

eBook

5 Ways to Ensure Data Quality

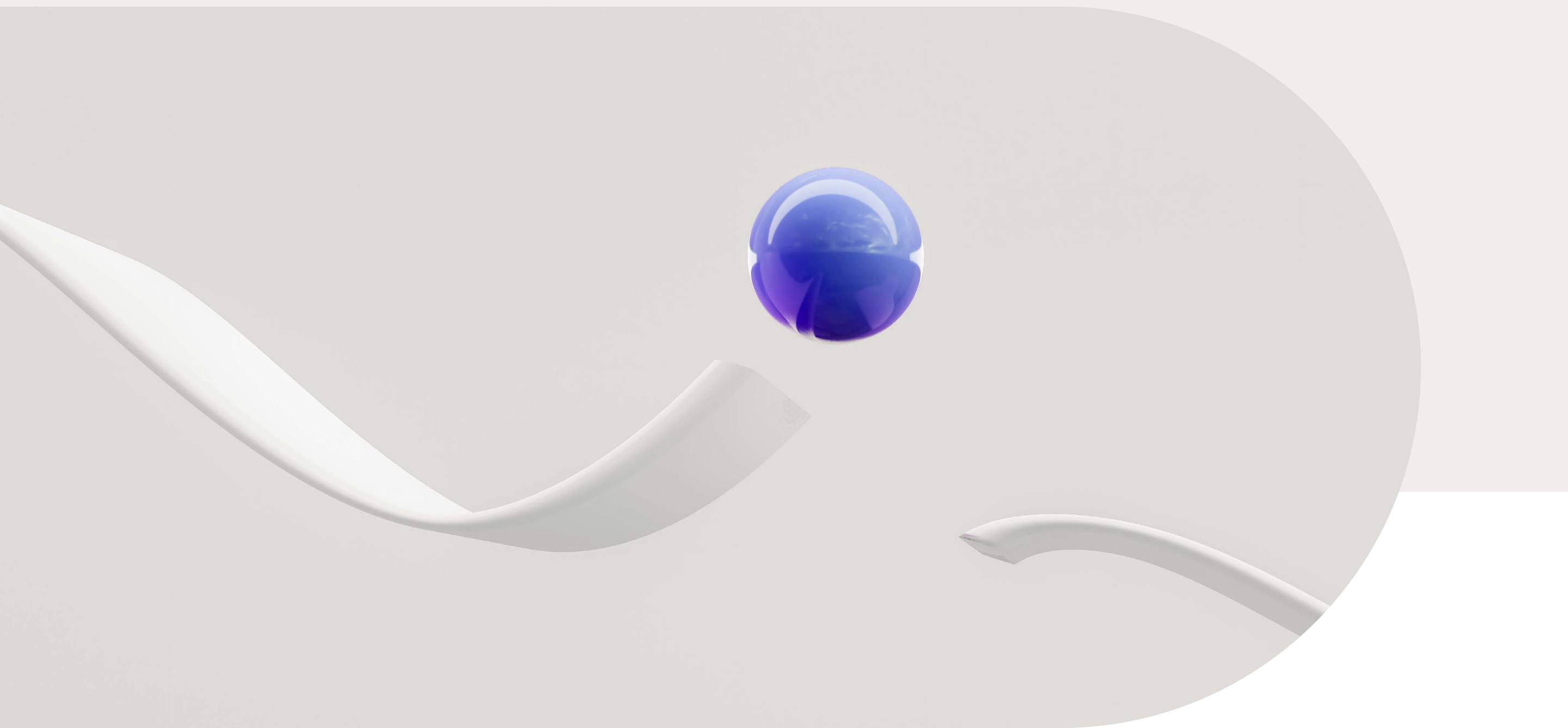
Transform Your Data from Liability to Asset



