TDWI WEBINAR

Drive Faster Insights from Your Data

David Stodder
Senior Director of TDWI Research
for Business Intelligence

GUEST SPEAKER
Thomas Ridings, Senior Director of Product Management, Matillion

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

Webinar Sponsor

TDWI WEBINAR

Drive Faster Insights from Your Data

TDWI Research Perspectives on Trends in Data Engineering Productivity and Collaboration with Best Practices and AI Innovation

David Stodder
Senior Director of TDWI Research
for Business Intelligence

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

Demand for Faster Data Insights – and Concerns

Maximizing the value of data

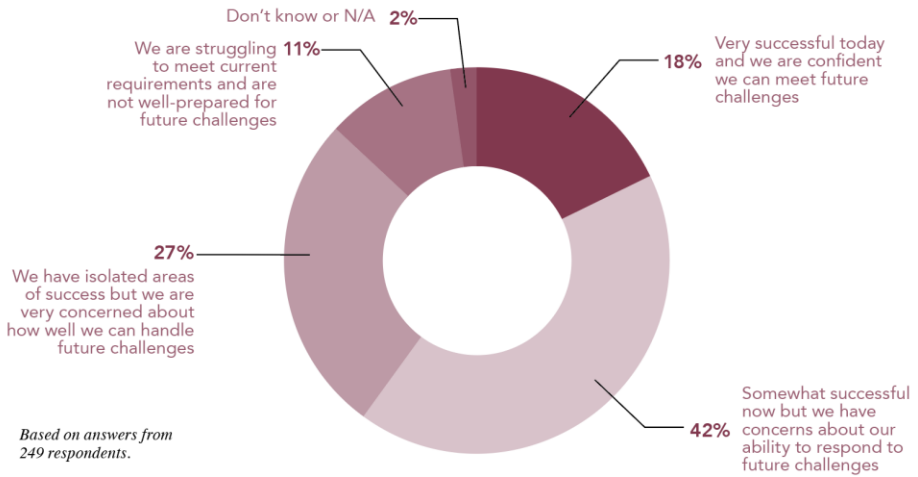
- Democratized users need contextual data to drive operational decisions, customer engagement, risk management, regulatory adherence, and more
- Data-rich applications and enriched transactions
- Automated decisioning in apps

Analytics and AI

- Data for model development and testing
- Data for LLMs & generative AI
- Augmentation of BI and reporting

Moderate success currently: Concerns about responding to future challenges

Overall, how successful is your organization in using data systems to enable users to gain faster data insights? How confident is your organization that it can succeed as data and workload volumes, speed, and/or complexity increase?



Research data source: TDWI Q1 2024 Best Practices Report. See https://bit.ly/43j2ajp

Data Engineering: The "Engine" of Faster Insights

Data engineering: Central to provisioning data for BI, applications, analytics, and AI

- Discover, collect, and prepare data
- Design and populate data warehouses and data lakes
- Develop and code data pipelines, connectivity, and transformations (ETL/ELT), often with significant technical expertise
- Validate data; manage, modify, and maintain data integration through multiple iterations
- Big data and data engineering services market size to be worth $169.9 billion by 2029, up from $75.55 billion in 2024 (CAGR of 17.6%) [Source: Market Data Forecast, www.marketdataforecast.com]

Data demands put pressure on data engineering

- To be efficient and effective as data demands escalate
- 50% of organizations surveyed say project teams spend more than 61% of their time on data integration, pipeline development, and preparation, with 16% spending over 81%
- Pressure to provision data for advancing analytics and AI
- Pressure to partner effectively in teams with data scientists, analysts, developers, and low- and no-code users

