StreamSets
A SOFTWARE AG COMPANY

EBOOK

Five Challenges Limiting the Impact of Transformative Analytics

streamsets.com

StreamSets
A SOFTWARE AG COMPANY

Five Challenges Limiting the Impact of Transformative Analytics

Contents

The evolution of analytics: From Advanced to Transformative

We're in an era of game-changing technologies. From machine learning and generative AI to large language models and neural networks, technology today has the power to revolutionize how enterprises operate.

For the field of data analysis, this has meant a rapid evolution of analytics from *advanced* to *transformative*. Now, enterprises have the power to uncover deep insights from their data that could change how they do business.

But they can only do this successfully with an **agile data infrastructure** in place. The reality is that most organizations lack the foundation to enable transformative analytics. And while tools such as Hadoop promised much, they ultimately weren't fit for purpose and have failed to deliver. This means the impact of transformative analytics has been limited to a small number of businesses that have highly skilled data professionals.

It doesn't have to be this way. The combination of cloud with modern data destinations and analytics engines, and the broader availability of generative AI tools, offers a new opportunity to enable transformative analytics for all.

If you are responsible for managing data infrastructure, read on to learn about five things holding your transformative analytics projects back and how an agile data infrastructure will unlock success.

What is Transformative Analytics?

Transformative analytics goes beyond traditional business intelligence approaches to uncover deeper insights, patterns, and predictions that inform a strategic vision to transform, anticipate risk, and grasp opportunities.

Typically undertaken by data scientists and involving large amounts of hand-coding data pipelines and manual data processing, transformative analytics techniques include predictive analytics, pattern matching, forecasting, semantic analytics, cluster analytics, and complex event processing.

Organizations use transformative analytics for a variety of complex and valuable tasks:

- **Diagnostic Analytics** – Why has something happened?
- **Predictive Analytics** – What is likely to happen in the future?
- **Prescriptive Analytics** – What is the best course of action?
- **Cognitive Analytics** – Applying human-like intelligence to tasks to get smarter and more effective over time.

Transformative analytics projects are often exploratory, but the result is an insight or prediction that can drive business strategy and act as the catalyst for achieving critical business outcomes.

What can Transformative Analytics help organizations achieve?

Transformative analytics has applications in every industry. **Here are six use cases:**

1. **Predictive maintenance in manufacturing**
   Manufacturers want to avoid delays. By analyzing performance metrics and data from key equipment, manufacturers can predict where in the lifecycle any given machine is. This insight allows them to forecast timelines for maintenance or replacement, so they can plan ahead, reduce maintenance costs, and slash downtime.

2. **Pre-clinical drug discovery**
   In drug development, the pre-clinical drug discovery stage is where researchers screen thousands of compounds to identify those that show potential. Transformative analytics processes vast volumes of biological and genetic data to help identify drug targets and optimize existing ones.

3. **Predictive planning in logistics**
   Logistics providers are under pressure to cut fuel costs and optimize routes. With transformative analytics, logistics providers can analyze years' worth of data on routes, fuel consumption, and vehicle usage to plan efficient routes in real time, eliminate wasted capacity, and determine if additional drivers and vehicles are needed.

4. **Marketing-mix modeling**
   Marketing teams can use transformative analytics to determine the effectiveness of spending by channel. This is done by linking marketing investments to other drivers of sales, as well as variables such as seasonality, promotional activity, and competitor campaigns. This helps teams understand the impact of marketing tactics and provides the ability to forecast the future impact of marketing tactics.

5. **Propensity to buy in eCommerce**
   Transformative analytics combines data sets on purchases with online behavior (e.g., browsing history or social media activity) to correlate the effectiveness of campaigns. This predicts which customers are more likely to buy products or services and the channels most likely to reach them to give the best chance of driving revenue.

6. **Automate mundane tasks**
   Transformative analytics can prompt generative AI to automate mundane tasks, improve decision-making with data-driven insights, and unlock new levels of creativity. With generative AI, anyone can effortlessly connect their daily analytical systems just by asking, which helps to streamline workflows and facilitate digital transformation.

