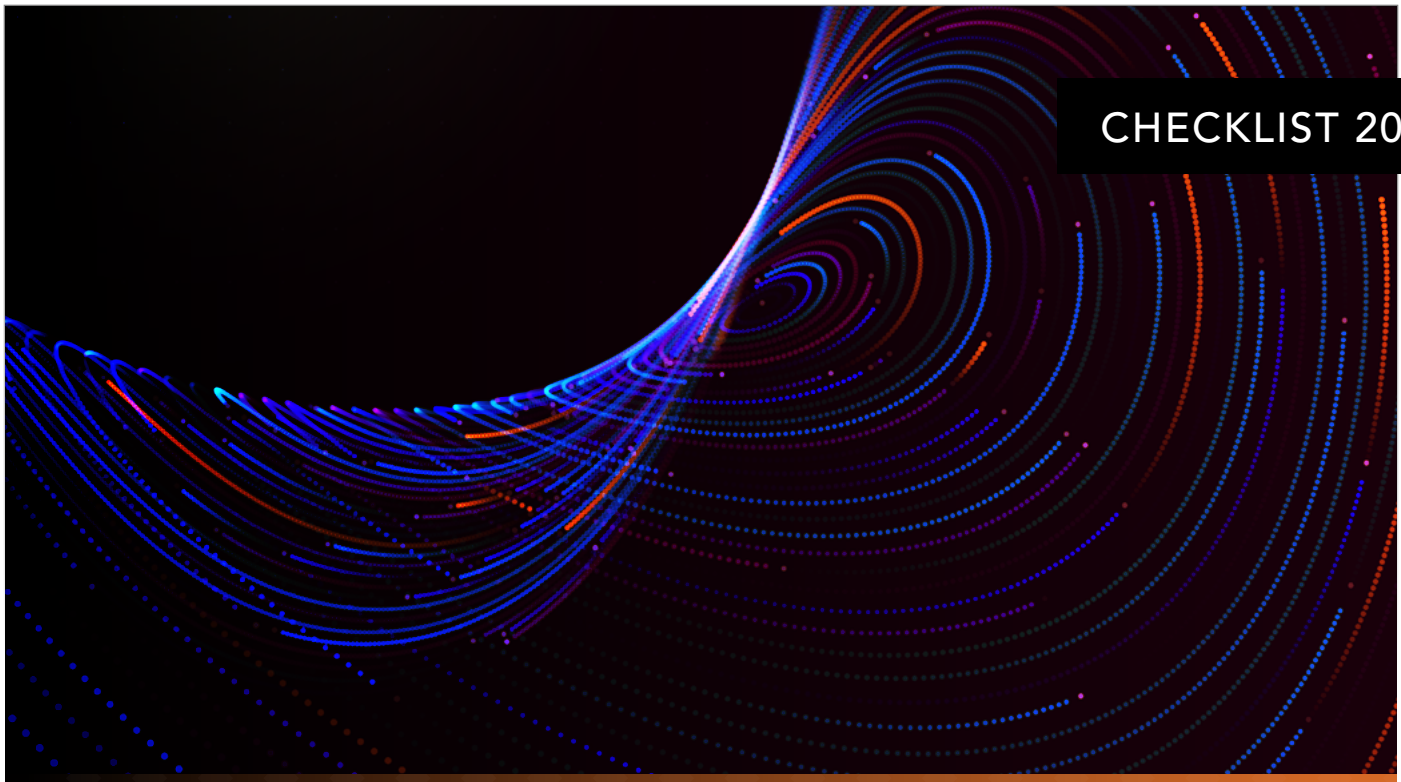


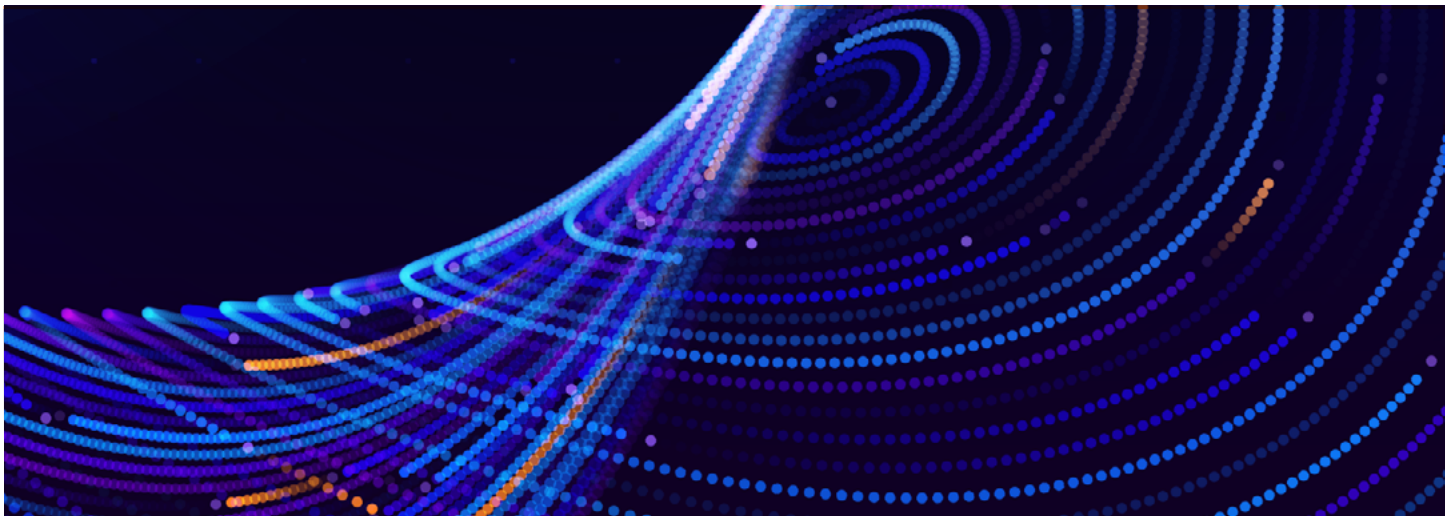
CHECKLIST 2025

Building Trusted Data and AI Governance in a Regulated World

By Fern Halper, Ph.D.

Informatica®

tdwi TRANSFORMING DATA WITH INTELLIGENCE™



Building Trusted Data and AI Governance in a Regulated World

By Fern Halper, Ph.D.

Since the introduction of GDPR, the regulatory landscape has transformed dramatically—driven by geopolitical shifts, technological acceleration, and growing concerns about data privacy, security, ethics, and transparency. From the European Union's AI Act to the Digital Operations Resilience Act (DORA) in financial services, a wave of regulations is reshaping how organizations collect, use, and manage data. For highly regulated industries such as pharmaceuticals and financial services, which already operate under strict compliance frameworks, these evolving regulations underscore the growing need to ensure data is accurate, traceable, and trusted across the enterprise.

In this environment, data governance becomes increasingly important for organizations wanting to manage their data more effectively. TDWI surveys consistently reveal that data governance is a top priority for companies, underscoring its critical role in building trusted data for successful data management and analytics.

Seven steps to establish trustworthy data and AI governance today:

1. Understand the current regulatory landscape and what's coming next
2. Actively address data quality
3. Manage metadata and data lineage
4. Prioritize privacy and security
5. Govern data throughout the AI life cycle
6. Investigate AI governance in addition to data governance
7. Don't forget about responsible governance

Yet data governance is a journey, and that journey has become more complex for several reasons. One is the fact that organizations are moving to the cloud, hybrid cloud, and multi-cloud environments to support their new initiatives around AI and other innovations. This complicates data governance by distributing data across diverse platforms and infrastructures, making it more challenging to ensure consistent policies, visibility, and control. Additionally, these initiatives often require new data types, such as unstructured data. In this environment, data for AI must be trustworthy, transparent, ethical, protected and private, safe, and performant. This requires expanded controls to ensure trust, transparency, compliance, and accountability across diverse and less predictable formats.

On top of this, with the pressure on companies to implement AI and now generative AI, organizations will need to move beyond data governance to model and AI application governance. Models will need to be versioned, documented, and controlled. There will need to be guardrails deployed in the organization as more users consume AI applications. As organizations build their own AI applications with their own company data, these will need to be governed as well.