Data Pipelines, Connectivity, and ETL/ELT Challenges

Data governance and security

- Difficulty monitoring exposure risks; interest in automation

Identifying and eliminating unnecessary workloads

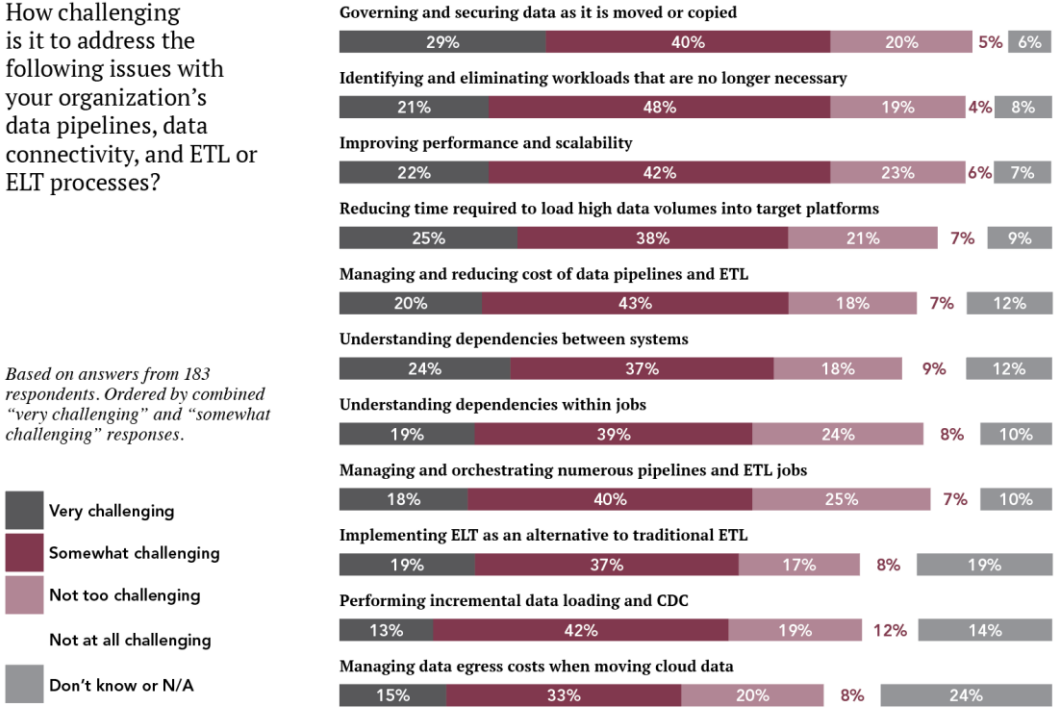
- Problems with overall visibility to reduce costs and resource usage
- 63% say managing and reducing costs is challenging

Improving performance, scalability, and orchestration

- Research shows challenges
- Challenges understanding dependencies

How challenging is it to address the following issues with your organization's data pipelines, data connectivity, and ETL or ELT processes?

Based on answers from 183 respondents. Ordered by combined "very challenging" and "somewhat challenging" responses.