Written by

Simon Ninan

Senior VP, Business Strategy,
Hitachi Vantara

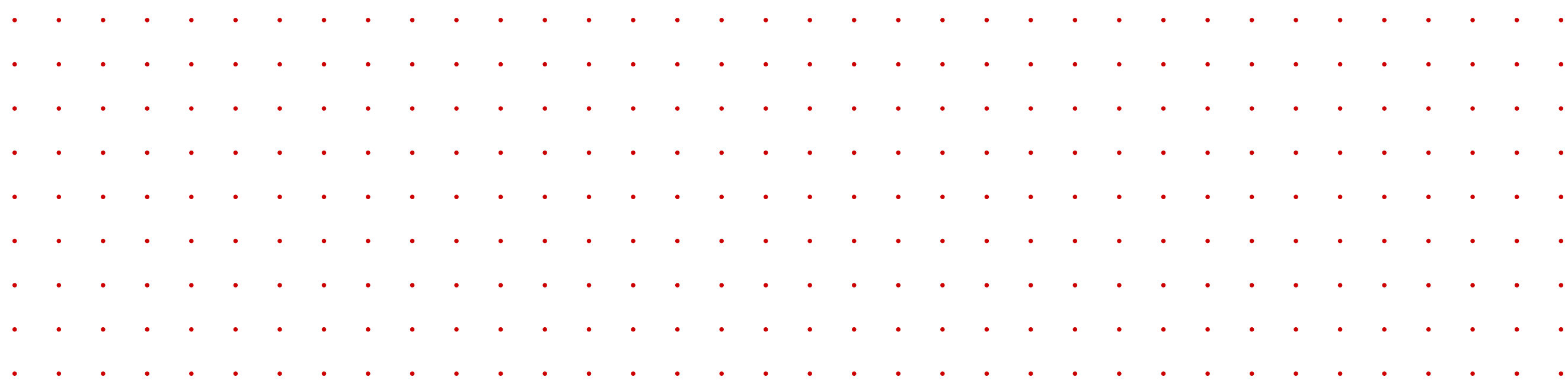


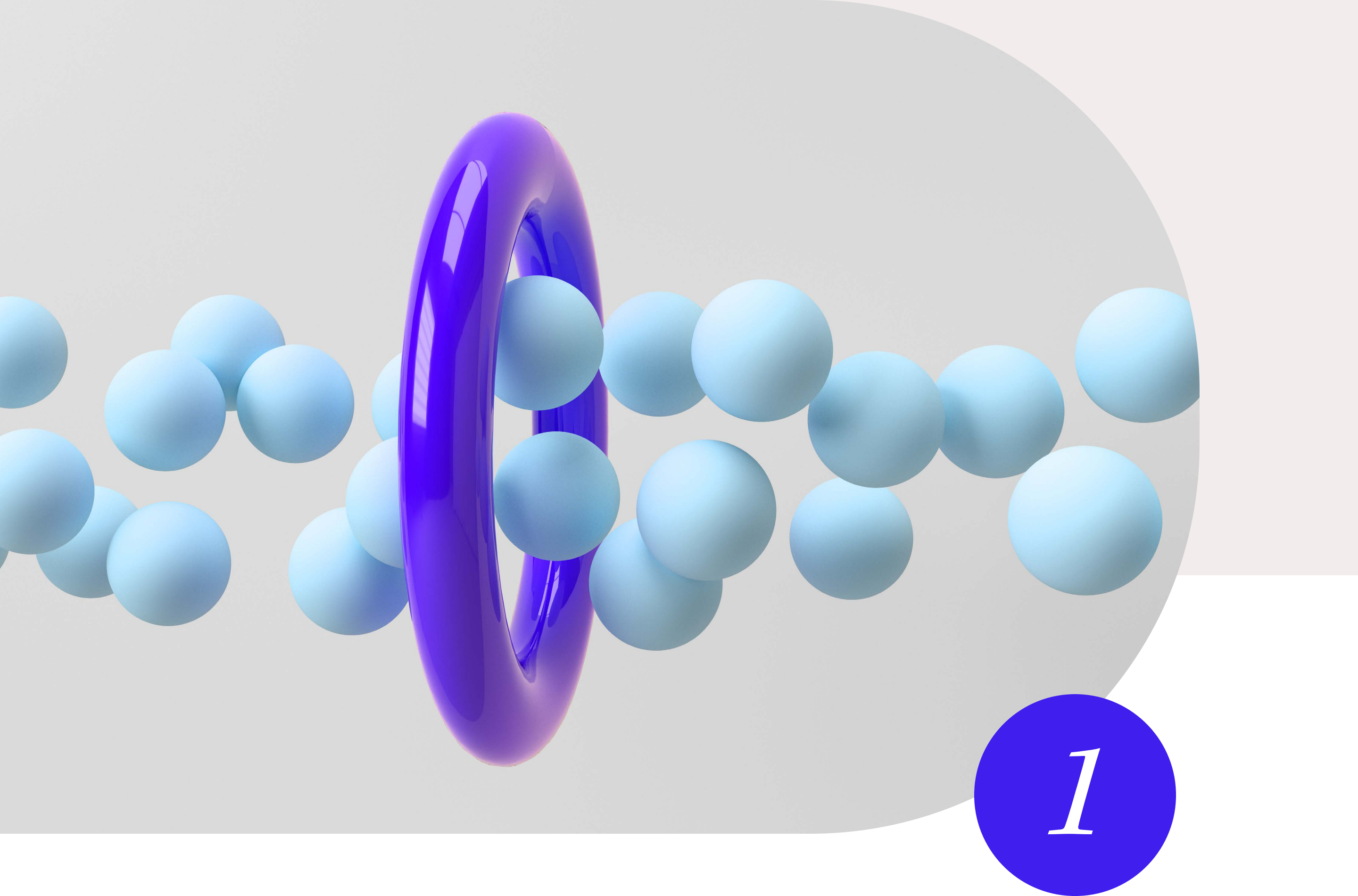
Transform Your Data from Liability to Asset

