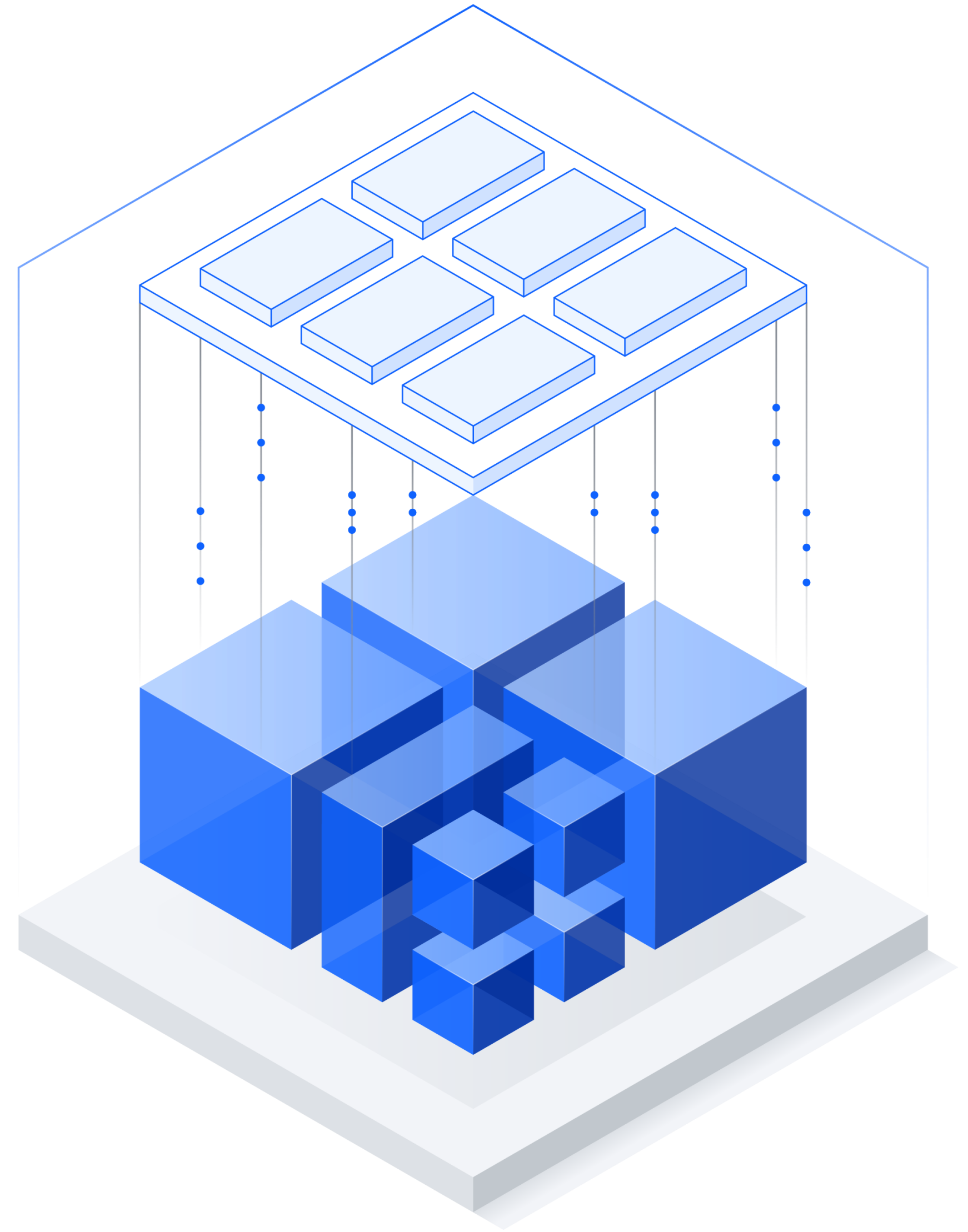


IBM Manta Data Lineage

Gain deeper visibility into the provenance of your data and its journey from source to end use

—



Data ecosystems are **increasingly hard to manage**. Information is spread across multiple technologies, teams, networks, and regions.

Data structures and teams are **continually changing**, and data volume is constantly growing.