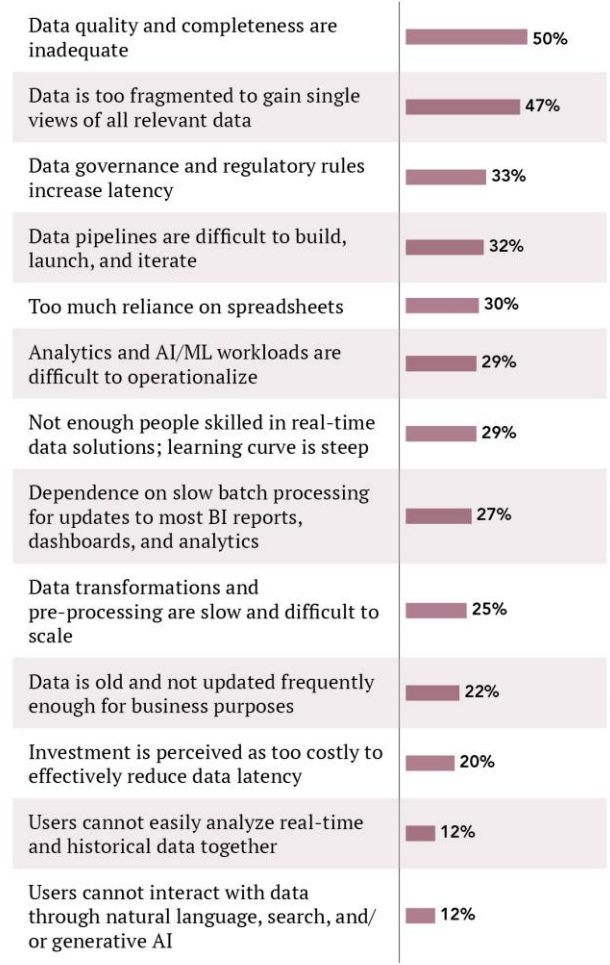


Research data source: TDWI Q1 2024 Best Practices Report. See https://bit.ly/43j2ajp

Top Issues Hindering Timely Data Insights

What are the biggest issues hindering your organization from enabling users to gain timely data insights for answering business questions and solving problems?

Based on answers from 245 respondents, who were asked to select at least their top five concerns.



Research data source: TDWI Q1 2024 Best Practices Report. See https://bit.ly/43j2ajp

Major Challenges Hindering Data Engineering Goals

Disconnected and siloed data integration
- Data accessibility barriers, especially for AI
- Too many tools and jobs to track and manage, including from a data governance perspective
- Technical expertise barriers for low- and no-code users
- Data lineage and documentation is incomplete and inconsistent
- Difficulty managing and reducing costs

Scalability and orchestration challenges
- Difficulty managing and locating all the pipelines, connectors, and ETL/ELT jobs
- Manual maintenance of manual coding: Time and complexity fixing errors and making updates

Support for AI growth, including generative AI
- More users want analytics and AI for data-driven decisions
- Feature engineering and vector embedding
- Selecting and validating data for LLMs and generative AI
- Data lineage and documentation for AI explainability

Performance and agility issues
- Difficulty performing data engineering responsibilities quickly to accelerate business use
- Need for more automation to reduce manual coding
- Accelerating costs create resource limitations

Poll Question

To enable faster data insights, which of the following currently your biggest data engineering challenge?

Technical challenges leading to delays in building and managing data pipelines, connectors, and ETL jobs
Having the talent and expertise to keep pace with accelerating data demands, including AI requirements
Data and tool silos make it difficult to centrally manage and orchestrate data pipelines, connectivity, and ETL, and to manage code
Team collaboration is lacking among data engineers and no-code and low-code users
Cost issues; we cannot manage and reduce costs as data demands escalate

Solving Challenges: Trends and Priorities

Streamlining data insights with tech modernization

- AI-driven automation; automation of processes for putting data pipelines into production
- Potential of generative AI for improving productivity
- Improving reuse by setting up pre-built connector libraries
- Easier interfaces for less technical users

Centralizing data integration management

- Orchestration of jobs to manage dependencies, address bottlenecks, and improve scalability
- Enabling rationalization of data pipelines, ETL, and connectors to ensure business value
- Data lineage and documentation for data and AI governance
- Reducing silos to enable faster and more complete tech modernization (rather than piecemeal)

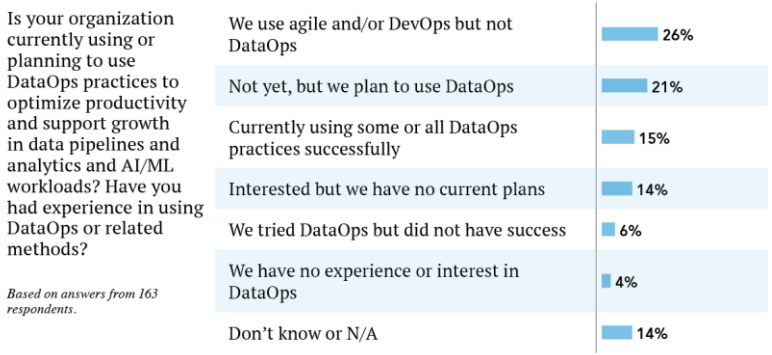
TDWI research: Importance of AI-driven automation