Data quality issues can emerge at any stage of the data lifecycle.

The exponential growth of data—driven by digital transformation, IoT, and artificial intelligence—has transformed data from a simple business resource into a critical strategic asset. Organizations must also navigate the complexities of AI and machine learning. This abundance of data and new tools brings unprecedented challenges: without careful management, data can quickly become a complex liability that drains resources, introduces risk, and obscures true business insights.

Read on for five essential approaches that can transform data from a potential liability into a powerful asset.





1

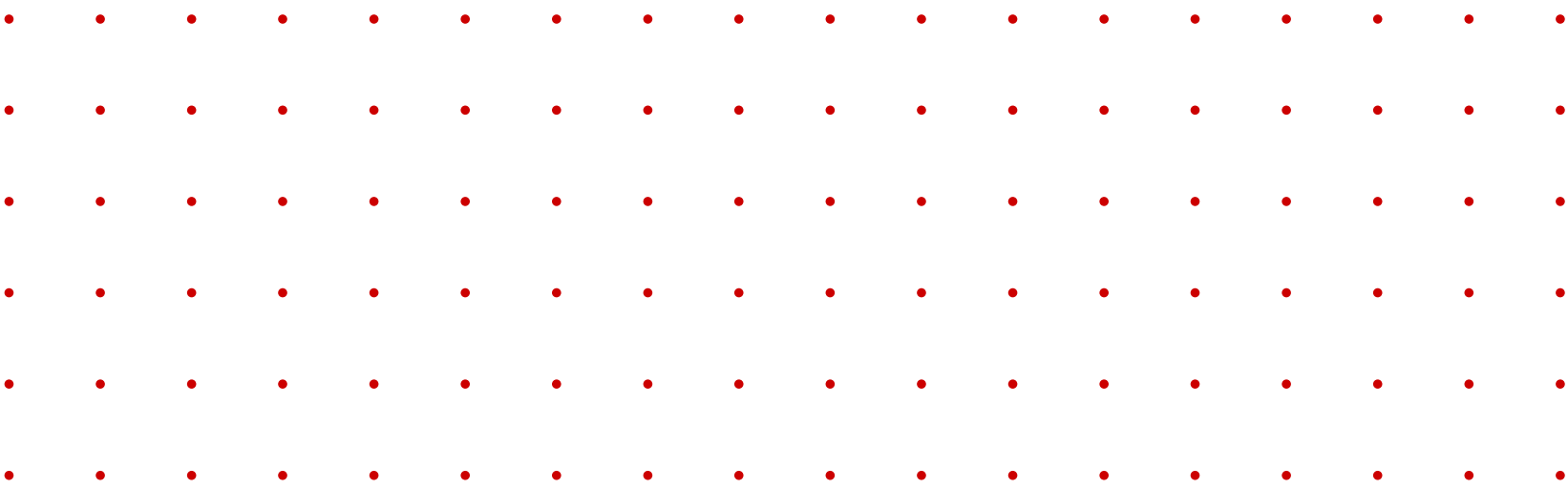
Master Your Data Inputs

Data quality operates as a cascade—problems at the source flow downstream, affecting every subsequent process. The key to prevention lies in understanding and addressing common input challenges.

Data incompleteness creates blind spots in decision-making. This isn’t just about having “enough” data. It’s about having the right data. Whether from data gaps or collection biases, incomplete data leads to faulty insights and missed opportunities. You can combat this by implementing robust data collection protocols and regular auditing data sources.

Both accidental and deliberate inaccuracies impact organizations. Consider how even well-intentioned data sources can introduce errors, as evidenced when major search engines began returning nonsense results because they were pulling from user-generated content that was meant to be satirical.

The lesson? Validate sources and implement robust verification processes.



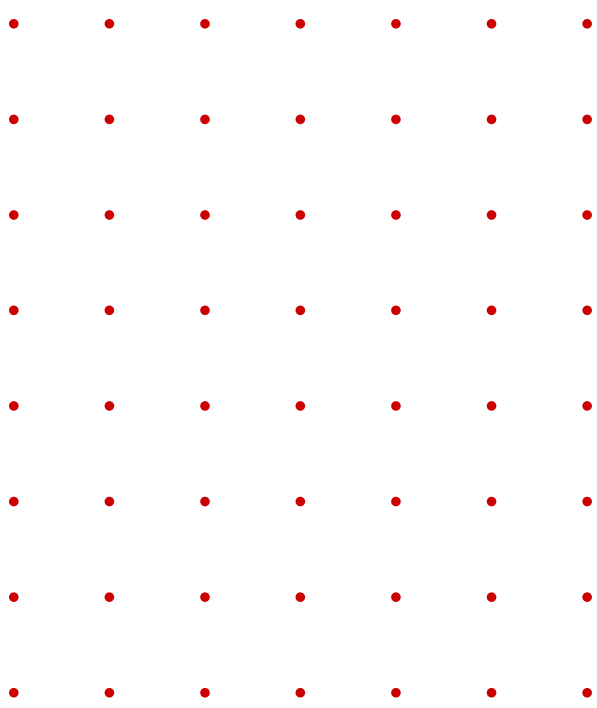


Secure Your Data Foundation

Data protection extends far beyond traditional cybersecurity.

Organizations need to think about protection across multiple dimensions:

- Comprehensive data protection that goes beyond basic security to include privacy controls and access management.
- Sophisticated privacy controls that prevent unintended data leaks and unauthorized access.
- Clear protocols for data retention and recovery to prevent costly data loss.
- Protection against emerging threats like adversarial attacks that can compromise AI systems.





