

The Downsides of the Modern Data Stack



Table of Contents

1. Introduction
2. What Is a Data Stack?
3. What is a Modern Data Stack?
4. What are the Components of a Modern Data Stack?
5. What Are The Characteristics of A Modern Data Stack?
6. Why Do Organizations Use A Modern Data Stack?
7. What are the Downsides to Using a Modern Data Stack?
8. How to Tackle The Challenges of A Modern Data Stack?
9. How Does Integrate.io Elevate Your Modern Data Stack?
10. How to Implement Integrate.io in Your Data Stack?

1. Introduction

Modern Data Stacks have filled the current data landscape with a vast variety of data tools and technologies.

This is evident in the growing popularity of Software-as-a-Service (SaaS) tools. A [study](#) shows that the SaaS tools market was valued at \$130.69 billion in 2021 and is expected to reach \$716.52 billion in 2028.

A Modern Data Stack is an innovative data science approach involving a strategically built stack of custom-selected tools that best fit an organization's needs. Although companies have been deriving significant value from modern data stacks, the workload of maintaining a long chain of stack tools is too heavy to ignore.

In this whitepaper, Integrate.io highlights the modern data stack, its prominent features and digs deeper into the dark sides of its features that can negatively affect an organization. Read on to explore the downsides of a modern data stack and how Integrate.io is the perfect solution to manage them.



2. What is a Data Stack?



A **Data Stack** is a set of software technologies that facilitate a smooth data flow from its inception to removal.

The tools in a data stack are used for data collection, transformation, and storage.

The number of tools in a stack varies for every organization, depending on their business requirements. There is no restriction on the number of platforms in a data stack.

The internal processes of a business define the flow of information through its data stack. The assortment of tools in a data stack is based on the business teams it is supposed to facilitate. For example, the tools in an organization's marketing stack will differ from those in its sales stack.

A data stack is designed for organizations to simplify systems integration and data analysis. The goal of a data stack is to improve business efficiency by disparately handling data management without impacting the business's computational resources and performance.

The legacy data stack proved to be an innovative step toward the increasing adoption of a data-driven culture in organizations. However, it needs help to keep up with modern organizations' rapidly evolving data demands.

3. What is a Modern Data Stack?

The Modern Data Stack (MDS) refers to a cloud-native data stack. It is a suite of software tools technology-forward organizations use for their modern data integration and analysis needs.

The tools in a modern data stack are specific to the different niches of data handling, corresponding to the various stages of the data lifecycle, including data management, transformation, analysis, and more. These tools patch together to build an organization's data ecosystem.

Organizations today require actionable data at the speed of thought and the flexibility to adapt to changes. In an attempt to modernize the legacy data stack, the modern data stack was introduced. The term 'modern' in a modern data stack points to its key feature of having all cloud-based tools as part of the stack. The reason behind an all-cloud stack is the capability of the cloud to manage large volumes of data that can scale your business without draining your precious resources in the process.

Engineers locate and identify relevant data from a data source during the Extract phase and then transform that data into a suitable format for data analysis.

The aim behind a modern data stack is to help organizations segment their tasks on top of a stack of interoperable cloud tools. This system is designed to streamline an organization's data architecture for a smooth flow of information and faster data activation. A modern data stack leverages the cloud's speed to enhance business throughput so data can match the speed of business.

The ultimate objective of a modern data stack is to give businesses real-time data access to proactively discover new opportunities and boost business efficiency.



4. What are the Components of a Modern Data Stack?

Before getting into the features of a modern data stack, it is essential to understand the independent components of the data stack and the tools that serve under each element. The primary components of a modern data stack include:

1. Data Collection

Data collection involves gathering raw, unstructured data from various applications, devices, and backend services. It collects two main types of organizational data: behavioral data and transactional data.

The modern data stack platforms involved in data collection focus on tracking and compiling data to ensure reduced quality issues in the accumulation of data.

2. Data Storage

Data storage is the process of storing information for all data platforms to access and utilize. It involves a data warehouse or a data lake that acts as a central repository for storing all the data in a modern data stack.

The modern data stack tools for data storage focus on supporting various data formats, real-time/batch data ingestion, auto-scaling, and good data governance.

3. Data Ingestion

Data ingestion is the process of loading raw, unstructured data from a source platform to a central data lake or warehouse.

