Data privacy and security with IBM

Establishing a unified framework for automated data governance, security, and compliance across all phases of the AI lifecycle

Highlights

- A unified privacy framework for the entire AI lifecycle
  - Easily deploy a comprehensive framework on the cloud environment of your choice
  - Collect, organize and govern all enterprise data with a data fabric
  - Layer on robust privacy tools for data masking, de-identification and active metadata for pervasive enforcement
  - Connect with OpenPages for a unified view of all private data assets alongside enforcement of the latest regulations and robust auditing

As organizations consolidate their information architecture stack, five considerations are critical to establishing a foundation for trusted data:

- Pervasive data governance
- Automated data discovery and cataloging
- Continuous auditing and reporting
- Timely response and assessment
- Deployment flexibility

IBM Cloud Pak® for Data delivers the critical capabilities necessary to build a unified privacy framework spanning the entire data and AI lifecycle. From hybrid data management, data governance and security to data privacy, risk and compliance, the collaborative platform provides a fully integrated solution that can be deployed on any cloud.

By leveraging the platform capabilities, businesses can achieve a real-time view of personally identifiable information (PII) across their enterprise and throughout every stage of analysis. Intelligent automation further simplifies how organizations understand and manage sensitive data. Embedded automation capabilities help accelerate the collection, cataloging and masking of sensitive data to build a foundation of trusted, compliant data that can be easily accessed by your teams and models.

Try it today for free →

A unified privacy framework via IBM Cloud Pak for Data

Organizations collect a tremendous amount of data from a variety of sources, and any of these data sources could potentially contain sensitive data. With data often being relocated for warehousing, reporting, analytics, storage, testing and application use, it becomes increasingly difficult to understand everywhere that sensitive information resides. As more focus is put on AI and the outcomes it drives it's no longer enough to simply understand who has had access to that information; business leaders now must also answer for what models have used it and what subsequent outcomes they have driven.

The emergence of newer technology platforms such as cloud and AI can exacerbate the issue. Organizations often feel a tension between their need for innovation and their responsibilities for data governance, security and regulatory compliance. However, in reality, a well-governed, secure environment can actually spur innovation and lead to an increase in organizational productivity. In order to understand the amount of sensitive data living across the organization and mitigate associated risks, it is important to examine the entire data landscape to ensure all regulatory requirements regarding its lifecycle and correct usage are met.

The data lifecycle should be managed from creation to disposal and include everything in between. "Gartner found, in the *2020 CISO Effectiveness Survey*, that 78% of CISOs have 16 or more tools in their cybersecurity vendor portfolio; 12% have 46 or more. Most organizations recognize vendor consolidation as an avenue for reduced costs and better security, with 80% of organizations interested in vendor consolidation strategy."[1] IBM Cloud Pak for Data provides the opportunity to make end-to-end data lifecycle management a reality while simultaneously allowing for vendor consolidation without lock in.

