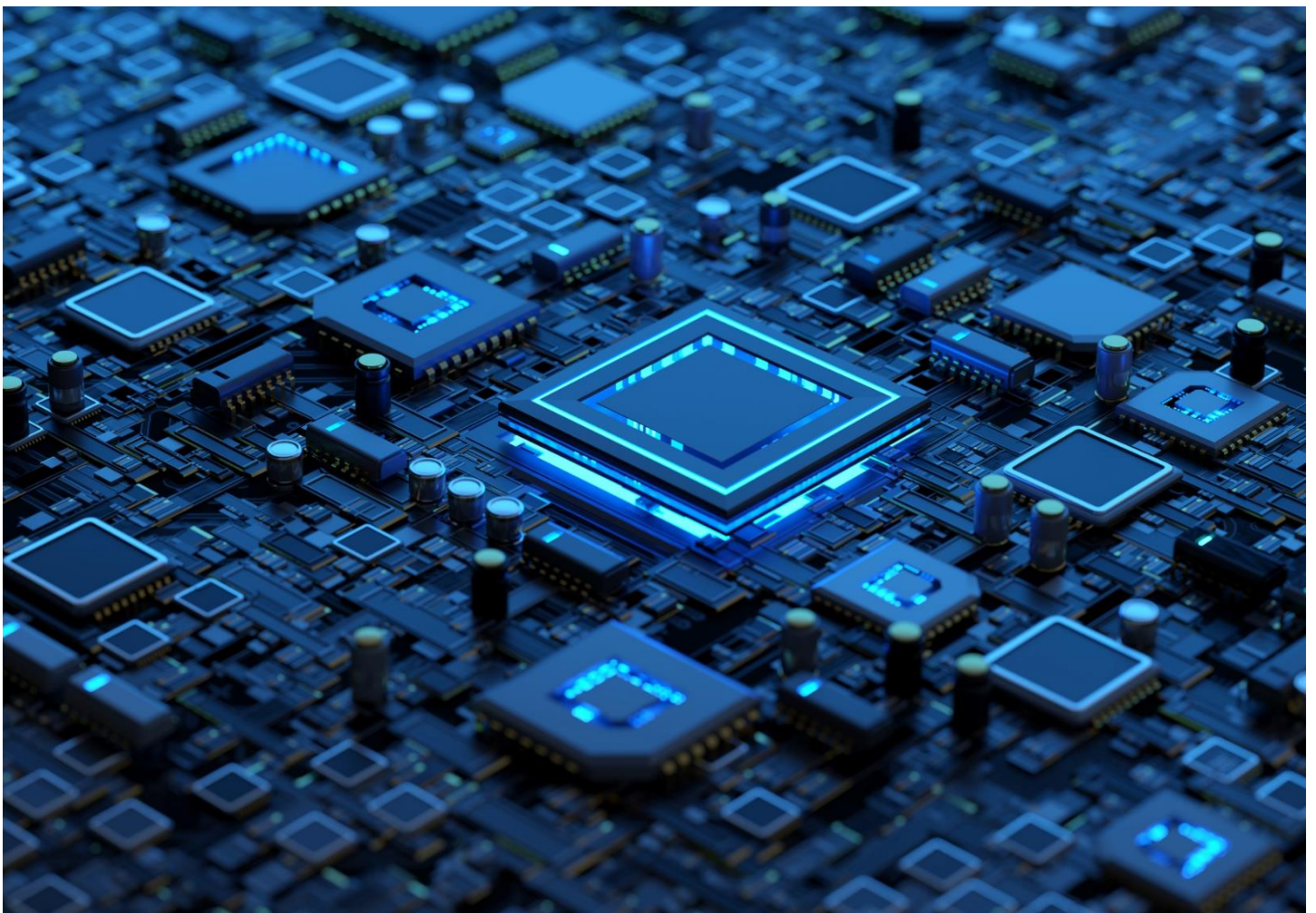


# Data governance for AI in financial services

Point of View | October 2023



## **CRISIL GR&RS analytical contacts**

### **Mohit Modi**

Global Head of Data & Analytics

[mohit.modi@crsil.com](mailto:mohit.modi@crsil.com)