The regulations mentioned above require transparency, accountability, and ethical oversight—requirements that generative AI, in particular, makes more complex due to its opaque decision-making, data dependency, and potential for bias, misinformation, and hallucinations. In this environment, compliance is no longer just a legal obligation—it's a strategic imperative for organizations to deliver trusted, well-governed data. Failure to meet these expectations introduces significant business and reputational risk, as noncompliance can lead to regulatory penalties, erosion of customer trust, investor support, and long-term damage to brand credibility.

This Checklist offers practical guidance for navigating these challenges. It outlines steps organizations can take to understand and respond to emerging global regulations, establish a foundation of trusted data across distributed environments, and expand governance to include AI and generative AI systems.

1 Understand the current regulatory landscape and what's coming next

For organizations to succeed, they need to stay current with evolving global and industry-specific regulations. Since GDPR was introduced in 2018, there have been a number of regulations and legislation introduced globally that impact both data and AI. To support responsible innovation, organizations must navigate an evolving array of global and industry-specific regulations that directly impact how data is governed, managed, and trusted across the enterprise. These include:

- **The EU AI Act.** The EU AI Act, passed in late 2023, gained final approval in March 2024. It is a global benchmark and applies to organizations operating both within and outside the EU whose systems impact EU citizens. The EU AI Act includes four risk levels (from minimal to unacceptable) associated with AI systems in terms of people's rights, safety, or security. It includes obligations for data governance, human oversight, and documentation. High-risk systems include generative AI systems with foundation models, such as chatbots.[1] The EU AI Act requires risk management, recordkeeping, and explainability, particularly for high-risk AI. For instance, chatbots need to be identified as such. Fines for violating

[1] https://time.com/6903563/eu-ai-act-law-aritificial-intelligence-passes

the EU AI Act can range from 35 million euro or 7% of global turnover to 7.5 million or 1.5% of turnover.[2] Additionally, failure to comply will also result in organizations likely needing to spend time, additional resources, and money correcting errors in their systems.

- **Digital Operations Resilience Act (DORA).** DORA is another EU regulation aimed at financial institutions that went into effect in January of 2023 and is now fully applicable as of January 2025. It addresses safeguarding information and communication technology (ICT) risks. DORA introduces requirements across five key areas: ICT risk management, incident reporting, digital operational resilience testing, ICT third-party risk management, and information sharing.[3] It includes obligations for data governance because it mandates that financial institutions maintain complete visibility, control, and accountability over their data. DORA impacts how data and models are governed in analytics systems that support financial decision-making. Like the EU AI Act, it can also apply to organizations outside of the EU.

- **California Privacy Rights Act (CPRA).** This U.S. state-level regulation, implemented in November 2020, expanded the California Consumer Protection Act (CCPA). It includes many of the ideas from GDPR provisions about automated decision-making, profiling, and algorithmic transparency. The CPRA requires organizations to disclose how automated decisions are made, with rights to opt out (Section 1798.185(a)(16)).[4] Other states have followed suit with their own privacy laws, including the Colorado Privacy Act in 2023 and the Virginia Consumer Data Protection Act (VCDPA), with others on the way.

- **Canada's Bill C-27 (includes the Artificial Intelligence and Data Act – AIDA).** Although this bill died in committee by 2025 due to concerns by some about its exclusionary public consultation process, its vague scope and requirements, and its lack of independent regulatory oversight (and proposed changes did not address concerns),[5] Canada has other regulations in effect such as its 2019 Directive on Automated Decision-Making. The Directive applies to any system that assists or replaces human decision-making in programs and services that affect legal rights, entitlements, or benefits. It includes provisions about data quality, bias mitigation, transparency and explainability.[6]

- **The EU Corporate Sustainability Reporting Directive (CSRD).** Beginning in phases starting in 2024, EU regulations require companies—including EU-based firms and non-EU subsidiaries operating in the region—to disclose their environmental and social impacts. This includes reporting on environmental criteria such as pollution, biodiversity, and resource use, as well as social standards. Companies must also provide data on how their products and services affect end users, including aspects like product safety,

[2] https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai

[3] https://www.digital-operational-resilience-act.com

[4] https://codes.findlaw.com/ca/civil-code/civ-sect-1798-185/#:~:text=(16)%20Issuing%20regulations%20governing%20access,processes%2C%20as%20well%20as%20a