As a part of a modern data stack, a data ingestion component consists of [data integration tools](#) responsible for laying quality data pipelines. In addition to data integration, pipeline monitoring is also essential to data ingestion.

4. Data Transformation

[Data transformation](#) is the mechanism of [cleansing](#), filtering, and processing raw data to retrieve structured, understandable data that can be queried for analytics.

In a modern data stack, data transformation tools are designed to support consistent data modeling frameworks that allow businesses to prepare their data for analysis.

5. BI & Analytics

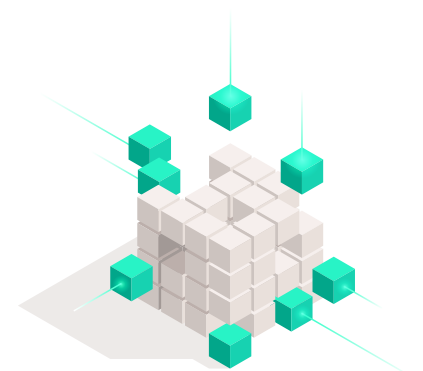
Business Intelligence & [analytics](#) are essential components of a modern data stack that allows businesses to understand their data and make informed decisions.

BI tools in the modern data stack involve data visualization and analytics tools. The reports, dashboards, and metrics from these tools make data easily understandable for all organizational teams to extract rich information.

6. Data Orchestration

Data orchestration corresponds to the consolidation and organizing of data to ensure data availability for analysis. Orchestration involves pipeline management tasks, including maintaining pipeline schedules, responding to failures, and managing all pipeline dependencies from one place.

The data orchestration tools in a modern data stack consist of workflow management tools and end-to-end pipelining solutions that support pipeline management and smoothly integrate with infrastructure services.



5. What Are The Characteristics of a Modern Data Stack?

A modern data stack is recognized by a combination of state-of-the-art features that deliver modern solutions to problems of present-day businesses. Following are some of the critical characteristics of the platforms that make up a modern data stack:

Cloud-native

Research predicts the [global cloud computing market](#) to reach \$1.6 trillion by 2030. This rising adoption of cloud technologies has been the driving force behind developing a modern data stack. The tools in a modern data stack are backed by cloud vendors and are highly scalable. Cloud-based technologies are also easier to integrate with an organization's existing cloud infrastructure.

Centralized Data Storage

A modern data stack is known for offering a single source of true data in the form of cloud data warehouses/lakes. A data warehouse/lake acts as a centralized storage location for a modern data stack that seamlessly integrates with all the stack tools and fuels data analytics. The most popular cloud data warehouses in modern data stacks include BigQuery, Redshift, Databricks, and Snowflake.

Plug and Play Architecture

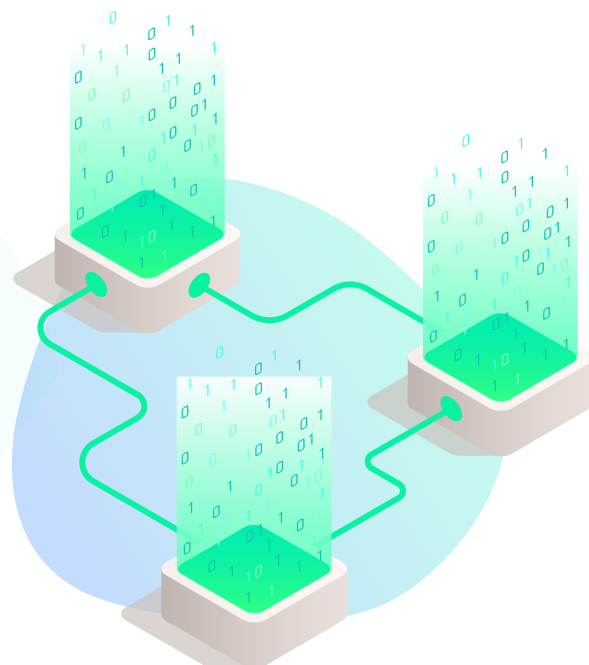
Another prominent characteristic of a modern data stack is its plug-and-play architecture. This feature enables tools to fit flexibly into various data models and seamlessly plug into an existing data stack. A plug-and-play architecture creates a network of tools that plug together without requiring changes in the stack. With a variety of domain-specific tools, a modern data stack is a chain of tools plugged together to implement different stages of data operations.

