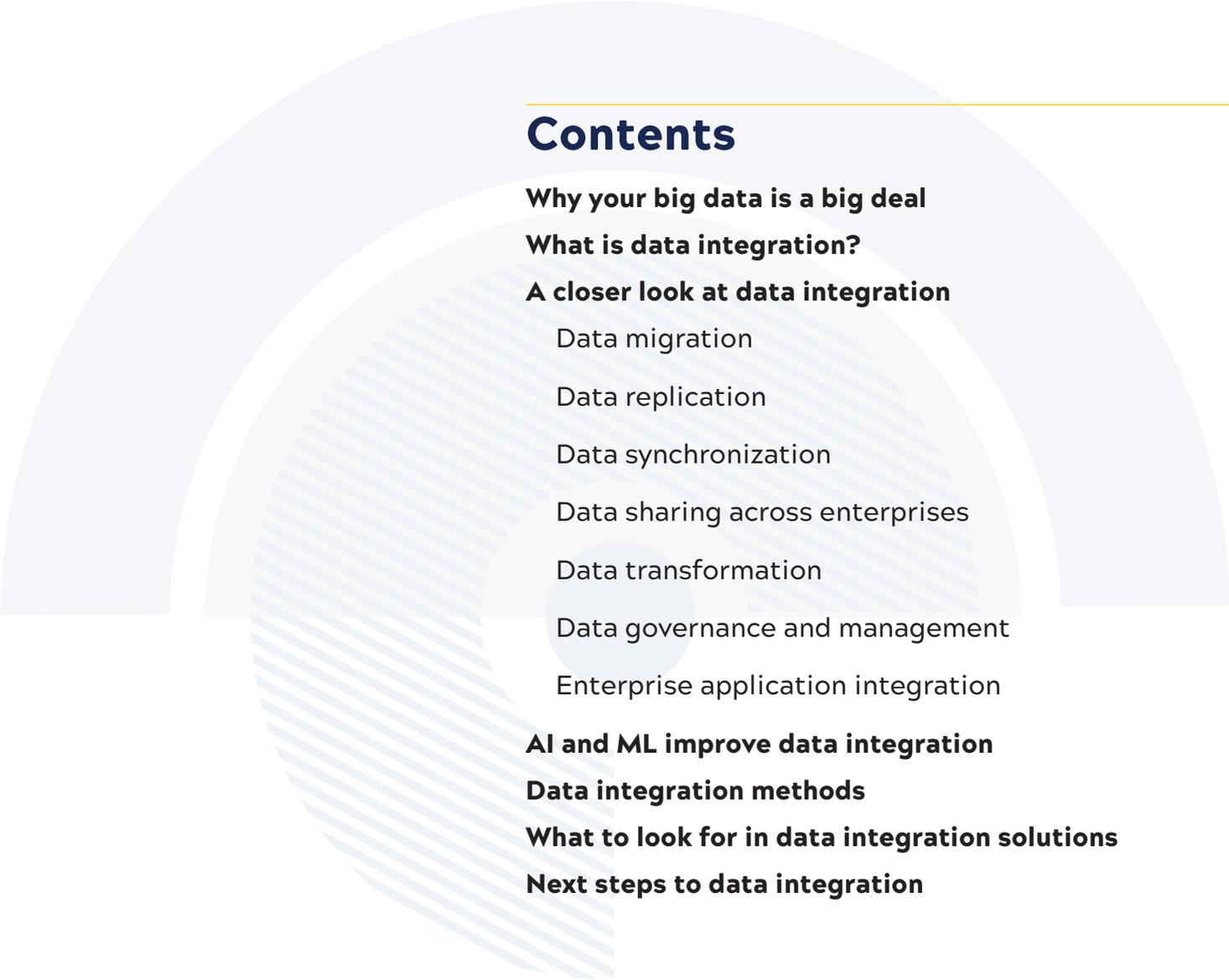


snapLogic®

The Ultimate Guide to **Data Integration**



Contents

Why your big data is a big deal	3
What is data integration?	4
A closer look at data integration	7
Data migration	8
Data replication	8
Data synchronization	9
Data sharing across enterprises	9
Data transformation	10
Data governance and management	10
Enterprise application integration	11
AI and ML improve data integration	12
Data integration methods	12
What to look for in data integration solutions	13
Next steps to data integration	14

“ [With SnapLogic] we’re scaling to enormous data volumes, analyzing more than 100 million records in a recent integration. Instead of taking five to six hours to push 150,000 records to an external endpoint, we were able to push three times that volume of information to the external endpoint in about two minutes.”