[5] For more about this bill, see https://srinstitute.utoronto.ca/news/whats-next-for-aida and https://montrealethics.ai/the-death-of-canadas-artificial-intelligence-and-data-act-what-happened-and-whats-next-for-ai-regulation-in-canada/#:~:text=Though%20the%20AIDA%20aimed%20to,independent%20regulatory%20enforcement%20and%20oversight

[6] https://www.canada.ca/en/treasury-board-secretariat/services/access-information-privacy/directive-automated-decision-making.html

privacy, and accessibility.[7] The Directive includes the concept of double materiality, which requires businesses to assess the impact of their activities on the environment and society as well as how sustainability affects their financial performance. To comply, companies will need to manage and report on more detailed data and new data, thereby impacting governance.

- **The enhancement and standardization of climate-related disclosures for investors.** In 2022, the Security and Exchange Commission proposed standardizing reporting, including Scope 1 and 2 greenhouse gas (GHG) reporting. Scope 3 would be required only if relevant or if the company has set a Scope 3 target. The SEC is expected to issue a final rule in 2025.

- **Basel Committee on Banking Supervision's standard 239: Principles for effective risk data aggregation and risk reporting (BCBS239).** The European Central Bank Risk Data Aggregation and Risk Reporting (RDARR) Guide outlines more detailed and prescriptive requirements for effective risk data aggregation and reporting frameworks in banking from May 2024.[8] Included is a specific focus on data governance, including enhanced data quality and reporting.

Across jurisdictions, these regulations share a common emphasis on transparency, accountability, and risk management in the use of trusted data and responsible AI. They increasingly require organizations to implement robust data governance frameworks that ensure data is accurate, traceable, ethically used, and compliant with evolving standards for privacy, security, and algorithmic oversight.

[7] https://service.betterregulation.com/document/705505

[8] https://www.bankingsupervision.europa.eu/press/supervisory-newsletters/newsletter/2025/html/ssm.nl250219.en.html#:~:text=Moreover%2C%20remediation%20of%20RDARR%20deficiencies,complex%20economic%20and%20financial%20landscape

To meet compliance obligations, it will be important to review the regulations to understand how they impact your company. Additionally, some companies are building a cross-functional compliance intelligence process. This may leverage a combination of technical, operational, and governance tools to ensure that organizations stay aligned with internal policies and external regulations. For instance, some tools help manage data policies, ownership, and traceability. Others manage how personal or sensitive data is collected, stored, and used. Still others provide visibility into compliance status and help track conformance to internal controls and external standards.

2 Actively address data quality

Data quality refers to the overall suitability of a data set to serve its intended purpose. It's a measure of how well the data meets the requirements and expectations of its users, particularly in terms of accuracy, completeness, reliability, and relevance. Poor data quality undermines the effectiveness of analytics and AI by producing inaccurate, incomplete, or inconsistent outputs—ultimately eroding trust in the results. When data cannot be relied upon to deliver meaningful insights, organizations begin to question the value of their analytics and AI investments. TDWI research consistently shows that many organizations remain dissatisfied with the quality of their data, and this issue is even more pronounced when dealing with unstructured data, which often lacks standardized formats or governance controls.

As regulatory expectations around transparency, fairness, and accountability grow—especially in

areas including AI governance and sustainability reporting—ensuring high-quality, trusted data is no longer optional; it's a critical requirement for achieving and demonstrating regulatory compliance. For example, GDPR Article 5(1)(d): the Accuracy Principle states that "data must be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy')."[9] The EU AI Act Article 10 discusses the importance of data quality with regard to training data for AI models and the potential for bias. CSRD also calls for data quality in ESG reporting.

Modern data quality management includes scaling data quality while keeping up with evolving regulations. This involves policies as well as technologies, such as automated data quality rules using AI-infused automation, given the scale of data needed for modern analytics such as AI.

As a first step, it will be important to determine data quality requirements. This means reviewing the regulations to specify certain data quality expectations. Organizations will also need to define data quality rules. These include format specifications, limitations to data values, and computation rules determining how values are calculated as well as timeliness, consistency, accuracy, completeness, and timeliness rules. With the introduction of newer data types, the meaning of some of these metrics may be different. There may also be new issues to check for, such as the plausibility of unstructured data. Tools must be able to extend to meet these requirements and provide custom rules for specific circumstances for sensitive and unstructured data.