Quickly understanding data's **journey from source to end use** is critical but has become nearly impossible.



Fundamental challenges inhibit enterprises from becoming data-driven

45%

Data access

Multiple locations, clouds, applications and data silos hinder timely delivery of data to the right users

Only **45%** of enterprise data useful for analysis was analyzed or fed into AI.¹

80%

Data quality

Poor data quality and consistency lowers trust in data

80% of business executives surveyed do not trust their data.²

26%

Data literacy

Lack of a consistent understanding of data across the organization can hinder data utilization

26% of respondents surveyed cite expanding companywide data literacy as a high or critical priority.³

18%

Data protection

Inadequate data protection increases the risk of non-compliance

Only **18%** of respondents surveyed report that their firm excelled at protecting organizational data.⁴

61%

Data lineage

Lack of visibility into the journey of data is a barrier to generative AI

61% of respondents surveyed think concerns about data lineage and provenance will be a barrier to generative AI adoption⁵

¹ “How Is Data Utilization Rate Changing in a Time of Uncertainty”, IDC, <https://www.idc.com/getdoc.jsp?containerId=US49286422>
² “The Data-Powered Enterprise”, Capgemini, <https://www.capgemini.com/insights/research-library/the-data-powered-enterprise/>
³ “Data Literacy Matters to your company’s success”, Forrester, <https://www.forrester.com/report/data-literacy-matters-to-your-companys-success/RES149295?objectid=RES149295>
⁴ “Forrester’s Data and Analytics Survey – 2022” – IBM authorized login required to access the survey contents.
⁵ “CEO’s guide to generative AI”, IBM Institute of Business Value Study, 2023, <https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/ceo-generative-ai/platforms>

Data blind spots and lack of visibility into the journey of data has consequences...

Decreased trust in analytics and AI

If organizations can't fully explain how data was collected or verify its origins, they can't truthfully respond to audit inquiries or trust data to produce accurate and quality analytical and AI outcomes.

Increased number of data incidents

Data engineers are limited and have a limited ability to assess the end-to-end impacts of to-be implemented changes making impact and root cause analysis demanding and nearly impossible.

Increased risk of non-compliance

With the growing number of regulations and policies, just one data incident can ruin a company's reputation and cost hundreds of millions of dollars in damages.

Slow delivery of predictive insights

Data engineers spend significant time on impact analysis due to lack of understanding of the data environment causing delays in the development and delivery of new features.