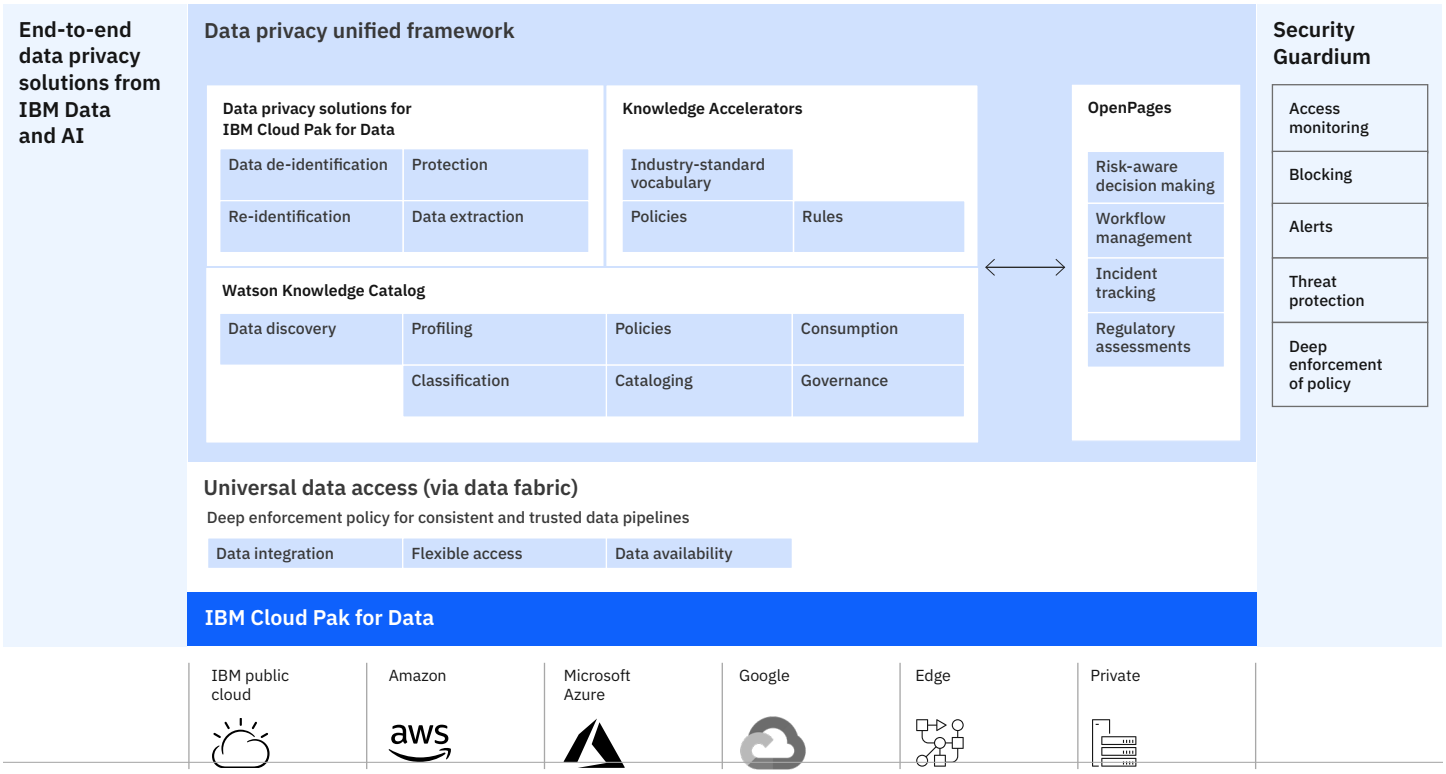

Figure 1. Capabilities represented in the unified data privacy framework deployed on IBM Cloud Pak for Data

How it works

Red Hat OpenShift

This unified data privacy framework is delivered via IBM Cloud Pak for Data, a fully integrated data and AI platform that enables organizations to accelerate AI-powered transformation by unleashing productivity and reducing complexity. Cloud-native by design, IBM Cloud Pak for Data is built on and takes advantage of the underlying resource and infrastructure optimization and management of Red Hat® OpenShift® Container Platform. The OpenShift foundation also provides true deployment flexibility, because it allows the platform to be deployed on any cloud and fully supports hybrid environments spanning AWS, Azure, Google Cloud Platform, IBM Cloud® and private cloud deployments.

Data fabric

The latest version of IBM Cloud Pak for Data further infuses automation throughout the platform, allowing users to build an intelligent data fabric that helps connect the right data to the right people at the right time. This new architecture makes up the foundation of the framework and uses the power of AI to automate complex data management tasks to help universally discover, integrate, catalog, secure and govern data across multiple environments. One of these automated workstreams is data privacy and security or AutoPrivacy as it is referred to. AutoPrivacy speaks to capabilities that employ AI to intelligently automate the identification and monitoring of sensitive data, and subsequently the enforcement of data policies throughout the organization. In conjunction with the broader framework, AutoPrivacy helps accelerate compliance and minimize risk.

IBM Watson Knowledge Catalog

One of the products that supports AutoPrivacy is IBM Watson® Knowledge Catalog, a data catalog backed by active metadata and policy management. It serves as the heart of this intelligent data fabric and the overarching data privacy framework. It tethers all other capabilities to a data governance experience that delivers data in context, while helping ensure data quality. Watson Knowledge Catalog in turn, leverages capabilities from the following products to advance data privacy programs:

- IBM Knowledge Accelerators help organize data along a common and known business vocabulary, in addition to automatically providing business context and definitions during the on-boarding of regulatory and industry data content within Watson Knowledge Catalog.
- IBM Data Privacy for Cloud Pak for Data provisions masked data quickly while maintaining data utility and relationships to achieve better and faster insights. The de-identification process is accelerated by importing sensitive data assets into Watson Knowledge Catalog to automatically discover and classify data to be masked.