Additionally, it will be important to implement data quality controls. In addition to the above metrics, these include procedures for data validation. Organizations are beginning to implement monitoring and observability tools. Data monitoring is the proactive process of reviewing and evaluating data and its quality using software that measures and tracks performance through dashboards, alerts, and reports to ensure the data is fit for purpose. Data observability ensures comprehensive and holistic visibility into the health of data systems using advanced analytics such as machine learning to detect anomalies, analyze root causes, and track data lineage. Data observability supports robust data governance frameworks by providing insights into data usage and quality. This helps organizations meet regulatory requirements for data management and reporting. With real-time monitoring and alerts, data observability enables proactive identification and resolution of compliance issues.

In TDWI research, "achievers" (those respondents who achieve significantly higher value from their data) are more likely to utilize advanced tools such as modern automated data quality tools and data observability tools.[10] These tools facilitate better data management by providing comprehensive visibility, automated data profiling, and real-time monitoring capabilities.

3 Manage metadata and data lineage

Like data quality, metadata management and data lineage are critical for trusted data as well as important for regulatory compliance. Metadata provides context about data—such as its origin, usage, and definitions—which helps organizations assess

[9] https://gdpr-text.com/read/article-5/

[10] TDWI 2024 *Best Practices Report: Data Monitoring, Management, and Observability*, online at https://tdwi.org/bpreports

its reliability and relevance for decision-making, fostering trust. For compliance, metadata supports auditability and accountability, making it easier to demonstrate adherence to regulatory requirements. Data lineage describes where data originated and how it has been transformed, consumed, and shared. This is obviously important for building trust—think how crucial it is to know if data has been altered in any way before analyzing it to make a business decision. In TDWI research, only about a third of respondents claim they have robust data management tools for metadata, data lineage, and data catalogs.[11]

Metadata and data lineage are important for compliance—in fact they are called out in some regulations. For instance, CPRA grants consumers rights over their personal data, including the right to know what data is collected and how it's used. To comply, businesses must have robust data governance frameworks that include metadata management to catalog data assets and data lineage to trace data origins and movements. DORA requires comprehensive data governance practices, including metadata management and data lineage, to ensure operational resilience. Organizations are turning to data catalogs and new tools to help to infer metadata and data lineage.

Data catalogs help organizations discover and understand their data. Modern catalogs include features such as a business glossary, data profiling, and classification, along with data rating, data usage, and data lineage. These catalogs help users understand the data and document data ownership and data definitions. Newer tools automatically classify data as sensitive, collect metadata from a wide range of sources (databases, BI tools, ETL pipelines, cloud platforms, etc.) and centralize it.

Modern catalogs also support data lineage. They provide audit trails including lineage to understand what data was accessed and how it was used. Some can monitor the flow of information from creation or acquisition to use to ensure that no process or transformation has introduced any errors or inconsistencies that might impact compliance. Newer tools use metadata to automatically map end-to-end data lineage, showing how data moves, is transformed, and is consumed across systems.

Additionally, some tools can combine data lineage with data quality insights. In other words, data quality metrics, such as completeness, accuracy, and validity, may be integrated directly into data lineage visualizations. This approach allows users to monitor data quality metrics and observe how they change in data flows.

To support compliance, some tools make use of what is known as "inferred data lineage"—the process of automatically deducing relationships and data flows between systems, tables, or fields (without explicit documentation) using techniques such as metadata analysis or machine learning. This avoids manual lineage processing, reducing the risk of user errors. By enabling end-to-end visibility into data flows, this capability helps organizations strengthen trust in their analytics and AI models while enhancing the overall customer experience.

4 Prioritize privacy and security

Effective data governance requires data to be trustworthy, ensuring it is complete, accurate, timely, relevant, and robust enough to solve intended problems. Additionally, data must be secure, with authorized access, encryption, masking, and protection against unauthorized actions like data leakage. Some key trends, such as shifting work

[11] Unpublished results from *TDWI's Data Ecosystems Assessment*, 2025

patterns post-COVID and remote working, pose additional security risks and require additional processes, investment, and governance. While the vast majority of respondents to TDWI surveys believe that their organization has the right processes in place to keep their data secure, no matter where it resides in the data ecosystem, only about 50% of respondents to TDWI surveys believe that their organization deletes data per regulatory/compliance requirements throughout the data ecosystem.[12]

