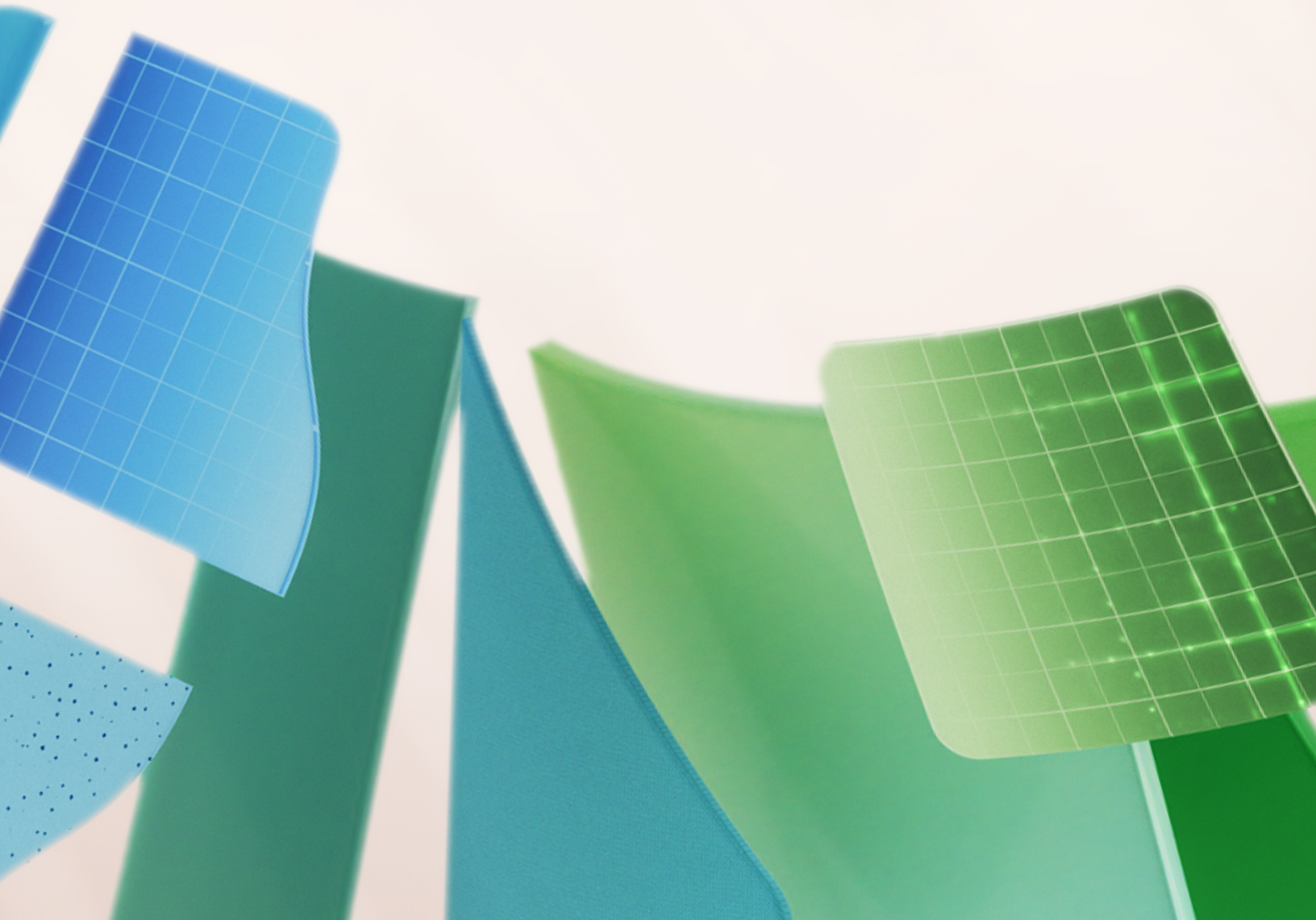
Microsoft 365

Data in Check

Mastering data governance for AI success



Contents

Introduction

Data governance is the backbone of secure AI adoption

Generative AI holds the promise of business transformation—but realizing that potential depends on the quality and availability of your business data.