### **Warren Moore**

EMEA Data & Analytics Lead

CRISIL GR&RS

[warren.moore@crsil.com](mailto:warren.moore@crsil.com)

### **Ben Lumsden**

EMEA Data & Analytics

CRISIL GR&RS

[ben.lumsden@crsil.com](mailto:ben.lumsden@crsil.com)

## Unlocking the potential of AI and Gen AI

The fast-paced nature of artificial intelligence (AI) development is exemplified by the recent rapid and widespread emergence of generative AI (Gen AI).

At CRISIL, we are excited about the unfolding potential, as we witness the sheer breadth and depth of use cases that these revolutionary technologies can unlock to reshape the way banks operate. They hold enormous commercial value and offer opportunities to enhance the customer journey.

As the AI journey to extract value commences and market players look to adopt the latest trends AI has to offer, we believe it is imperative to emphasise the importance of underpinning this evolution with effective data governance.

Anshuman Prasad, Global Head of Risk Analytics, CRISIL, recently published an article<sup>1</sup> identifying several governance challenges that financial institutions must overcome before widespread AI adoption becomes commonplace in the industry.

Ensuring the transparency of training models, holding firms accountable to high ethical standards, and establishing specific and robust governance structures are three essential areas that must be addressed for the effective implementation of AI.

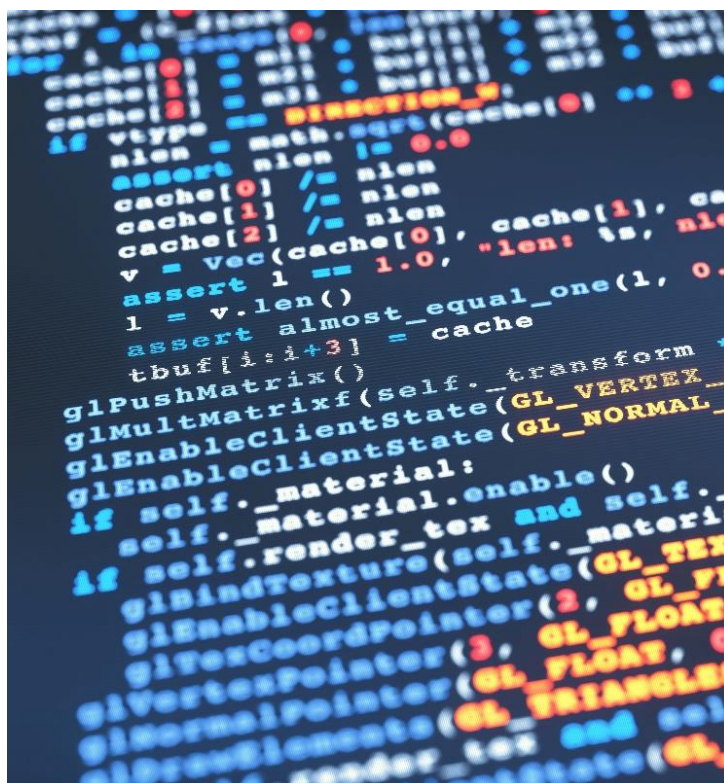
Banks are striving to overcome these challenges, expand AI capabilities, and enhance traditional governance tools and capabilities to handle the extensive and diverse datasets associated with AI techniques. It is crucial for them to align their business strategies, understand the inherent risks of adopting new technologies, and match their risk appetite with the AI use cases they want to develop.

Foundational governance pillars such as data quality, metadata management and taxonomy management will not only require more emphasis but also need to adapt to support the underlying data that fuels these capabilities. This paper delves into the symbiotic relationship between the scaling of AI and data governance, highlighting how it can unleash unparalleled potential for financial institutions.