Numerous regulations exist around data privacy. In the U.S., data privacy is primarily governed by state-level legislation, with the most prominent being the California Consumer Privacy Act (CCPA) and its amendment, the California Privacy Rights Act (CPRA). These laws grant consumers rights over their personal data, including the right to know, delete, and opt out of the sale of personal information. Other states, including Virginia, Colorado, Connecticut, and Utah, have enacted similar laws, and more states are following suit. There are also sector-specific laws including HIPAA (healthcare) and GLBA (financial services). In 2024, Washington State implemented its My Health My Data Act (MHMDA) which broadens privacy around certain newer kinds of consumer health information from devices, which may be unstructured. Globally, in addition to GDPR, there is Brazil's LGPD, Canada's PIPEDA, China's PIPL, and Australia's Privacy Act, all reflecting a growing global consensus on the importance of protecting personal data and holding organizations accountable for how it is collected, processed, and shared. However, the lack of consistency and uniformity across regulations adds to the complexity, requiring greater compliance management, including data management for international organizations.

There are a number of steps that organizations can take to ensure that their data remains private. These include:

- Ensure that your vendors comply with standards and regulations such as ISO/IEC 27001, SOC 2 Type II, HIPAA, FedRAMP, GDPR, and CCPA, among others. That means that they are implementing certain certified controls, undergoing independent audits to demonstrate compliance, and providing authorized services.

- Ensure that vendor products enable organizations to define, manage, and enforce fine-grained user access policies based on roles, responsibilities, and data sensitivity. At a minimum, role-based access controls (RBAC) should be supported. Some vendors provide attribute-based access control (ABAC) for more granular policy enforcement. For example, access can be restricted based on time of day or geographical location of the user.

- If your company is implementing data governance across hybrid or multi-cloud environments, consider federated access management solutions to streamline secure, consistent user authentication and authorization across diverse systems. This approach can enable centralized policy enforcement and integrates with existing identity providers, reducing the risk of access control gaps while improving regulatory compliance and operational efficiency.

- Traditional security mechanisms such as encryption and masking will continue to be important for data both at rest and in transit to deal with sensitive data in any platform.

- In line with the monitoring and observability features described above, platforms should provide capabilities to alert users about anomalous behaviors, unauthorized access, or policy violations.

[12] Ibid.

5 Govern data throughout the AI life cycle

With the advent of AI and generative AI, organizations need to think about governing data across the AI life cycle. That life cycle consists of data collection, preparing data for training AI models, model building, operationalizing AI, and monitoring the data in production. For instance, Article 10 of the EU AI Act mandates data governance practices for high-risk AI systems, including documentation of data collection, preprocessing, and bias mitigation. The CPRA requires governance of data used in AI systems, including profiling. Understanding the flow of data also helps in identifying and mitigating risks associated with data handling and processing. It allows organizations to demonstrate to auditors or regulatory bodies how data is managed, transformed, and protected.

The first data governance checkpoint in the AI life cycle involves managing the quality, sensitivity, and completeness of input data. Organizations must evaluate whether personally identifiable information (PII) or protected attributes are present, ensure proper anonymization or minimization where needed, and assess data sets for bias or representational gaps. Data lineage tools are important here, as they help track the source, transformations, and versioning of data sets used in model training. Establishing and enforcing policies for data sourcing, documentation, and access rights helps reduce downstream risks and supports compliance audits. In high-risk AI systems, organizations must be able to explain how training data was selected and what steps were taken to mitigate harmful outcomes.

Once models are trained and deployed, data governance must extend to outputs and their downstream use. Output data—such as credit scores, risk ratings, or predictive classifications—may feed into other decision systems, making accuracy and completeness important. Monitoring is critical to detect quality drift, model degradation, or violations of organizational policies (see below). For example, a model trained on clean data may begin to fail if the input distribution changes. Organizations should establish alerts for anomalies, log model decisions, and track the provenance of output data to ensure consistency and traceability. Additionally, data should be tagged appropriately to inform downstream users of its origin, sensitivity, and any associated risks. AI-infused observability tools can aid in this process by automating monitoring, reporting, and compliance checks.

