Cloud transformation

SPONSORED CONTENT | WHITE PAPER

How regulated firms can have it all: **The cloud, AI, and security wherever the data resides**

Highly regulated companies can use artificial intelligence, data sharing, and cloud-based technologies for even their most sensitive data while meeting stringent regulatory requirements for data privacy and security.

CIO

SPONSORED BY CLOUDERA PROTEGRITY

Companies in highly regulated industries want to make effective use of the cloud, including for artificial intelligence (AI) and machine learning (ML) workloads, but are often hamstrung by security, compliance, and privacy concerns.

This is a challenge to overcome, because taking advantage of cloud offerings is a clear priority. A recent Foundry survey found that analytics and business intelligence (BI) capabilities are the top cloud investment priorities for financial firms, for example, cited by 64% of the respondents. AI/ML closely follows, at 62%. But respondents are also mindful of security concerns, as cloud-based security is not far behind, at 54%.[1]

Addressing these issues requires solutions that address not just security but also data privacy. The answer is a data-centric, fit-for-purpose approach that can adopt different technologies, depending on the use case. Solutions should go beyond mere encryption to include tokenization, masking, anonymization, and synthetic data generation.

Regulated industries have requirements that now necessitate leveraging AI to manage data anywhere — in clouds, data centers, and at the edge — to access 100% of their data wherever it resides.

Your strategy should include adoption of platforms that make effective use of existing data as well as advanced AI-driven analytics that complement those provided by cloud providers such as Amazon Web Services (AWS).

[**64%** of financial firms say **analytics and business intelligence (BI) capabilities** are their top cloud investment priorities.]

[1] Foundry, 2025 Cloud Computing Study, June 2025; https://foundryco.com/research/cloud-computing/

Cloud challenges abound

Flexible, all-encompassing platforms are required to address the diversity of challenges that highly regulated firms face by spanning hybrid cloud environments and supporting multiple cloud regulations, including:

- **European Union General Data Protection Regulation (GDPR)**
- **Payment Card Industry Data Security Standard (PCI-DSS)**
- **U.S. Health Insurance Portability and Accountability Act (HIPAA)**

These regulations have varying penalties for noncompliance. GDPR fines can run up to €20 million (~US$23 million) or 4% of a company's annual revenue, and HIPAA violations can result in fines such as $100,000 and five years in prison. In addition, companies can suffer severe reputational harm, with loss of customers and hence revenues.

It's challenging to migrate critical data workloads to the cloud while adhering to the varied regulatory standards – especially adhering to data residency requirements and avoiding the exposure of sensitive data.

Also, the shared responsibility models created by cloud providers make it clear: Enterprises share the onus for security. Hyperscalers typically provide infrastructure security, but their customers are responsible for data security, including data classification, encryption, and access control.

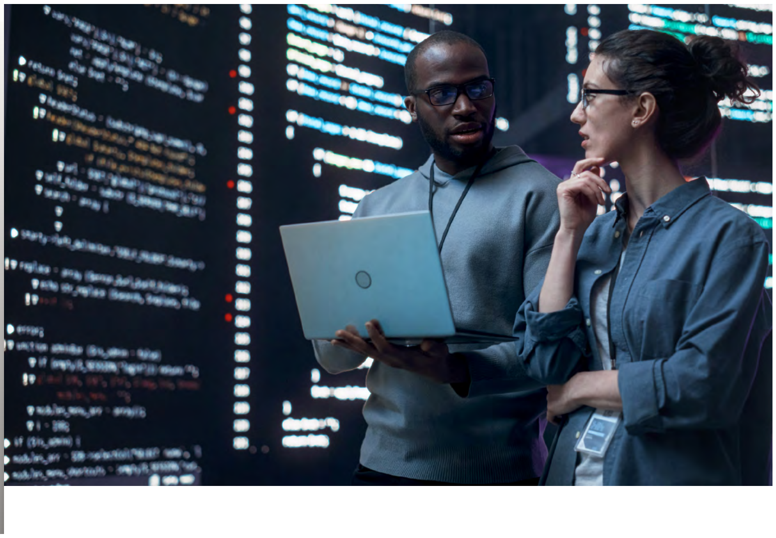
AI adds complexity

As organizations add AI capabilities, data security and privacy challenges proliferate.