Organizations prioritizing strong data governance position themselves to gain a competitive edge with AI. Data governance helps ensure that data queried and generated by AI is correct, secure, controlled, and compliant.

Data quality and availability: AI systems rely on large, high-quality datasets to provide the best response possible with proper context from your organization's unique workflows. Better data quality leads to more effective AI insights and predictions.

Compliance: Strong data governance ensures adherence to regulatory requirements, reducing the risk of legal issues and fines.

Security: Proper data labeling and management protects sensitive information from unauthorized access, inappropriate sharing, and breaches.

Trust: Reliable data governance builds stakeholder confidence in AI outputs, fostering greater adoption and support for AI initiatives.

This e-book introduces critical data governance practices that create an AI-ready organization. It covers how to set up robust data quality standards, ensure compliance with regulatory requirements, and implement data protection and security measures. Whether it's keeping sensitive meeting content from being shared externally or guarding against data breaches, you'll learn strategies for building trust in AI systems and making data readily available to drive actionable insights. By adopting these practices, your organization can apply the full potential of AI for continuous innovation and competitive advantage.

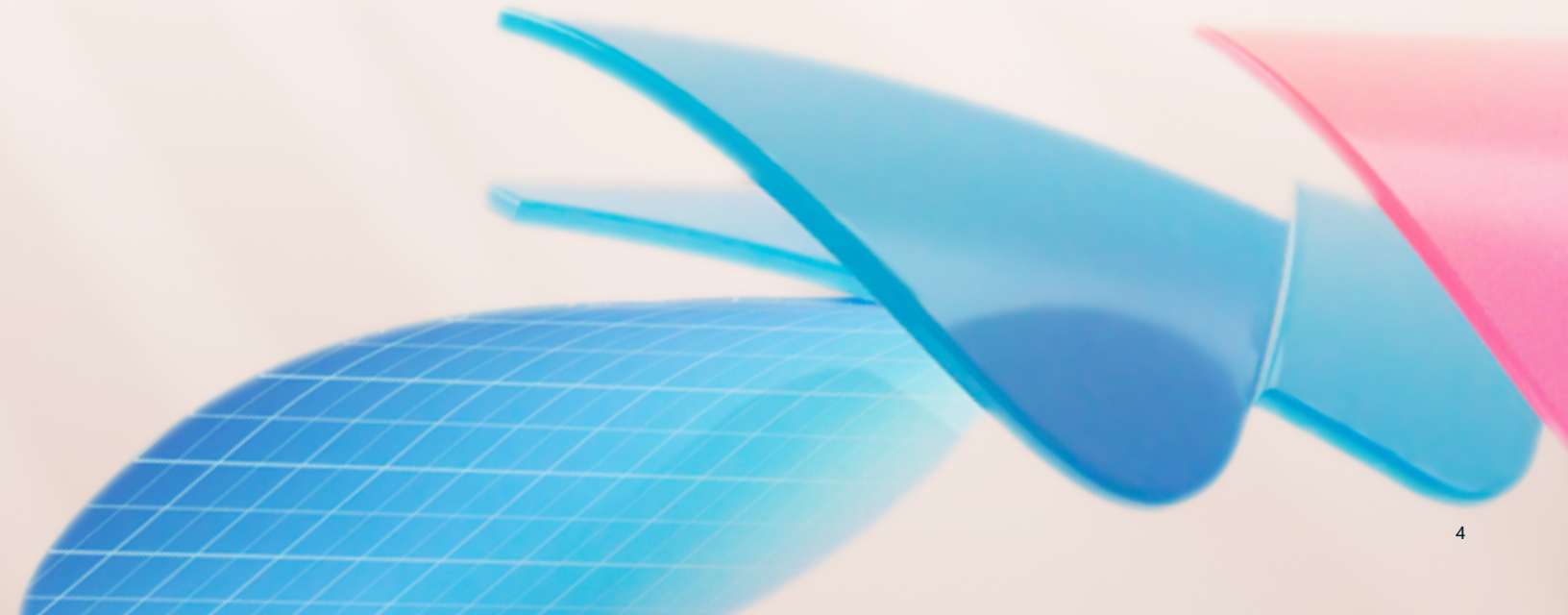
Generative AI use cases

Accelerate communication: Draft personalized content faster, freeing up time to build relationships and collaborate.

Improve efficiency: Reduce time spent on routine tasks, enhance productivity, and cut costs.

Enable innovation: Help generate ideas and proposals for new products and services.

Personalize customer experiences: Tailor content and recommendations to drive loyalty and engagement.



1

Strengthening data governance for AI transformation

Every enterprise has policies and processes governing data usage. Data governance frameworks vary in maturity and comprehensiveness, but few have been fully optimized for AI. While many data governance best practices stay the same, such as ensuring data accuracy and consistency, other aspects need updates to maximize AI investments. Let's look at a few key areas.

Data visibility

Detailed knowledge of data flow within AI systems makes it possible to find and mitigate unauthorized or inappropriate use. This visibility helps support security and compliance, protecting sensitive data to maximize AI value.

Traditional data governance has focused on knowing where data lives and controlling access. However, as AI becomes more integrated into business operations, data governance needs to keep up with evolving security needs. For instance, with a highly sensitive product launch, governance must now extend to managing AI-generated content, ensuring that project-related documents, communications, and insights produced by AI are secure. This means implementing safeguards so that only authorized team members can access and use AI to analyze or summarize project information.

Additionally, by managing data processing and storage volumes, organizations can better control the operational costs associated with AI.