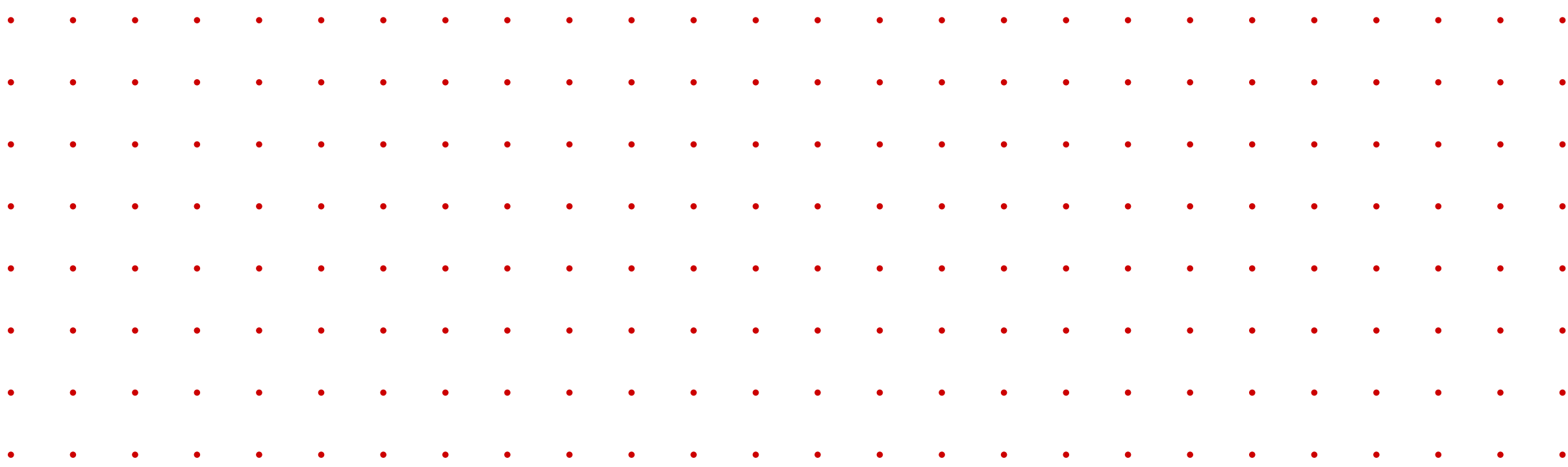
Optimize Your Data Processing

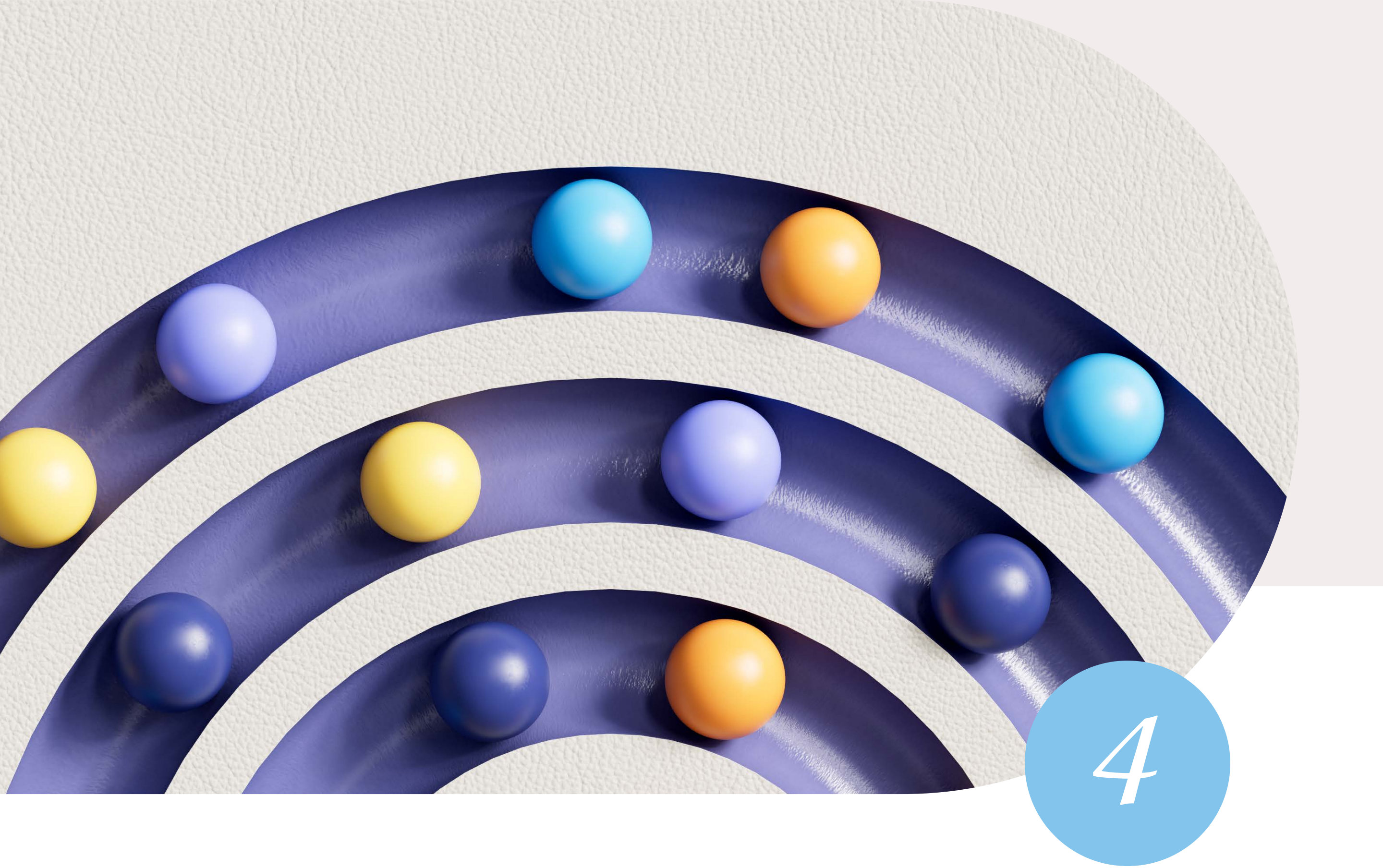
Even perfect input data can produce faulty insights if processed incorrectly.