Data governance is a significant issue. Organizations must ensure that employees don't share private data with public AI models. Controls should be put in place to prevent users from uploading sensitive data, inadvertently or not, to public AI tools.

In addition, AI can make it increasingly challenging to comply with data privacy, transparency, and fairness regulations. Companies need the ability to demonstrate that they're not sharing private customer data with public AI models.

There's also the likelihood that employees will increase their use of AI tools, compared to previous generations of BI solutions, which traditionally have been adopted by data analysts, financial professionals, and marketing professionals. However, the ease of use of generative AI technologies makes AI tools more appealing to users across the organization.



The issue here is that AI solutions typically incorporate unstructured data – including, for example, email messages and PDF documents with sensitive data – which has largely been out of reach for traditional BI tools. AI solutions may potentially be able to evade legacy security controls.

Finally, AI models tend to pull in data from various sources, not just a single repository or data lake as with BI tools, thus expanding the potential attack surface. Additionally, the dynamic nature of the data makes it more difficult to secure, compared to data coming from a single known source.

Industry-specific cloud offerings pique interest

Given these challenges, it's not surprising that highly regulated companies are interested in industry-specific cloud offerings.

Regarding selecting cloud solutions, respondents to the Foundry Cloud Computing study rated the importance of industry-specific offerings at 7.6 on a 10-point scale. IT leaders from financial services and healthcare designated these solutions as being slightly more important: 8.0 on average. Furthermore, respondents said security and governance are the most-sought-after benefits of vertical offerings.

Elements of an integrated approach

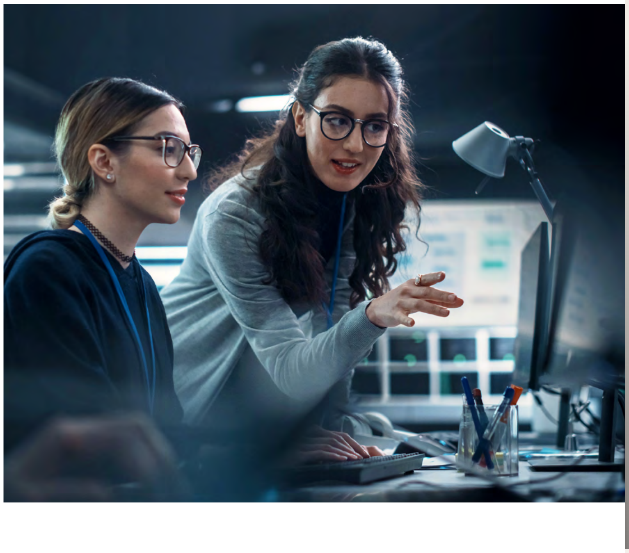
Regulated industries require a multifaceted approach to address cloud-based security requirements. The strategy should include:

- **Tokenization**, which replaces sensitive data elements with random, nonsensitive equivalents, or tokens, that have no relationship to the original data. Even if the tokens were stolen, they would have no value. A sample use case: using tokens to replace credit card or bank account numbers in financial transactions, thereby reducing the risk of fraud while remaining in compliance with PCI-DSS.
- **Format-preserving encryption** (FPE) ensures that encrypted output maintains the same format and length as the original data, such as the 16-digit structure of a credit card number. FPE enables secure data to integrate seamlessly with legacy systems and validation processes.
- **Data masking** involves obscuring specific data such that the original information is protected but the data is still usable for testing or analysis. For example, masking individual patient data makes the data usable for analytics.
- **Data anonymization** is the irreversible transformation of data to remove personally identifiable information. Governments may anonymize citizen information to conduct population studies or policy analyses while remaining in compliance with rules such as GDPR.
- **Synthetic data generation** is the creation of artificial data sets that mimic the statistical properties of real data. Companies can then use the synthetic data for testing or analysis, or to train ML models without exposing sensitive information.