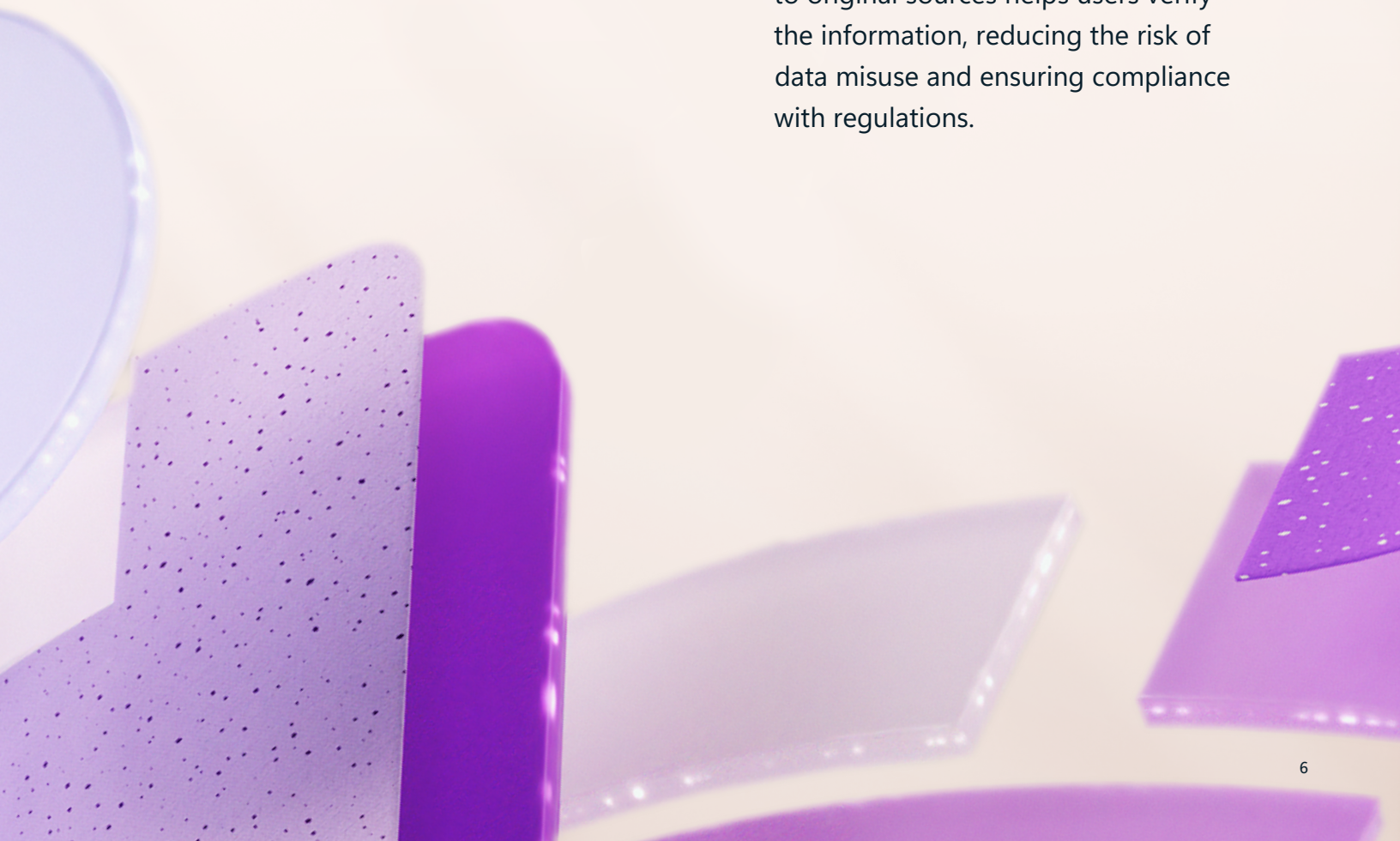
Data quality

AI amplifies the importance of high-quality data, as poor data quality directly affects AI outcomes. When applying AI tools to query your business data, you want to make sure that data is current and trustworthy. Equally important is understanding the provenance and quality of data your teams use to build their own AI models and apps. Regular data audits, stringent validation processes, and proactive management help ensure data integrity. By focusing on these areas, organizations can achieve reliable AI results that drive better decision-making and innovation.

User management

With AI, users interact with data in sophisticated ways. They communicate with AI systems like Microsoft 365 Copilot through natural language prompts. With permission and protections in place, AI can access relevant context from data—such as files, chats, and emails—along with external sources via plugins to generate a response.

It's crucial to monitor the data used in these processes to ensure the AI provides correct, grounded responses. Providing transparency through footnotes or links to original sources helps users verify the information, reducing the risk of data misuse and ensuring compliance with regulations.



Compliance and eDiscovery

The rise of AI introduces new challenges in compliance and eDiscovery (the handling of data for legal cases), particularly in managing AI-generated data and adapting to evolving legal requirements. Updating data governance frameworks to address these challenges involves developing policies that cover AI-generated content, such as ensuring AI-produced documents or communications are tracked, categorized, and stored securely. For example, policies may need to be updated to ensure that AI-generated emails or reports are tagged and archived properly for future retrieval.

Enhancing eDiscovery capabilities would include integrating AI tools that can search and identify AI-generated content across various platforms. For instance, during a legal inquiry, eDiscovery tools must be able to find and retrieve specific AI-generated documents, summarize relevant communications, and provide clear audit trails to prove compliance. By updating these capabilities, organizations can better manage data during legal audits or investigations, ensuring that all relevant AI-generated information is accessible and defensible in court.

Data security

Securing AI-driven operations should involve Zero Trust principles at the identity level, which minimizes the risk of unauthorized access. Regular updates to endpoints, including devices and applications, reduce vulnerabilities that could be exploited. Awareness of the generative AI tools in use within the organization allows for the blocking of unsanctioned or insecure applications, which in turn prevents potential security breaches. By limiting access to AI tools and data to only trusted personnel, organizations can achieve greater data integrity and protect their AI operations from potential threats.