Two areas deserve special attention:

- Model overfitting is a particular concern in AI applications. An overfit model is like an invention that performs flawlessly in the lab but fails in the real world. Success requires striking the right balance between training data and real-world application.
- Algorithms need to account for context. Organizations often assume their data processing is sound, only to discover their algorithms aren't accounting for crucial contextual variables that completely change the meaning of their results.

Poor processing can turn good data into bad insights, while intelligent processing can reveal hidden patterns and opportunities.



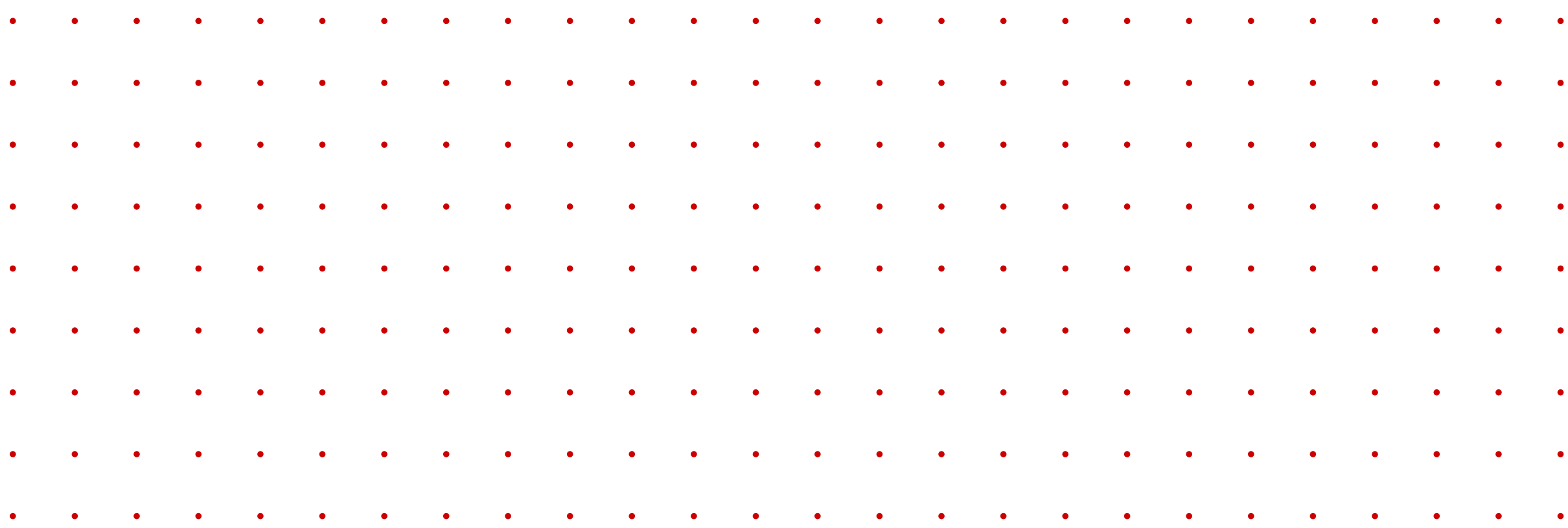


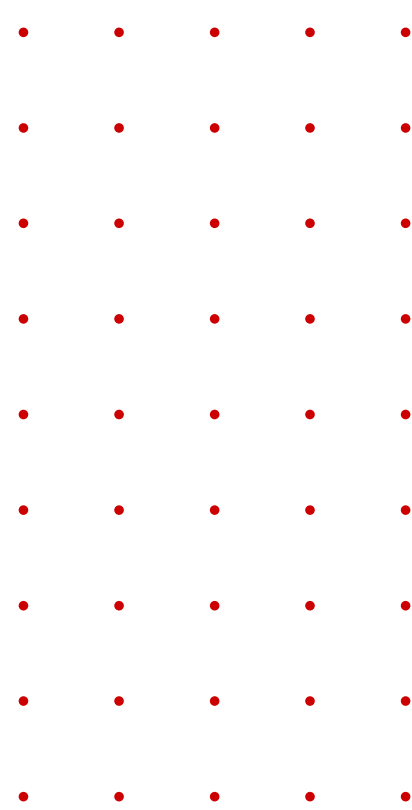
Ensure Output Quality

The true test of data quality lies in its outputs.

Enterprise data strategies should focus on:

- Building robust traceability systems that reveal how insights are generated.
- Developing clear protocols to identify and manage AI hallucinations—a growing concern as generative AI becomes more prevalent.
- Creating feedback loops to validate outputs against real-world results.





Drive Efficiency and Sustainability

Consider the following physics framework to data:

Matter (data types), Space (locations), Time (performance and latency), Life (users of data), and Law (data regulations) combine to create exponentially growing challenges.

To manage this complexity while maintaining sustainability, focus on:

- Reducing data waste through intelligent storage and processing.
- Optimizing resource utilization to minimize environmental impact.
- Implementing automation to handle growing data volumes.
- Building systems that scale efficiently with business needs.



The Path Forward

The most successful organizations think about Return on Data (ROD) rather than just ROI. Progress from basic data management to advanced orchestration and monetization makes data quality the competitive advantage.