Five challenges limiting Transformative Analytics

Many organizations lack mature enterprise data infrastructure. At the same time, data is becoming more complex – with more of it and more data types to consider. This combination can cause five common challenges that limit transformative analytics.

Challenge 1

The lack of a centralized "data mission control" makes it hard for organizations to gain deep observability.

- A lack of observability makes it almost impossible for organizations to manage data pipelines and infrastructure across cloud, on-premise, and legacy environments.
- When organizations don't have deep observability over data architecture, they can't understand how systems are connected and how data flows across the enterprise.
- This means organizations are prevented from accessing data locked away in legacy systems – this data can be the "secret sauce" that is invaluable for transformative analytics.
- They are also unable to detect, control, and adjust to data drift – the unexpected and undocumented changes to data pipelines supplying transformative analytics.
- This stops teams from detecting when there are changes to data quality, sizing, throughput performance, error rates, and private or sensitive information leakage.

78% of organizations say a single platform that can handle the complexity of data spanning across cloud and on-premises worlds would be a huge benefit. [1]

[1] Creating Order from Chaos: Governance in the Data Wild West

- Organizations need a single, smart interface to act as a centralized data mission control for managing data pipelines and infrastructure.

Challenge 2

Teams are held back by data integration friction, which stops the flow of data from source to destination.

- Data pipelines are the plumbing that enables effective transformative analytics, but the lack of clear central mission control means teams can't identify and capture changes to data or data infrastructure, resulting in broken pipelines and data not being delivered downstream.
- Processes like change data capture (CDC), which continuously identifies and captures changes to data schema, structure, and dependencies, can be impossible due to the chaos of modern data ecosystems.
- These breakages are fatal for transformative analytics – imagine a machine learning model trying to predict customer purchasing habits based on age and income. This can't happen if the pipelines break, leaving the model unable to reliably predict the likelihood of purchase.
- Organizations need to develop malleable infrastructure based on resilient data pipelines that embed the ability to resist change and overcome data integration friction.interface to act as a centralized data mission control for managing data pipelines and infrastructure.

87% of organizations say data pipelines break at least once a year. Of these, more than a third (36%) say their pipelines break every week, and 14% say pipelines break once a day. [2]

[2] Lifting the Lid on the Hidden Data Integration Problem

Challenge 3

Data transformations are ineffective, making it difficult to deliver analytics-ready data.

- Data transformation is the process of converting, cleaning, and structuring data into a usable format – it's a key step in Extract, Load, Transform (ELT) data integration processes.
- Data engineers often work with large datasets and complex transformations, but these require efficient methods for processing and analyzing data.
- SQL is often used to manipulate data but can be tedious and time-consuming to use and struggles to handle the complexity of transformations required.
- Teams are often blind to the inconsistency and incompleteness of data sets, leading to errors and inaccuracies in transformative analytics.
- Organizations need a modern data transformation tool to operationalize data transformation to make data machine-readable and analytics-ready. This allows the business to make decisions based on accurate, reliable insights.

38% of organizations say the variety of data formats, both structured and unstructured, is the main cause of data friction. [3]

[3] Lifting the Lid on the Hidden Data Integration Problem

Challenge 4

Data has become a challenge to visualize, making it difficult to describe data logic.

- Rising complexity in environments without a data mission control is having a knock-on effect that results in a lack of visibility over data logic.
- This prevents organizations from modeling and describing data elements in detail, leaving them blind to data entities, attributes, and relationships.
- It can also cause lost or orphaned data, as departments or people trusted to maintain data reorganize or leave, and ownership becomes hazy or disputed.
- Organizations are running the risk of blindly feeding data into analytics tools without knowing how this could impact other efforts or whether more accurate data is available to prevent the garbage in, garbage out truism hitting their projects.
- Organizations need data infrastructure that can visually describe data logic, making it easier to manage and power transformative analytics.