What is Zero Trust?

Zero Trust is a security model that focuses on verifying every request as if it came from an untrusted network. Rather than assuming everything within the corporate firewall is safe, this approach adopts the principle "never trust, always verify."

Zero Trust principles

Verify explicitly: Always authenticate and authorize based on all available data points.

Use least privilege access: Limit user access with Just-In-Time and Just-Enough-Access (JIT/JEA), risk-based adaptive policies, and data protection.

Assume breach: Limit damage, control access, ensure encryption, and use data to spot threats and strengthen defenses.

2

The role of data stewardship in AI transformation: Origins, quality, and reliability

As organizations adopt AI to assist in business decisions, they must ensure that these systems can provide correct and reliable results. The data grounding AI responses must be available, consistent, and well-documented.

Data stewardship supports trustworthy AI by implementing governance policies that track where data comes from, check its quality, and ensure its reliability and accuracy. These practices build the foundation for AI systems that deliver dependable insights, enabling informed and confident decision-making.

Data lineage: Understanding origins and changes to information

Data lineage tracks data's journey through an organization. It documents the data's origins, transformations, and destinations. For AI-powered businesses, understanding data lineage is vital for several reasons:

- **Transparency:** Knowing where data comes from and how it changes helps confirm the accuracy of AI-generated outputs and ensures regulatory compliance by providing a clear understanding of the data's history.
- **Traceability:** Data lineage allows organizations to trace errors or inconsistencies back to their source, simplifying troubleshooting and data correction.
- **Impact analysis:** Understanding how data is used by generative AI tools helps assess the potential impact of changes in data sources or processing methods on output accuracy and business outcomes.

Data quality: Quality drives the value of AI responses

Data quality directly affects the reliability of AI outputs. High-quality data provides the context for AI models to deliver correct and valuable responses to user inputs. Key aspects of data quality include:

- **Accuracy:** Data must correctly represent real-world conditions.
- **Completeness:** All necessary data must be present and accounted for.
- **Consistency:** Data must be consistent across different systems and over time.
- **Timeliness:** Data must be up to date and available when needed.

Data reliability: Ensuring trustworthy data

Data reliability means data consistently meets quality standards and is available when needed. For AI-powered businesses, reliable data is crucial for promoting trust in AI tools and the decisions they inform. Ensuring data reliability involves:

- **Data redundancy:** Implementing backup systems that prevent loss and increase availability.
- **Regular backups:** Conducting frequent backups to safeguard against data corruption or loss.
- **Monitoring and alerts:** Setting up monitoring systems to detect and alert stakeholders to data issues in real time.
- **Disaster recovery plans:** Developing and testing plans to recover data and resume operations quickly after disruptions.

3

The relationship between governance, security, and responsible AI

Together, data governance and security create the backbone for responsible AI use. High-quality, secure data ensures AI works ethically and effectively. Key areas to evaluate include data classification, access controls, encryption, incident response, and regulatory compliance.

Data classification

Classifying data is a critical part of controlling how AI tools handle sensitive information. Typically, data and meetings are classified as general, confidential, or highly confidential. Proper classification ensures that AI only accesses appropriate information, reducing the risk of exposing sensitive data to unauthorized users.

Misclassification, on the other hand, can lead to AI processing data that should be restricted, resulting in security breaches or compliance issues. Effective data governance, whether through automated tools or end-user policies, ensures that data is classified correctly, safeguarding sensitive information and supporting the ability of AI to deliver reliable and compliant outputs.

Access controls

These controls regulate who can access AI-relevant data and which applications or identities have permission to interact with that data. Weak access controls can lead to unauthorized exposure of sensitive information, increasing the risk of breaches and misuse.