These five steps serve as a stairway to value. Each one builds on the last to create a comprehensive approach to data quality. By addressing these five key areas, organizations can transform data quality challenges into opportunities for innovation and growth.

Take the Next Step

Clear the data silo chaos of hybrid cloud. Hitachi Vantara's Virtual Storage Platform One (VSP One) brings you ultimate data clarity by harmonizing all your block, file and object data into one simple platform. With data intelligence at your fingertips, you're ready to transform your business for infinite growth.

Explore Hitachi Vantara's solutions and contact us to discuss a hybrid cloud strategy to meet your needs.

Clear the Chaos →

One Hybrid Cloud Data Platform



State of Global Data Infrastructure

Building a robust data infrastructure is critical to meet AI's intense demands.

AI adopters are raking in revenue—at least, the ones who have laid the groundwork for security, data quality and implementation. Avoid the surprising mistakes most leaders make by taking advantage of this survey's findings. Download our State of Data Infrastructure Global Report to unlock key strategies to lead in an AI era.

Read the Report →

State of Data Infrastructure Global Report



About Hitachi Vantara

Hitachi Vantara is transforming the way data fuels innovation. A wholly owned subsidiary of Hitachi Ltd., we're the data foundation the world's leading innovators rely on. Through data storage, infrastructure systems, cloud management and digital expertise, we build the foundation for sustainable business growth.

© Hitachi Vantara LLC 2024. All Rights Reserved. HITACHI and Lumada are trademarks or registered trademarks of Hitachi, Ltd. All other trademarks, service marks and company names are properties of their respective owners. HV-GBS-EB-5-Way-To-Ensure-Data-Quality-21Feb25-A

Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA
hitachivantara.com | community.hitachivantara.com

Contact Information
USA: 1-800-446-0744
Global: 1-858-547-4526
hitachivantara.com/contact