IBM OpenPages with Watson

Watson Knowledge Catalog then enables the loading of asset metadata into IBM® OpenPages® Data Privacy Management, which gives users a unified view of all private data assets being stored across their organization and allows those users to run privacy assessments on them. Combined, both products cover the spectrum of discovery and usage scanning to identify sensitive information.

A Total Economic Impact™ (TEI) study commissioned by IBM from Forrester Consulting demonstrated a three-year 218% ROI for IBM OpenPages with Watson, which totals USD 5.1 million in benefits. This is comprised of roughly $1.65 million in regulatory fine avoidance, $1.65 million in reduced risk management effort and $1.8 million in avoided legacy solution costs.[2]

IBM Security™ Guardium solutions

The details provide a robust data audit framework that can be sent to IBM Guardium® Insights and monitored there to protect data in motion. Serving as the security operations center, Guardium Insights can help continuously uncover suspicious activity, plus store years of security and compliance data.

A Total Economic Impact™ (TEI) study commissioned by IBM from Forrester Consulting demonstrated a three-year 401% ROI for IBM Security™ Guardium, which totals USD 5 million in benefits. This is comprised of roughly $2.1 in increased audit efficiencies, $1.1 million in increased ability to meet compliance regulations, $1 million in increased database security and $0.8 million in increased database analysis automation.[3]

What it provides

Data privacy is no longer a compliance checkbox; it has become a strategic competitive advantage to raise the bar on brand trust with consumers. With IBM Cloud Pak for Data and the unified privacy framework it supports, enterprises can undertake risk-aware business decisions, incorporate them into their workflows, protect enterprise-wide data and be compliant with regulatory requirements. Some of the benefits are as follows:

Continuous auditing

Create visibility across sources to manage related data. Greater availability of data, facilitated by real-time integrations, makes it possible for compliance experts to monitor a wide variety of sources. A governance, risk and compliance (GRC) solution helps to provide aggregate insights into systemic issues in controls, processes and compliance for dynamic areas such as cyber risk and data protection. For this reason, the solution also provides a connected library of risk and compliance items that can be viewed across different business dimensions.

Data governance

With a data catalog at the core, you can create and automate policies for enterprise-wide categorizing and classification of data. This occurs everywhere data resides to ensure the appropriate data protection measures are applied and triggered when accessing, using, or transferring data that is classified as sensitive. Additional capabilities such as data masking, user-based access controls for discovery, and risk assessment of unstructured data are also available for an even more robust approach to data governance.

Data discovery

Focus on streamlining data operations with increased efficiency, data quality, findability and governing rules in order to provide an efficient, self-service data pipeline to the right people at the right time from any source. This will prove invaluable as more departments within the organization express the need to manage and access data.

Timely response and assessment

Implement changes to governance artifacts quickly by automating the reporting of personally identifiable information (PII) in order to improve accuracy and reduce audit times. Give data citizens a holistic, real-time view of how private data is being used throughout the organization, from applications to AI models.

Deployment flexibility

Make your information architecture efficient and agile to meet the demands of today and stay competitive tomorrow. Help chief data and privacy officers build collaborative workflows and automate their AI lifecycles across an array of contributors with an agile and resilient cloud-native platform. Enable data citizens to succeed with AI irrespective of their unique data and cloud landscape. Container-based platforms, such as Red Hat OpenShift, help realize these benefits anywhere through containerized services, container management and orchestration that can lower IT infrastructure and development costs by up to 38% per application. Plus, they can be deployed across any environment—whether on-premises or across multiple clouds and multiple vendors.

Learn more

IBM provides a unified framework for data privacy and security through an integrated set of automated and market-leading capabilities running on IBM Cloud Pak for Data. The auditing, governance, discovery, response and assessment capabilities that can be deployed anywhere establish the foundation needed for trusted data.

Try it today to see for yourself or schedule a consultation with an IBM expert to have your questions answered.

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America
July 2021

1 Smarter with Gartner, *Gartner Top Security and Risk Trends for 2021*, Kasey Panetta, April 5, 2021

2 *The Forrester Total Economic Impact™ Of IBM OpenPages With Watson*, IBM, October 2020

3 *The Total Economic ImpactTM Of IBM Security Guardium*, IBM, 2020.