Data governance plays a key role in enforcing these controls by restricting data access to authorized personnel and specific applications, ensuring that sensitive data is handled appropriately. This not only protects data integrity but also ensures that AI systems operate on secure, trusted datasets, enhancing their reliability and compliance.

Encryption

Securing data from interception and tampering using encryption helps ensure that generative AI tools can ground their responses in the correct context—such as work-related data, files, chats, and emails—without risking data leakage.

Data governance policies can mandate robust encryption practices, protecting data throughout its lifecycle. This approach ensures that AI tools can deliver reliable responses while keeping your data secure and maintaining trust.

Incident response

Sensitive organizational data can be exposed through incidents involving generative AI tools that grant unauthorized access to files, emails, or other business data that systems use to generate responses.

A proactive incident response plan is crucial in these scenarios. Without such a plan, the organization risks not only exposing sensitive data but also relying on compromised outputs from the AI. Data governance includes having detailed response protocols to quickly address breaches, minimizing their impact and preserving the reliability of AI systems.

Regulatory compliance

AI applications must adhere to regulations such as GDPR or CCPA, which govern data protection and privacy. Non-compliance can lead to significant penalties and erode trust. It's also crucial to understand where AI tools process data, as many free tools may handle data globally or outside of your company's usual storage locations. Data governance ensures that AI applications not only run within legal frameworks but also keep data within the right service boundaries, aligning with your organization's compliance standards. This approach supports ethical AI use and helps build trust in AI technology.

4

Updating your data governance framework to support AI adoption

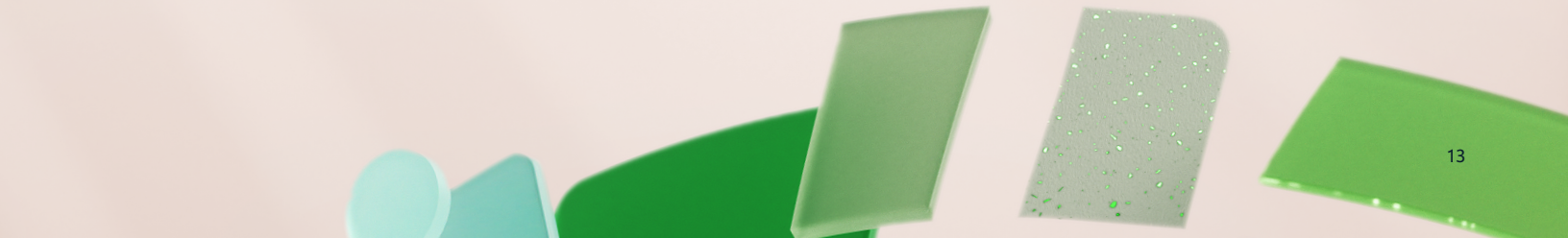
In the evolving AI landscape, it's important to identify specific areas of your existing data governance framework that might require extra attention. Rather than overhauling the entire framework, you can focus on areas that are most likely to evolve, directing your efforts where they will most enhance the value of AI while ensuring your data stays protected and secure. Here are some key insights to consider:

Adapting policies and procedures

AI involves new kinds of data collection, processing, and usage. Updating data policies can help address relevant needs around data privacy, regulatory compliance, and ethical use. For instance, requiring data anonymization in specific instances can protect personal information, making it safer to use throughout the AI lifecycle.

Roles and responsibilities

It's essential to incorporate AI readiness across the roles involved with data governance. Employees familiar with AI data requirements can act as stewards to support data quality and compliance. Cross-functional collaboration among legal, IT, and data science teams can help tackle AI-specific challenges more effectively.



Adapting data standards and definitions

Standardizing data formats, definitions, and quality metrics simplifies the implementation of policies governing AI tool usage within an organization. Clear data standards make it easier to decide which datasets AI tools can reason over to provide business context and which data users can upload for analysis. This ensures that AI applications use the most relevant and trusted data, enhancing their effectiveness while supporting compliance with organizational policies.