Effective governance also requires cross-functional AI literacy/fluency—a foundational understanding of data quality, privacy, model behavior, and risk. Organizations should develop an AI skills matrix and training plan that fosters literacy across teams, from data scientists to compliance officers and business users. Leadership buy-in is essential to institutionalize responsible AI practices and align governance with strategic goals. As the consumerization of AI continues, equally important is the ability to democratize data and AI responsibly: making insights accessible across the organization, while ensuring trust, transparency, and adherence to regulatory policies.

6 Investigate AI governance in addition to data governance

Organizations will need to evolve their data governance to include AI governance, which involves the responsible use of the technology. It is one thing to govern the data that goes into and out of AI models, however, organizations will also need to govern the models and applications built from the models themselves. For organizations, this means not only adhering to existing legal and compliance frameworks but taking a holistic approach to foster trust and confidence among users and stakeholders.

AI model governance involves establishing policies, procedures, and tools to ensure that AI models are developed, deployed, and maintained responsibly and ethically. It includes practices such as monitoring model performance, ensuring compliance with regulatory standards, and addressing issues such as bias, fairness, and explainability. Effective model governance also involves regular audits, version control, and tracking to maintain transparency and accountability throughout the AI life cycle.

While AI governance is relatively new on many companies' radar, it is important. Governing AI models was implied in Article 22(1) of the GDPR, "The data subject shall have the right not to be subject to a decision based solely on automated processing."[13] In the EU AI Act, high-risk systems including health and credit scoring require traceabililty and explainability.

There are a number of considerations that are important for AI governance. The first is that AI governance isn't the same as data governance. An AI model is a piece of software (albeit different than a simple codebase) and needs to be governed. This includes maintaining a robust versioning system for models, ensuring that every iteration is tracked, reproducible, and linked to the specific training data and parameters used. Models should be stored in a repository that includes documentation of their intended purpose, a feature repository, performance benchmarks, and any ethical or regulatory considerations. Lineage tracking is also important—not just for the data used to train the model, but for the model itself, including dependencies, code artifacts, and deployment context. These practices enable greater transparency, accountability, and control throughout the AI life cycle.

Additionally, model decay/drift is inevitable in AI systems. As environmental conditions, user behavior, or input data distributions change over time, models will degrade in performance—no model is static. Regulations such as the EU AI Act and FTC oversight in the U.S. require organizations to ensure that AI models remain accurate and fair over time. Failing to detect and address model drift can result in biased or erroneous decisions that impact consumers, especially in areas such as lending, healthcare, or hiring. To maintain compliance and build trust, organizations must implement robust monitoring systems to continuously evaluate model outputs, detect drift early, and take corrective action to mitigate harm.

Explainability involves understanding the why behind an AI model in a way a human can understand. If a model is built and put into production, the output of this model—and what went into determining the output—need to be understandable. For instance, a customer should be able to understand why his loan or credit card application was rejected. This is important for fairness and transparency. In some instances, organizations are starting to utilize tools that help with explainability of model output. Two popular techniques include LIME, (local interpretable

[13] https://gdpr-info.eu/art-22-gdpr/

model-agnostic explanations) and Shapley values. In some instances, vendors have incorporated these into their software to produce charts where users can determine the most important features in predicting the outcome of interest.

Of course, success requires cross-functional coordination to align on policies and developing a culture of compliance. It will require MLOps, governance teams, compliance and legal, and business leadership all playing critical parts.

7 Don't forget about responsible governance

Responsible data and AI considers the ethical, societal, compliance, legal and environmental ramifications of using data in a wide variety of applications and processes. Responsible data and AI governance can be a strategic framework for proactively addressing broader planetary, societal, and business concerns, or a set of tactics for mitigating the risks and downsides from inadvertent, accidental, or ill-advised uses of data and analytics. Tied to the notion of responsible data and AI is the notion of environmental, social, and governance (ESG) reporting. On the environmental front, organizations are dealing with issues such as their carbon footprint to support sustainability goals. Societal issues include fairness, bias, privacy, and respect for human rights. Governance includes clear policies around ethics and transparency.