SaaS Tools

The tools in a modern data stack fall into Software-as-a-Service (SaaS) platforms. They come as pre-configured tools ready to fit into a modern data stack. SaaS tools are known for their self-service capabilities, and their vendors manage their upkeep and maintenance.

Low Code/ No Code

Modern data stack tools are purpose-built to simplify data operations for non-technical organizations. The low-code and no-code designs of modern data stack tools fulfill this purpose. These tools can be set up and operational in a few hours without technical expertise.



6. Why Do Organizations Use a Modern Data Stack?

The booming digital world requires modern organizations to equip themselves with technologies that empower them to embrace the changing data demands.

However, even in today's digital pool of endless opportunities and possibilities, organizations need help finding their competitive edge in the market. When the modern data stack was conceptualized, it gave organizations hope to create a business advantage for themselves. From saving costs and time to enhancing business performance and scalability, implementing a modern data stack benefits businesses in numerous ways. Following are some of the primary reasons why organizations are attracted to using a modern data stack for their data operations:

Cost Effective

A modern data stack comprises mainly SaaS tools that come with consumption-based pricing and maintenance support. Organizations pay for what they use, and the costs only rise as their consumption scales. Moreover, these tools eliminate the need for businesses to hire dedicated developers to build and maintain their data pipelines, enabling them to lower their data engineering expenses. With the platform vendors handling their tool's maintenance, organizations can spend their surplus capital on critical business tasks. Even as organizations scale, the modern data stack automatically grows with them.

Scalability

The cloud-native capability of a modern data stack empowers organizations to reach almost infinite scalability. The time, costs, and computational power associated with growing data needs are no longer a burden on an organization's infrastructure. With the cloud transparently and quickly executing complex computations, organizations can focus on their business growth without worrying about the scalability of their modern data stack.

Interoperability

A modern data stack's plug-and-play feature is handy for organizations to have a collaborative system of tools. The tools in this stack are designed to integrate with a variety of cloud services and custom applications that patch together as a single data system. These tools exchange data and communicate using an Application Programming Interface (API) built for smooth interoperability.

Low Entry Barrier

Another major reason organizations are moving towards a modern data stack is its low technical barrier. A low entry barrier is beneficial for implementing and using the modern data stack, with inclusivity to all organizations and not just the data experts. Non-technical professionals can easily build and manage data stacks with the low-code/no-code features of the modern data stack. Moreover, a modern data stack offers better data comprehension than traditional data stacks, allowing businesses to quickly understand data and derive meaningful insights. The technical barrier is eventually lower than the traditional data stack.

Business-Focused

With the complex IT tasks handled mostly by platform vendors, a modern data stack enables organizations to take a business-focused approach to reach their data goals. The tools in a modern data stack are designed for self-service, so organizations do not have to depend on data professionals. By removing the technical bottleneck, a modern data stack enables data teams to focus on higher-value data tasks while business teams extract and analyze information to improve business efficiency.

Operational Analytics

Organizations today highly value technologies that can help them get faster and better results from their data to get ahead of their rival businesses. It is one of the reasons why organizations are more compelled towards a modern data stack. Modern data stacks have evolved traditional analytics to iteratively deliver operational analytics to businesses for quick and efficient data outcomes. With a modern data stack, even non-technical companies can gain actionable insights from their data within a matter of hours.

7. What are the Downsides to Using a Modern Data Stack?



Data operationalization is faster and simpler than ever before, thanks to the evolution of the modern data stack.

It has enabled businesses of all sizes to become data-driven. However, the benefits of a modern data stack come at a price. Implementing a modern data stack comes with various obstacles that can hinder the competitive success of a business.

The concept of a modern data stack sounds like a flawless improvement in an organization's data experience. However, the problem lies in the ways many organizations implement their modern data stack. Although the modern data stack resides in the cloud, every tool in the stack takes up a part of the organization's infrastructure for its configuration and management. Whether it is the burden of tool maintenance or the frustrating end-user experience of jumping between tools, a lot falls flat with using a modern data stack.