- Senior Manager, Data Strategy and Architecture, Box

Why your big data is a big deal

Data is information that is generated by your business, employees, customers, partners, and technology. Data presents insights about what has happened in the past, what is happening now in real-time, and what may happen in the future. When enterprise data sources are brought together to create a complete, holistic perspective, it gives you the insights you need to make better business decisions and take faster actions.

Today’s businesses compete based on how fast and how well they can glean meaningful insights from their data assets to generate better products, services, and, ultimately, experiences that customers seek out to determine whether they will stay loyal to your brand or buy from your competitor.

The faster you can access the insights from your data, the faster you can take action and make bold moves in your market. Speed and responsiveness anchored on accurate, trustworthy data is the goal.

The problem is that as technology, data repositories, and applications have exponentially multiplied and continue to do so - the average enterprise has 115 applications in use at any given time - and they generate endless amounts of data. A typical business generates vast amounts of data from daily operations such as activities on e-commerce websites, transactions at point-of-sale (POS) systems at retail stores, sensor data from machines at manufacturing plants, inventory backlogs at warehouses, and many more. The more data you have, the harder it gets to synchronize data across data stores, replicate data, ingest data, maintain a single source of truth, ensure data quality, and find the meaningful data points that can impact your success.

This is where data integration comes in.

Yes, it’s a complex topic, but like everything we do at SnapLogic, we aim to make it easy to understand and work with.

Our goal? To help you discover how data integration is foundational to your success today and crucial to building the automated enterprise of the future. That future, by the way, starts today.

What is data integration?

Data integration is the process of bringing together various sources of data to make that data more useful. For example, by integrating data from your CRM, contact center, website, mobile app, marketing, and sales software, you can create a 360-degree view of your customer data. Having a 360-degree view is more useful to the business (and ultimately to the customer experience) than just knowing separate data points that don't seem to connect in each of the systems.

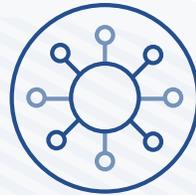
Data integration involves techniques, tools, and practices that ingest, transform, combine, and provision data across all data types, wherever that data may live to meet the data needs of your applications and business processes. Data integration provides a way to access data, transform the data, and then deliver it to a destination data repository or application.

So, at its core, data integration is about making data more useful.

To do that, it needs to deal with all types of data, combine and transform data efficiently. Ultimately, effective data integration helps your business leverage accurate, holistic, up-to-date data to make better strategic decisions.



Benefits of data integration



A single, trustworthy source of truth across the business



Improved analysis, forecasting, and predictive analytics



360 view of the customer and business



Faster innovation and time to market



Increased agility and responsiveness



Improved ROI from tech stack investments

Data integration is foundational to driving deeper collaboration and to automating business processes and workflows that improve efficiency, decrease human effort, and create the type of enterprise that delivers exceptional experiences. Wherever you are in your digital transformation journey, data integration is the key to moving forward.

What happens without data integration

Nearly every enterprise today has undergone efforts to digitally transform and harness the power of their data. But, it's not uncommon to still be struggling.

With massive amounts of data being generated, the rapid pace of tech innovation, the costs of change, growing sprawl of application and data silos, and a plethora of available data management and analytics tools to choose from - it's easy to see why so many businesses wrestle with trying to effectively manage and glean real value from their data.