- Data quality, validation, and enrichment: 40% very important, 27% important but not a current priority
- Data management (including for data warehouse and data lake): 36% very important, 26% important but not a current priority
- Data governance and regulatory adherence: 33% very important, 31% important but not a current priority
- Data ingestion, pipelines, preparation, and transformation: 30% very important, 29% important but not a current priority
- DevOps, programming, code generation, and container orchestration: 21% very important, 32% important but not a current priority

Optimizing Productivity with DevOps and DataOps

Collaborative data culture for more effective teams

- DevOps and DataOps for data engineering and code management
- DataOps: Building on agile and DevOps to provide a structure similar to software development life cycles but for data processes
- Data observability: Important to move beyond limited, siloed monitoring; visibility to see how problems impact business
- Leadership: Key for encouraging cross-functional collaboration



Research data source: TDWI Q3 2023 Best Practices Report

TDWI WEBINAR

Thank You

David Stodder
Senior Director of TDWI Research
for Business Intelligence
dstodder@tdwi.org
@dbstodder

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

Thomas Ridings

Senior Product Director
Matillion

research

tdwi
Transforming Data With Intelligence™

Matillion is the **Data Productivity Cloud**

Empowering teams to make their data business-ready faster

matillion

The landscape is more complicated

More data, from ever increasing sources, with different data formats, structures and approaches

The infrastructure (scaling, integration, security)

Budgets aren't keeping up

Overall demand shortage for data skills is increasing the cost of recruitment and staffing

Infrastructure and software costs are rising.

There's growing demand

As businesses are demanding more from their data, the demand increases.

50% increase in demand year-on-year.

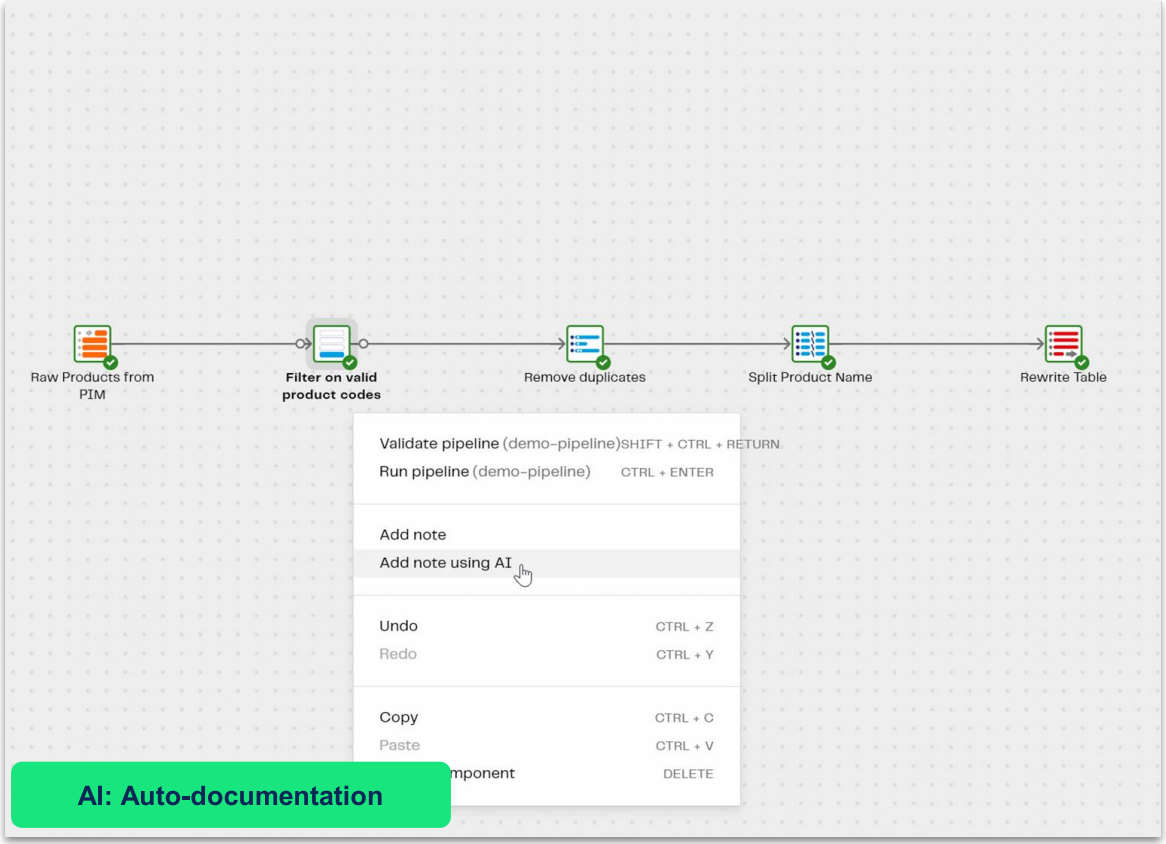
Teams and organizations need to be equipped to be **productive** with data