Following are some of the downsides to using a modern data stack for your business:

Weak architectural design

Organizations that pursue a modern data stack without backing it with a strong data architecture are setting themselves up to fail. A weak architectural design and a lack of iterative data modeling processes result in the implementation of an unstable data stack that will never scale. The long-term effects of this failure show up as technology debts, and the cloud services end up costing businesses an arm and a leg.

Data democratization issues & Data Silos

Each tool in a modern data stack acts as a miniature data ecosystem that resides in a more extensive system with other tools. It means that bulks of data and metadata are associated with a single tool, creating an accessibility issue. When all organizational stakeholders do not have easy access to all information they need, the data gets siloed in the different tools. Data silos create technical barriers that can inhibit data democratization in an organization. Consequently, a modern data stack acts as an obstacle between business teams and organizational data, eventually threatening a business's long-term success.

Compromised Data Quality

The quality of an organization's data is controlled by so many factors that it can be challenging to track and maintain. With a long chain of platforms to pass through, data's journey through an organization's modern data stack can negatively impact its quality. As a result, businesses end up with compromised, incomplete data, which can be a massive setback for operational analytics. This creates grave challenges for data analysts who must spend additional resources to transform the data into actionable information.

Data Governance Issues

A modern data stack is designed to grow and accommodate changing data requirements of a business. Ultimately as more and more data accumulates and more applications and teams start getting involved in the data operations, it becomes critical to know when, how, and who is working with data. The more a modern data stack expands to cultivate productivity, the more data governance issues will arise for organizations. Consequently, it becomes highly likely for organizations to lose control of their data and its operations.

7. What are the Downsides to Using a Modern Data Stack?



Slow Decision-Making

Incorporating data changes and error handling in a long line of tools is a tedious task. The fragmented data is not easy to observe holistically. In addition to the resources involved, tracing back to data origin issues requires switching between multiple tools and their specific standards. A modern data stack's engineering and compliance overhead leads to delays in critical business decisions.

Data Security

The [security controls and permission settings](#) vary from tool to tool. When more and more tools are added to a modern data stack, the security circle of that organization expands. Managing data security in this vast circle of software platforms is challenging, with crucial security risks involved. The bigger the circle, the more security gaps exist with larger areas of potential leakages in an organization's data ecosystem. In this way, a modern data stack leaves an organization's security at the mercy of the tools in its data stack, posing major security threats to critical business information.

Hidden Expenses

When adding new tools to their modern data stack, most organizations are fooled by the visibly low costs of individual tools. They must know the hidden expenses that a fragmented data stack withholds. These expenses are of two kinds; opportunity costs and risk costs.

The opportunity costs come from spending business time and resources on managing tools. Every hour spent on management slices away the time a business can spend building data products and driving growth.

The risk costs arise from specialized stack tools with a high technical barrier. It puts business stakeholders at risk of using tools with no available learning resources. Additionally, the budget overheads associated with error handling and change management can quickly spiral to extravagant rates that organizations may no longer be able to afford.

8. How to Tackle The Challenges of a Modern Data Stack?

There are many ways to get the most value out of your modern data stack without overcrowding it with disparate tools.

Following are some of the practices that can help organizations prevent or even solve a lot of the problems associated with a modern data stack:

Invest in building a solid data architecture

Investing in data modeling and architecture is crucial for organizations to find the right tools and glue these disparate tools together as a stack. Focusing on data architecture allows businesses to leverage their modern data stack for a modern data experience. Whether companies choose a centralized data lake architecture or a full-blown data mesh of domain-specific microservices, an organization should keep its MDS from exploding with tools.

Prioritize good data governance policies

An elevated data experience is about more than just software and its infrastructure. Businesses must improve their organizational processes as well. As an organization, **governance policies and procedures** should be of top priority for achieving optimal compliance. Good governance standards give businesses more control over their data so they can scale and maintain their operations smoothly.

End-to-End stack solutions

An end-to-end stack is a Datastack-as-a-Service (DaaS) solution that consolidates data operations like governance, infrastructure, configuration, and security. On top of being stand-alone data stacks, these solutions also offer the flexibility to integrate with other tools. Using end-to-end stack solutions reduces the requirement for separate tools for every data operation. Organizations can save significant time and resources when they don't have to purchase and maintain numerous tools.