### Overcoming key roadblocks: What can financial services CDOs do?

The journey towards achieving commercial and effective use of AI presents banks with several key milestones they must overcome. Investment in data governance has mostly taken a backseat relative to technology investment. Demand for extensive curated data continues to grow exponentially. Hence, it is essential for banks to establish strategic data governance capabilities for the successful implementation of AI, and not only for one-off use cases.

At CRISIL, we believe there are five imperatives for a financial services CDO when building the foundations of successful AI implementation:



<sup>1</sup> <https://a-teaminsight.com/blog/building-trust-in-ai-an-imperative-for-widespread-adoption/?brand=ati>



**Establish data quality standards and assurance processes:** As with all data processes and assets, ensuring data quality is crucial to maximise AI potential. Many banks still struggle with managing data quality, resulting in inaccurate or unreliable outcomes that erode business users' confidence and trust. Such outcomes can arise from various issues, including inadequate source quality, noise stemming from grammatical errors and language anomalies, and inconsistencies in data across formats, structures and content.



**Acquire high-quality training data in abundance:** After ensuring data quality, banks must focus on accessing enough training data for effective model output. Acquiring this data in abundance is a hurdle that requires more attention, especially in the financial services industry. The sector is highly regulated with a strong emphasis on data confidentiality, privacy and security. The scraping of data from public domains represents just one aspect of data privacy concern. It could prove to be a costly endeavour for firms that do not understand how it relates to data protection laws such as the General Data Protection Regulation (GDPR).

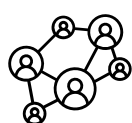
AI model output can be maximised by shifting towards a data-centric model with a focus on data quality, and not just quantity. Additionally, utilising private cloud environments can help address the privacy obstacles that large language models (LLMs) pose.



**Establish data lineage to enhance auditability:** Transparency and standardisation of the models must be established through comprehensive metadata and lineage tracking, allowing banks to understand the origin and use of the data inputs. When aggregating datasets, data users may label data inconsistently, and misaligned definitions can create confusion for business users seeking specific datasets within the necessary context for model training. Through the accurate capture and mapping of data origins, characteristics, categories, relationships and other useful information, a complete data trail can be mapped, with direct implications for key AI methodologies. Additionally, an incomplete lineage of data reduces transparency, making it more difficult for business users to understand how models have arrived at a particular outcome.

Thoroughly integrating these components into the organisation means that even though the business user may remain blind to the black box within the model, there remains a certain degree of understanding and explainability of the outcome based on the underlying datasets used.

The need for effective lineage is amplified by the sheer volume and variety of data required for certain AI use cases. This calls for a robust data versioning mechanism to ensure continuous clarity of up-to-date training data and to solidify this process through a clear data ownership structure. As AI models iteratively evolve using various data versions, maintaining a clear lineage of the data that trains each model becomes crucial.



**Evolve a shared responsibility model:** The development of a robust data governance model for successful AI implementation requires a cross-functional team. This team should assume responsibility for the full spectrum of AI potential, identifying risks and implementing effective mitigation strategies. The deployment of an expert governing body allows for a trusted team to set risk standards, audit AI systems, and guide business and development teams in complying with regulatory and organisational standards.

This ensures that the AI strategy is inherently linked to the risk appetite of the business, an essential aspect in the current uncharted territory of AI auditing and the resulting uncertainty surrounding risk management.

Effective AI implementation demands more than just top-down governance. It requires a cultural shift, organisational change, and technical expertise.



**Invest in data literacy across the organisation:** These adaptations to the business can be realised by addressing the skills gap through education and data literacy, enabling leaders to circumvent the shortage and cost of talent within this market segment. Additionally, third-party governance providers offer quick access to tried-and-tested solutions that can smooth out the learning curve across the organisation.