Organizations lack visibility into data environments, with **67% saying data fragmentation has made it hard to understand, govern, and manage data.** [4]

[4] Creating Order from Chaos: Governance in the Data Wild West

Challenge 5

There is low data literacy among the wider business, and a lack of self-service capabilities locks them out from participating.

- Most users in line-of-business teams don't have the skills required to use transformative analytics tools and techniques.
- This limits the power of transformative analytics to a small number of expensive and skilled data professionals who tend to spend their time manually finding data and hand-coding pipelines.
- Self-service tools that serve data intuitively and simply to end-users overcome data literacy challenges, but effective self-service relies on the ability to give users access to data within the parameters of robust guardrails.
- Today, many central IT teams can't give users unfettered self-service access to data and analytics tools without causing governance and security challenges.
- Organizations need low code/no code platforms to open analytics to more people and enable self-service with the guardrails that ensure data is secure and fits with governance processes.

80% of organizations want to enable a self-service data model to allow users to create and access datasets. [5]

[5] Creating Order from Chaos: Governance in the Data Wild West

How can Agile Data Infrastructure deliver analytics-ready data?

An agile data infrastructure drives order and discipline in previously chaotic data environments to power transformative analytics. It makes sense of complexity, eliminates data integration friction, and is ideal for complex data ecosystems spanning the cloud, legacy, and on-premises assets.

Agile data infrastructures work with whatever platforms teams choose for their transformative analytics projects. The flexibility of the infrastructure means it can manage varied data formats from multiple data sources (e.g., CRMs, supplier systems, etc.) and deliver data to modern data destinations such as Snowflake, Databricks, Azure Databricks, or Amazon EMR.

With an agile data infrastructure, line of business users no longer need to "get in the queue" for IT and data engineering teams to support complex requests for transformative analytics projects.

Key benefits of an agile data infrastructure for transformative analytics include:

- **Ability to visualize and describe data and data logic**
- **Total observability across data and pipelines**
- **Able to run powerful data transformations**
- **High resilience and flexibility**
- **Prevents pipeline breakages and data drift**
- **Low code/no code platforms that ensure ease of use and lower the barrier to entry**
- **Opens up analytics for all via self-service**
- **Central mission control covering the entire data infrastructure**
- **Enforces good governance**
- **Reduces reliance on hand coding, points solutions, and legacy tools**

Powering your organizations future

Transformative analytics gives organizations a glimpse of the future. To take control of their destiny, organizations must embrace new data infrastructure and adopt modern data platforms to unleash their potential. This will future-proof their transformative analytics, allowing organizations to reduce, improve and transform.

Reduce

- ✔ **Regulatory risk**
- ✔ **Compliance risk**
- ✔ **Fraud**
- ✔ **Waste**
- ✔ **Costs**

Improve

- ✔ **Enablement of strategic initiatives**
- ✔ **Forecasting**
- ✔ **Customer satisfaction and relationships**
- ✔ **Customer and market insights**
- ✔ **Risk management**
- ✔ **Financial performance**
- ✔ **Decision making**

Transform

- ✔ **Productivity**
- ✔ **Business operations**
- ✔ **Business processes**
- ✔ **Strategy and planning**

If you'd like to learn more about how **operational analytics** and effective **enterprise reporting** can drive organizational change and lay the foundations for transformative analytics, download our other eBooks now.

[Building an Agile Data Infrastructure for Enterprise Reporting]

[Five Data Principles for Ensuring Effective Operational Analytics]

StreamSets
A SOFTWARE AG COMPANY

StreamSets
A SOFTWARE AG COMPANY

About StreamSets

StreamSets, a Software AG company, eliminates data integration friction in complex hybrid and multi-cloud environments to keep pace with need-it-now business data demands. Our platform lets data teams unlock data—without ceding control—to enable a data-driven enterprise. Resilient and repeatable pipelines deliver analytics-ready data that improve real-time decision-making and reduce the costs and risks associated with data flow across an organization. That's why the largest companies in the world trust StreamSets to power millions of data pipelines for modern analytics, data science, smart applications, and hybrid integration.

streamsets.com