Continuous improvement

Data governance is an ongoing process, especially with AI. Regular audits and updates to your governance framework can help adapt it to new AI developments and regulatory changes. AI itself can be used to check and enhance governance practices, potentially finding gaps and suggesting improvements. This proactive approach enhances compliance and efficiency as AI technologies evolve.

Tools and techniques for effective data governance

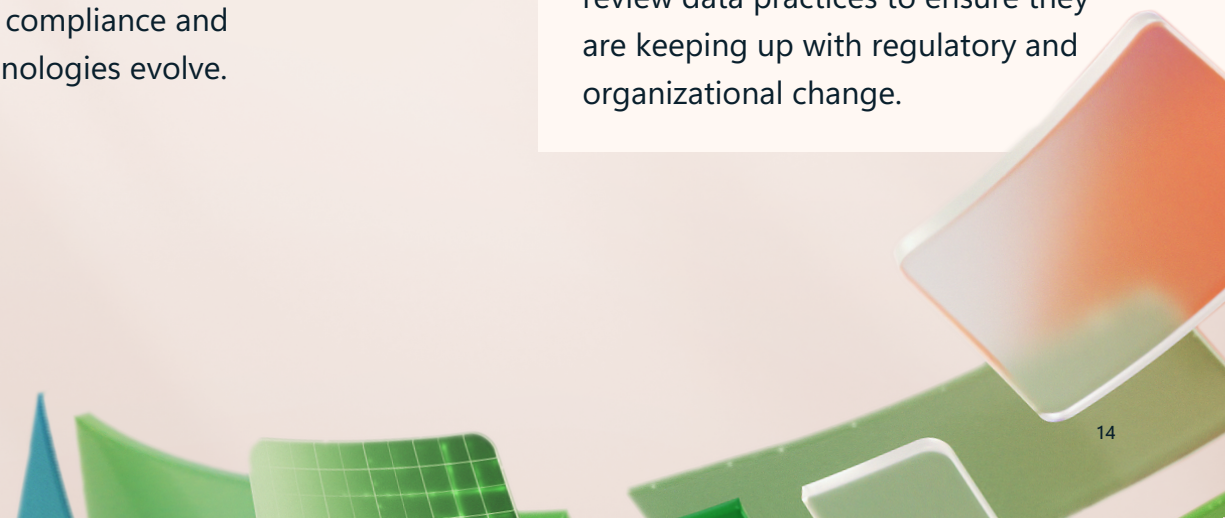
Align with business goals: Ensure your data governance strategy supports organizational aims to enhance decision-making and efficiency.

Automate routine tasks: Use automation for repetitive tasks like data classification and access management to reduce manual errors and improve security.

Ensure high standards of accuracy: Use validation and cleansing tools to support data integrity over time.

Train users: Offer training in data governance tools and procedures to prevent mismanagement and ensure compliance with policies.

Conduct regular audits: Regularly review data practices to ensure they are keeping up with regulatory and organizational change.



Conclusion

Effective data governance for AI enablement with Microsoft 365

Using AI responsibly requires robust data governance. Effective data governance ensures data availability, accuracy, and security, enabling AI to deliver reliable insights and foster innovation. Prioritizing data quality, compliance, and security enhances AI capabilities, driving better decision-making and maintaining a competitive edge.

Microsoft 365: Empowering AI readiness

Microsoft 365 offers best-in-class productivity apps with built-in tools for data classification, control, and protection. These features support your data governance framework, facilitating confident AI adoption when you're ready. Microsoft 365 helps you ensure:

- **Data quality:** Support accuracy and reliability for AI outputs.
- **Compliance:** Adhere to regulatory requirements, reducing legal risks.
- **Zero Trust security:** Protect sensitive information from unauthorized access, breaches, and cyberthreats.

Discover comprehensive productivity tools enhanced with AI options and robust protection to help your organization work efficiently and securely.

Explore Microsoft 365