Data that is not integrated remains siloed in the places it resides. It takes a lot of time and effort to write code and manually gather and integrate data from each system or application, copy the data, reformat it, cleanse it, and then ultimately analyze it. Because it takes so long to do this, and skilled resources in IT who can do it are scarce, the data itself may easily be outdated and rendered useless by the time the analysis is complete. Businesses don't have time to wait anymore.



A closer look at data integration

Now that you know the answer to “What is data integration?”, let’s take a look at it in a bit more detail.

There are three broad categories of data integration:



Operational data integration:

the goal is to synchronize and replicate operational data and make it available across applications, systems, and databases



Analytical data integration:

here you want to integrate data for all your systems of record, transform and summarize data and push it to your analytics stack which typically consists of a data warehouse and an analytics or business intelligence (BI) endpoint



Hybrid data integration:

this is a combination of operational data integration and analytical data integration where insights from BI tools are fed into operational systems to improve customer experience and drive operational efficiency

Businesses can have data spread across on-premises and cloud systems and need all three types of data integration to optimize performance and enable automation across this hybrid environment.

Data integration also has several subcategories. Below are the ones you'll hear most about:



Data migration

Data replication involves replicating data from where it is generated - such as a POS system or a warehouse inventory record from a particular region - to where it needs to be analyzed for planning, forecasting, and insights. There are different types of data replication, such as:

- Bring it closer to other data assets that are similar so that they can be combined to get useful insights
- Reduce the cost of data storage
- Improve data access performance
- Improve data availability



Data replication

Data replication involves replicating data from where it is generated - such as a POS system or a warehouse inventory record from a particular region - to where it needs to be analyzed for planning, forecasting, and insights. There are different types of data replication, such as:

- A full table replication copies data from source table to the destination in its entirety. It is time-consuming and requires significant network bandwidth
- Incremental replication, which can be key-based or log-based, identifies changes in the source data and propagates them to the destination

Effective data replication improves:

- Reliability and availability of data
- Performance for analytical workloads and helps generate holistic insights



Data synchronization

Data synchronization is the process of synchronizing similar objects or data structures (schemas) across different data stores or applications. There are two ways of looking at data synchronization:

- Incremental data replication can be viewed as data synchronization from a source to a destination
- Data synchronization can also be between two different applications - for example, a CRM system (such as Salesforce, Microsoft Dynamics CRM) and a Service Management system (such as ServiceNow, Zendesk) both hold important customer records. But the data in the CRM system will often be viewed as the master record. In that case, customer details in a Service Management system need to be synchronized periodically with the CRM system

Data synchronization is crucial to make sure that all departments in an organization, who may rely on different systems of records, are working with the most up-to-date data.



Data sharing across enterprises

Organizations deal with many external entities such as suppliers, distributors, customers, and partners as part of normal business operations. Data sharing across enterprises includes systems such as Electronic Data Interchange (EDI) that enable business partners to agree on data formats, acquire data, exchange messages, collaborate, and execute end-to-end business processes such as catalog discovery, procure-to-pay, order-to-cash, transportation of goods, and more.

Effective data sharing across enterprises provides the following benefits:

- Reduces time-to-market
- Reduces manual errors and improves productivity
- Improves revenue and earnings by selling goods and services through more channels and partners



Data transformation

Data integration tools are often known as Extract, Transform, and Load (ETL) tools, and the key functionality they provide is the ability to transform data. Data transformation includes but is not limited to changing data formats, combining data across multiple data sources, filtering or excluding certain data entries from the combined data set, summarizing values across data sets, and so on.

Data transformation can be done using code, SQL scripts, or visually. Data transformations are so fundamental to any data integration flow that the ability to transform data with ease is a crucial differentiator for any data integration tool.