#1 Design for discovery and learning

Documentation is key to discovery, make your pipelines readable

Multiple tools and systems increase the learning time, keep it simple

Column level lineage can accelerate learning

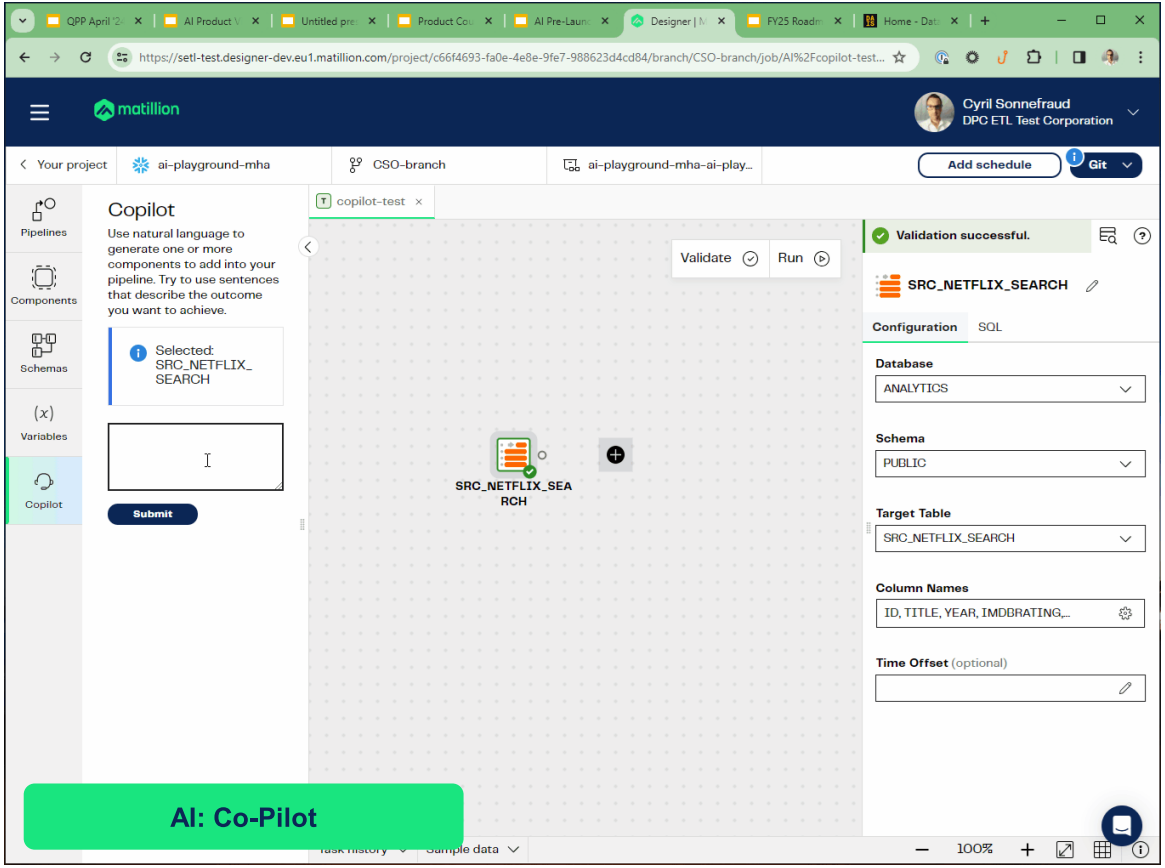


#2 Make your development time productive

Build what's unique to your business, leverage off-the-shelf or open-source where possible

Low-Code and High-Code development can co-exist

Reduce barriers so others can contribute but ensure there are guardrails



AI: Co-Pilot

matillion

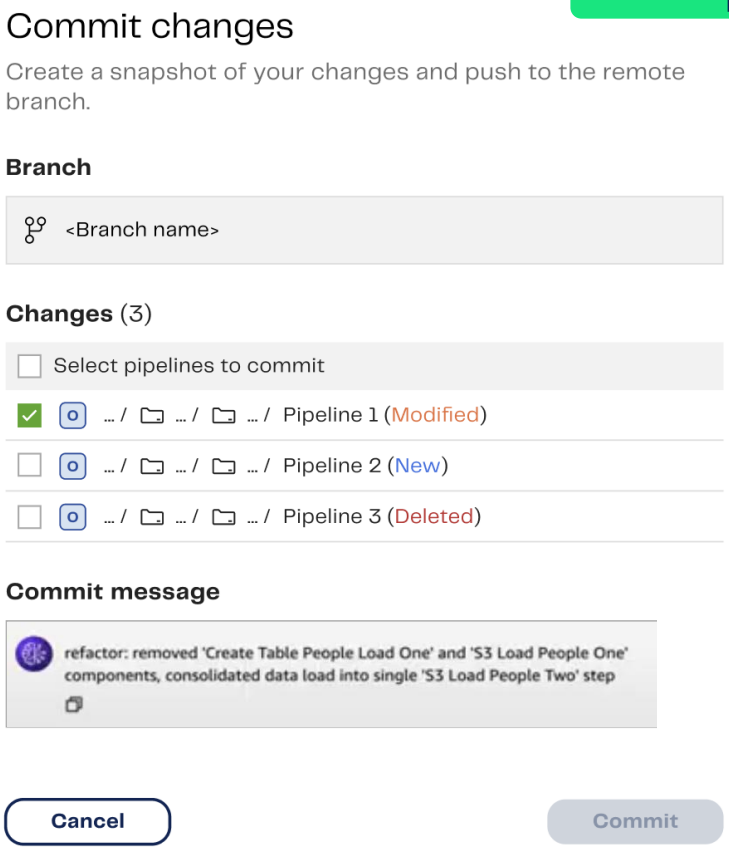
#3 Optimise for management

Ensure you have strongly versioned system

Automate your tests, mandate them before deployment

Work to identify and remove what slows you down from delivering value

AI: Auto-generating commit messages



matillion

Fireside Chat Discussion

Thomas Ridings
Senior Director of
Product Management
Matillion

David Stodder
Senior Director of Research
Business Intelligence
TDWI

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

TDWI WEBINAR

Audience Q&A with Speakers

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

Speaker Contact Information

Thomas Ridings
Senior Director of Product Management
Matillion
Thomas.ridings@matillion.com

David Stodder
Senior Director of Research
Business Intelligence
TDWI
dstodder@tdwi.org

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

Thank You to Our Webinar Sponsor

Follow Us on Social Media

tdwi.org/linkedin

tdwi.org/x

tdwi.org/facebook

tdwi.org/youtube

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

Thank You for attending!

tdwi TRANSFORMING DATA WITH INTELLIGENCE™