These data protection methods are essential for organizations that need to use 100% of their data — wherever it resides — to securely govern it and derive real-time, predictive insights without compromise.

"All of these technologies incorporate security and privacy into the data itself, as opposed to surrounding the data with layers of security," says James Rice, vice president of security and analytics with data security company Protegrity. "That reduces friction and makes the data more easily consumable."

The key is to incorporate these capabilities into a single data platform, such as Cloudera's data and AI platform, which delivers a consistent cloud experience that converges public clouds, data centers, and the edge. Doing so provides a powerful way for regulated companies to remain in compliance while taking advantage of various third-party data services, such as real-time streaming analytics, data engineering, AI/ML offerings, and data governance solutions.



These technologies are crucial for modern data protection. When deployed within the Cloudera platform and in partnership with Protegrity, they ensure comprehensive data-centric security across the entire data lifecycle. Protegrity integrates with Cloudera services, allowing for robust policy enforcement and protected data at rest, in transit, and in use.

Finally, the platform should be built to take advantage of multiple cloud offerings, including a range of AI solutions.

Benefits of an integrated platform

A single-platform approach has multiple benefits, beginning with accelerated business innovation. Embedding security and privacy into the data opens it up for use with advanced analytics and AI technologies while remaining in compliance.

Such a solution makes centralized policy management, or a "define once, protect everywhere" approach to data security and privacy, possible, Rice says. It's an approach that can also reduce the scope of audits by removing some systems from compliance requirements.

EU courts have ruled that tokenized or pseudonymized data is exempt from GDPR requirements, provided that the process is irreversible. If a company tokenizes data about its German customers and transfers that tokenized data to the U.S., for instance, the data is no longer subject to GDPR, Rice says.

The single-platform approach also gives companies access to more data sources and services, including real-time data for analytics and AI pipelines. That includes new AI models, third-party risk services, and data sharing with third-party providers. The ability to protect data also enables increased use of outsourcing and offshoring services.

> We work with a large credit reporting agency that is saving about $60 million per year in auditing **because they've descoped hundreds of systems from compliance requirements.**
>
> – **James Rice,** vice president of security and analytics, Protegrity

Opening up these avenues makes more data available for driving strategic decision-making without sacrificing privacy or operational flexibility.

Customer results

Another benefit: the ability to descope some data from regulatory requirements to reduce costs and total cost of ownership, in some cases dramatically.

"We work with a large credit reporting agency that is saving about $60 million per year in auditing because they've descoped hundreds of systems from compliance requirements," Rice says.

Similarly, a large insurance company reduced operational expenses by 25% after implementing a new architecture on AWS that processed billions of tokenization operations per week. The technology enabled the company to explore new cloud-based analytics, AI, and ML initiatives, which have improved customer experience while maintaining security and compliance.

Another example: A large bank ranked in the global top five achieved a 126% return on investment in just eight months by deploying Protegrity's platform on AWS Cloud. The solution protects 27 billion transactions daily across more than 100 countries.

"Protegrity provides this bank with data-centric security at scale," Rice says, "simplifying security policies across jurisdictions to support the bank's need for secure analytics, ML, and AI for global business processes."

Cloudera, Protegrity, and AWS: The path forward

These results highlight the "better together" value from the partnership of AWS, Cloudera, and Protegrity. Companies that integrate Protegrity security solutions into the Cloudera data and AI platform can bring AI to data anywhere and achieve a consistent cloud experience across public clouds, data centers, and the edge. This means the heavy lifting is already done. The platform, because it is AWS-ready, ensures customers can confidently use 100% of their data in all forms.

Customers can be confident that their data will be safe, no matter where it resides, in the cloud or in on-premises data centers. Cloudera and Protegrity also work closely with AWS to incorporate new AI developments, so organizations can keep pace with the rapid evolution of AI technologies.

Highly regulated enterprises can likewise rest assured that compliance regulations won't stand in the way of innovation, having addressed all the security and privacy requirements from the start. ◆

Safely unlock the value of your business's data, even your most sensitive information, by working with **Cloudera**, **Protegrity**, and **AWS**.