Focus on Data quality

Maintaining the quality of data that travels through a crowd of stack tools is difficult. Instead of investing in multiple platforms for speed and productivity, organizations should focus on value-driven technologies that not only reduce data silos but also make sure that businesses receive quality data to derive meaningful actions.



9. How Does Integrate.io Elevate Your Modern Data Stack?

Integrate.io is a data integration platform for E-commerce retailers that supports automation of the various implementations of data operations, including

- **Extract, Load, Transform (ELT)**
- **Low-code Extract, Transform, Load (ETL)**
- **Change Data Capture (CDC),**
- **Reverse Extract, Transform, Load (ETL)**
- **Representational State Transfer (REST) API**

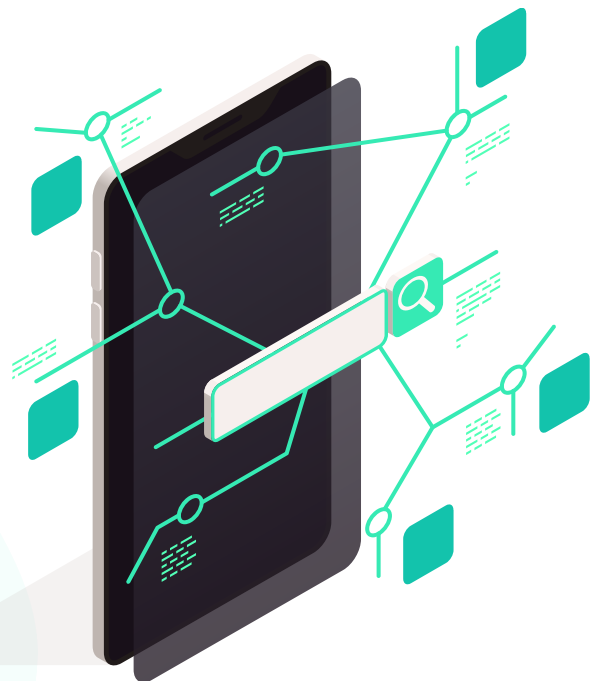
It is an all-in-one data operationalization solution based on your specific circumstances and business needs. It is designed to be readily usable by IT and business users alike. Integrate's end-to-end stack solutions make it an excellent fit for your modern data stack.

Integrate comes with built-in connectors for a wide variety of data sources, including a REST API component for specialized tools for your custom needs. It gives you complete control over your data integration processes and the flexibility to freely set up transformations using its Database Transformation Toolkit (DBT).

Low-code processes, an easy-to-use interface, and a cost-effective approach to handling data operations are just a few of the many features of Integrate.

On top of all this, you can ensure maximum security for your sensitive data in your stack with Integrate.io's data masking, hashing, and encryption/decryption standards.

Modern technology-forward companies can significantly benefit from utilizing Integrate.io in their modern data stack to promote a growth-oriented business culture. From data integration projects using ELT and master data management jobs to big data warehousing jobs and more, Integrate.io is a one-stop solution capable of efficiently managing all your data stack needs.



10.How to Implement Integrate.io in Your Data Stack?



Integrate.io allows you to set up a sophisticated data stack, even if you have limited in-house expertise. But how do you get started with Integrate.io?

Before implementing a modern data stack, it can be beneficial to reevaluate your approach to data. Following are some of the steps you can take to [reinvent your modern data stack](#):

- Keep all your data stakeholders in mind before building a data infrastructure. Think of every department that consumes organizational data.
- Maintain a catalog of all the existing data sources within your organization.
- Prioritize your data based on its operational value to ensure high-quality data flow in your pipelines.
- Stay one step ahead of compliance and security risks with good data governance policies in place to ensure consistent data flow in your stack.

Once you have realigned your strategies, the next step is to implement Integrate.io and let it take care of the rest of your modern data stack needs.

If you are ready to elevate your modern data experience, [contact our support team](#) at Integrate.io. You can schedule a demo or get a risk-free, 14-day pilot and experience the Integrate.io platform.

**We hope that you have enjoyed
reading this white paper.**

If so, be sure to check out our other [White papers](#) related to
"The Downsides of The Modern Data Stack".

**Learn more about how Integrate.io can help
make your data work for you!**

[GET A DEMO!](#)



Integrate.io