Application modernization delays

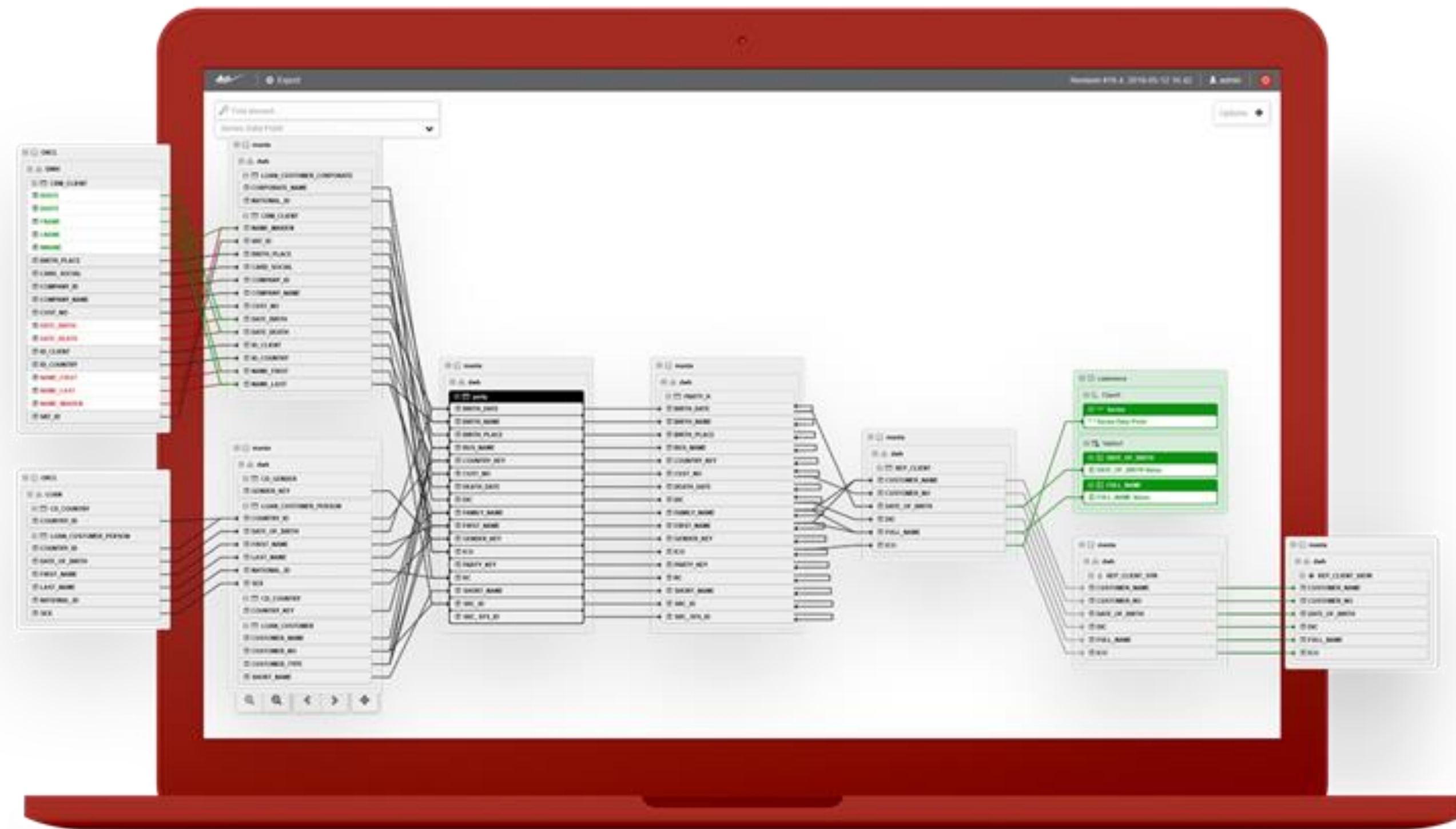
Renovating systems that have been developed, enhanced, and changed over decades is challenging. These systems have multiple sources and were built by many contributors.

Shortage of data engineering talent

Data engineers have become more critical than any other role as they oversee data pipelines and data structures. This requires a unique skill set making good engineers hard to find and keep.

Having a strong **data lineage foundation** maximizes business value from data, analytics and AI

Lineage provides **transparency** into enterprise data flow sources, targets, and transformations



IBM Manta Data Lineage

↓

Benefits



Improved incident prevention

Diagnose issues and discrepancies in data, analytics and AI



Ensure regulatory compliance

Meet regulatory requirements to avoid non-compliance penalties



Faster and more efficient migrations

Removal of manual tasks. Auto scan to-be-modernized systems



Established trust in data

Show data's journey clearly for all levels of the organization



Improved change management

Gain greater visibility into the impact of proposed changes



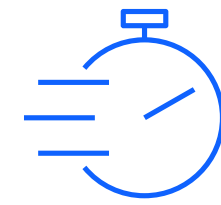
Retention of data engineering talent

Automate routine and manual tasks and enable self-service

IBM Manta Data Lineage



Advantages



Automation

Advanced, low touch, back-end engine

Speed at scale

- Full lineage in hours (not weeks or months)

Coverage

- Out of the box connectors
- 3rd party data catalog integrations *
- Integrated with IBM Knowledge Catalog
- Support for Information Governance Catalog *

Accuracy and completeness

- Entire code base analyzed by default
- Removes human error

Detailed collection

- Column and attribute level, including transformation logic



Innovation

User interface built with a lineage-first mindset

Step-by-step flow analysis

- Full understanding of your data's path

Color coding

- Make sense of all the noise

Dynamic filtering

- Present at user's desired level of detail

Historical lineage*

- View changes in pipelines over time

Indirect lineage*

- View filters that impact flows

Active tags*

- Sticky notes to label assets

Note:

* These features are available on versions: Manta Automated Data Lineage for IBM CP4D and IBM Automatic Data Lineage (Standalone, non-containerized)