Data governance and management

Data governance and management is a broad capability that consists of the ability to audit, control access to, profile, govern, share, and monitor data. It encompasses areas such as:

- Data Catalogs allow organizations to create an orderly list of data assets in an organization. Data catalogs use metadata associated with data assets to uncover context with various repositories of data. This metadata is then used for data discovery and to uncover data relationships
- Data Virtualization allows users and applications to access and manipulate data without any knowledge of how the data is structured, or where it is located

Data governance tools enable organizations to:

- Build trust in data
- Maintain data privacy by controlling access
- Provide audit capability so that organizations can proactively comply with regulations
- Allow collaboration between users so that collectively they can make the most of the data



Enterprise application integration

What is enterprise application integration (EAI)? Just like the name suggests, EAI is all about creating interoperability among applications and systems. This is where you get Salesforce, Workday, Microsoft Dynamics, ServiceNow, and NetSuite to play nice with each other. Traditionally, organizations managed their EAI tools distinct from their data integration tools. But now organizations are increasingly using a single unified platform for both application and data integration. EAI is critical to creating omnichannel customer experiences, streamlining workflows and processes, and creating seamless experiences for customers, partners, and employees.

These are the most common aspects of data integration. Next, we turn to the role artificial intelligence and machine learning play in data integration.



AI and ML improve data integration

Artificial intelligence (AI) and machine learning (ML) capabilities are increasingly being built into data integration platforms, significantly improving integrator productivity and time to value. AI-augmented data integration platforms speed up the ability to identify, access, connect, and move data.

Data integration solutions that incorporate AI are able to more readily find useful pipeline patterns, more relevant data from a given source, and drive faster, more accurate analysis and insights. Part of data integration is dealing with sensitive or personally identifiable information, identifying what should be masked or anonymized, and also discerning what is useful and what isn't. AI is able to do this automatically to help ensure compliance with HIPAA, GDPR, and other regulations.

Data integration methods

So, how do you do data integration? There are several approaches ranging from manual integration to low-code data integration platforms:

- Code it manually. This is a time-consuming and resource-intensive method where integrations are manually coded from a source to a destination and must be monitored and continually maintained by IT
- Use middleware. Middleware data integration serves as a mediator between data that needs to be normalized
- Let an integration platform as a service (iPaaS) simplify it for you. An iPaaS, such as the SnapLogic Intelligent Integration Platform, provides out-of-the-box connectivity to thousands of data and application endpoints, simplifies data transformations, and makes it easy to manage and govern that data

What to look for in data integration solutions

Choosing among data integration tools is about selecting a method that will make the process smarter, faster, and easier - and work within your budget.

It's also about harnessing the power of AI to ensure your integration capabilities can keep your business moving with as much speed and agility as possible.

Here are some things to look for in a data integration platform:

- Is it purpose-built for the cloud? No legacy components? Is it self-upgrading, with an elastic execution grid? Can it scale up or out? Manage environments from public to private and on-premises?
- Does it offer a clicks not code approach for faster, easier integration? Drag-and-drop, and snap-and-assemble? Is it robust enough for developers, but easy enough for your business teams to use?
- Is it AI- and ML-enabled to bring speed, quality, and accurate predictability to data-driven decision-making?
- Is the pricing transparent and predictable? As your team builds more integrations, moves more data, will you have to pay more?
- Does it enable integrations beyond data integration to provide an easy, complete way to integrate both data and applications?





Next steps to data integration

The SnapLogic Intelligent Integration Platform is here to make your data integration process smarter, faster, and easier. With an intuitive interface and more than 500 pre-configured Snaps, your business teams and developers can integrate data silos and applications with clicks, not code. Our Iris AI provides proven recommendations and guidance for smarter integration.

Learn more about [SnapLogic's data integration capabilities](#), and start your [free trial today](#) or [contact us to get a custom demo](#).

SnapLogic provides the #1 intelligent integration platform. The company's AI-powered workflows and self-service integration capabilities make it fast and easy for organizations to manage all their application integration, data integration, and data engineering projects on a single, scalable platform. Hundreds of Global 2000 customers – including Adobe, AstraZeneca, Box, Emirates, GameStop, and Wendy's – rely on SnapLogic to automate business processes, accelerate analytics, and drive digital transformation. Learn more at snaplogic.com.