ESG reporting has shifted from a voluntary practice to a regulatory requirement for modern enterprises, driven by mandates such as the EU's Corporate Sustainability Reporting Directive (CSRD) and the U.S. SEC's Climate Disclosure Rule, making sustainability a core business imperative. The EU's Corporate Sustainability Reporting Directive (CSRD) mandates transparent and auditable ESG data reporting on a number of issues around climate change and human rights. It impacts all companies with securities listed on an EU-regulated market.[14] Financial standards in the U.S. include SASB for disclosing sustainability initiatives to investors and SEC standards for reporting.

Companies must regularly monitor their compliance with CSRD requirements to avoid penalties, which can include substantial fines and legal consequences, which will vary by country (for example, in the U.K. fines could be up to €580,000 and in Germany it could reach up to €10 million or 5% of company turnover).[15]

To effectively manage this data and create an audit-ready strategy, organizations may need a trusted repository or data hub, where data can be profiled, cleansed, and validated to ensure consistency and reliability across reporting metrics. To promote transparency and traceability, the platform may offer automated data lineage, showing how ESG data flows from source systems (e.g., emissions tracking, HR platforms) through transformation pipelines to final dashboards or reports. Additionally, data integration tools would allow ESG data to be combined with operational, financial, or customer data, enabling more informed and sustainable business decisions.

[14] https://www.pwc.com/us/en/services/esg/library/eu-corporate-sustainability-reporting-directive.html

[15] https://senecaesg.com/insights/csrd-penalties-for-non-compliance-understanding-the-stakes/

Concluding thoughts

Organizations today face an increasingly complex array of regulatory expectations and business pressures that demand not only better data—but better governance of that data across its entire life cycle. From evolving global privacy laws and ESG mandates to the rise of AI and generative AI systems, the governance landscape demands transparency, accountability, and ethical oversight. This means that organizations must understand emerging regulations, elevate data quality, implement metadata and lineage, secure sensitive data, and extend governance into the AI and model development process.

In this environment, governance is no longer a back-office function—it is a critical enabler of innovation, trust, and competitive advantage. The organizations that succeed will be those that embed governance into their culture and infrastructure, making it front and center for responsible, data-driven transformation. Strong governance and compliance not only reduce risk but also enhance operational efficiency, strengthen investor confidence, and build brand trust with customers—ultimately positioning organizations to compete more effectively.

About our sponsor

Informatica

Informatica (NYSE: INFA), a leader in enterprise AI-powered cloud data management, brings data and AI to life by empowering businesses to realize the transformative power of their most critical assets. We have created a new category of software, the Informatica Intelligent Data Management Cloud (IDMC), which is the only platform that manages the entire life cycle of data and interoperates with everything. It is an end-to-end data management platform, powered by CLAIRE AI, that connects, manages, and unifies data across virtually any multi-cloud, hybrid system, democratizing data and enabling enterprises to modernize their business strategies. Customers in approximately 100 countries and more than 80 of the *Fortune* 100 rely on Informatica to drive data-led digital transformation.

Informatica. "Where data and AI come to life."

About the author

Fern Halper, Ph.D., is vice president and senior director of TDWI Research for advanced analytics. She is well known in the analytics community, having been published hundreds of times on data mining and information technology over the past 20 years. Halper is also coauthor of several Dummies books on cloud computing and big data. She focuses on advanced analytics, including predictive analytics, machine learning, AI, cognitive computing, and big data analytics approaches. She has been a partner at industry analyst firm Hurwitz & Associates and a lead data analyst for Bell Labs. She has taught at both Colgate University and Bentley University. Her Ph.D. is from Texas A&M University.

You can reach her by email (fhalper@tdwi.org) and on LinkedIn (linkedin.com/in/fbhalper).

About TDWI Research

TDWI Research provides industry-leading research and advice for data and analytics professionals worldwide. TDWI Research focuses on modern data management, analytics, and data science approaches and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of business and technical challenges surrounding the deployment and use of data and analytics. TDWI Research offers in-depth research reports, commentary, assessments, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.

About TDWI Checklist Reports

TDWI Checklist Reports provide an overview of success factors for a specific project in business intelligence, data warehousing, analytics, or a related data management discipline. Companies may use this overview to get organized before beginning a project or to identify goals and areas of improvement for current projects.

Transforming Data With Intelligence™

A Division of 1105 Media
6300 Canoga Avenue, Suite 1150
Woodland Hills, CA 91367

info@tdwi.org

tdwi.org