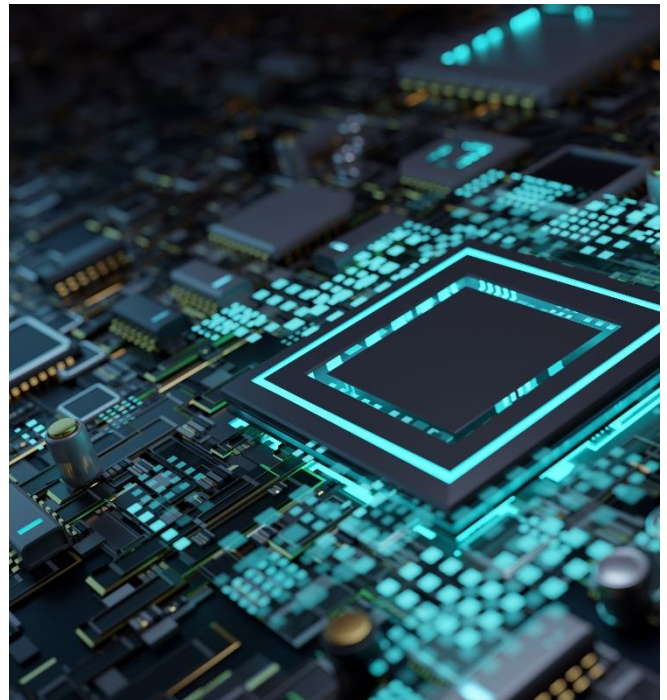
## Embracing Gen AI

As the financial services industry prepares for the powerful emergence of Gen AI, CRISIL recognises the immense potential it brings to the financial sector. With this potential, however, comes the need for pragmatism and, of course, robust governance.

Just as AI-oriented data governance foundations have been established, there is a need to proactively address the unique data governance challenges that Gen AI will present. Gen AI has the potential to introduce heightened concerns. Managing bias and fairness, pivotal in data governance for AI, now extends to applications built on these models.

Gen AI not only demands extended governance capabilities but also is capable of undertaking lower-value tasks in risk management, governance, and data collection, thereby supporting the development of the operating model it operates within.

CRISIL envisions a proactive approach to capturing Gen AI value, which would set the stage for strategic growth across the generative landscape.



## Conclusion: Seizing opportunities with data governance for AI

To summarise, CRISIL acknowledges that as AI and Gen AI redefine the financial landscape, governance will play a pivotal role in maximising their potential. By harmonising AI scaling with robust governance frameworks (including data governance), banks can confidently navigate this transformative era.

The power of AI and Gen AI can steer financial institutions towards efficiency, innovation, and responsible AI-powered growth.

## About Global Research & Risk Solutions

CRISIL GR&RS is a leading provider of high-end research, risk and analytics services. We are the world's largest provider of equity and fixed-income research support to banks and buy-side firms. We are also the foremost provider of end-to-end risk and analytics services that include quantitative support, front and middle office support, and regulatory and business process change management support to trading, risk management, regulatory and CFO functions at world's leading financial institutions. We also provide extensive support to banks in financial crime and compliance analytics. We are leaders in research support, and risk and analytics support, providing it to more than 75 global banks, 50 buy-side firms covering hedge funds, private equity, and asset management firms. Our research support enables coverage of over 3,300 stocks and 3,400 corporates and financial institutions globally. We support more than 15 bank holding companies in their regulatory requirements and submissions. We operate from 7 research centers in Argentina, China, India, and Poland, and across several time zones and languages.

## About CRISIL Limited

CRISIL is a leading, agile and innovative global analytics company driven by its mission of making markets function better.

It is India's foremost provider of ratings, data, research, analytics and solutions with a strong track record of growth, culture of innovation, and global footprint.

It has delivered independent opinions, actionable insights, and efficient solutions to over 100,000 customers through businesses that operate from India, the United States (US), the United Kingdom (UK), Argentina, Poland, China, Hong Kong, Singapore, Australia, Switzerland, Japan and the United Arab Emirates (UAE).

It is majority owned by S&P Global Inc, a leading provider of transparent and independent ratings, benchmarks, analytics and data to the capital and commodity markets worldwide.

## CRISIL Privacy Notice

CRISIL respects your privacy. We may use your personal information, such as your name, location, contact number and email id to fulfil your request, service your account and to provide you with additional information from CRISIL. For further information on CRISIL's privacy policy please visit [www.crisil.com/privacy](http://www.crisil.com/privacy).