# Data Integration Design Patterns

In the age of digital marvels, data is more important than ever before. Businesses rely on data to make critical decisions, and the ability to quickly and effectively integrate data from various sources is essential to success. An infrastructure that transfers data between the systems is necessary for connecting multiple systems. However, you frequently want the solution to be able to do more than merely distribute data. And this is where data integration comes in.

**Overview of Data Integration**

Modern businesses often face the issue of data overabundance. They have more data than ever before, but it’s spread across various silos and systems. [Data integration is the process](https://www.divami.com/blog/guide-to-data-integration-design) of combing through this data and assembling it into a format that can help make better business decisions.

Integration involves procedures like cleansing, [ETL mapping](https://www.astera.com/type/blog/understanding-data-mapping-and-its-techniques/), and transformation and starts with the ingestion process. It also paves the way for analytics to create valuable, actionable business knowledge.

A network of data sources, a controller server, and clients that access data from the primary server are the standard components of data integration systems. A typical data integration process involves the client requesting data from the controller server. The primary server subsequently ingests the required data from both internal and external sources. The information is gathered from many sources and then assembled into a single, comprehensive data collection. Finally, the client receives this and uses it.

**What is Data Integration Pattern?**

A standardized approach to integrating data is the Data Integration Pattern. DIP aids in standardizing the entire data integration process.

Data Integration patterns can be divided into five categories:

– Data Migration Pattern

– The Broadcast Pattern

– Bi-directional Pattern

– Correlation Pattern

– Aggregation Pattern

**Why is Data Integration Pattern required?**

The days of simply integrating data using conventional techniques are long gone. Instead, we must adapt our tools as we enter the new era of data to keep up with rapid technological advancements.

This is where [data integration patterns](https://www.divami.com/blog/data-integration-challenges-faced-by-manufacturers/) come in. They provide us with a standardized approach that we can use to quickly and effectively integrate data, regardless of the source.

The following are some benefits of developing data integration patterns for various forms of data provided by data sources:

**1. Time-Saving**

As we create an integration pattern for specific circumstances, Data integration patterns allow us to save significant time and effort.

**2. Better Business Decisions**

Using data integration patterns may be beneficial for business growth as it allows for a unified view of all the data in one location. DIP can also be used to synchronize data between different departments. This aids in the collaboration process and helps decision-makers to have a clear understanding of what is happening in other departments.

**3. Adaptability**

Data integration patterns ensure that the systems can adapt to emerging technology. They also help us to take advantage of new opportunities as they arise.

**4. Reusability**

DIP offers a method of integrating data that can be used in applications with a single click. This avoids having to develop new plans for each application.

**5. Reliability**

DIP ensures the dependability of data integration. With well-tested patterns, the chances of errors are significantly reduced.

**6. Better Communication**

Since DIP solves the issue of data silos, communication between multiple departments can be improved.

**What are the types of Data Integration Patterns?**

Similar to a hiking trail, patterns are found and established over time. They are always imperfect to varying degrees, but they can be improved or adapted to meet specific business needs.

The business use case, which is used for the general data transportation and handling process, can be thought of as an instantiation of the pattern. There are five data integration patterns based on the business use cases, including cloud integration patterns.

**1. Data Migration Pattern**

A specific data set is permanently transferred from one system to another using the data integration pattern known as data migration. Data is already contained within a source system before data migration happens.

Key steps of the data migration process include:

**– Data selection**: The data that is required for the target system is identified.

**– Data cleansing and preparation**: The selected data is cleansed to ensure accuracy and completeness.

**– Preparation**: A mapping is created between the source and target data structures.

**– Extraction**: The data is extracted from the source system.

**– Transformation**: The data is transformed to fit the target system’s requirements.

**– Loading**: The data load is inserted into the target system.

**– Data transfer**: The data is transferred from the source system to the target system.

**Why does it add value?**

The need for data migration often arises because businesses frequently use multiple systems for data management. When business needs or processes change, the data might need to be migrated to another system more suited to the new requirement. Data migration is essential to preserve our data independent of the tools we use to produce, view, and manage it.

Without migration, every time we wanted to switch tools, we would have to start over from scratch, hindering cost-efficiency, performance, and scale in the digital age.

**When is it useful?**

– Moving from one system to another

– Switching to a newer instance of that system

– Launching a new strategy to expand your current infrastructure

– Backing up a dataset

– Adding nodes to database clusters

– Upgrading database hardware

– Consolidating systems

**2. The Broadcast Pattern**

The broadcast data integration pattern distributes data in real-time from one source system to numerous destination systems. This is a recurring theme in the integration of broadcast data. Only data integration methods such as broadcast, correlation, or bidirectional sync allow for real-time data access between many systems.

The broadcast data integration pattern differs from other patterns in that it only transfers data in one direction – from the source to the target. Therefore, the pattern for broadcast data integration is transactional.

**Why does it add value?**

The broadcast pattern is quite helpful when system B needs to know certain information from system A that originates or lives in system A in close to real-time. For instance, you might wish to develop a real-time reporting dashboard, which would serve as a receiver for several broadcast-type integrations.

**When is it useful?**

– Propagate changes made in the source system to the destination systems in real-time

– Synchronize data between multiple destination systems

– Build a reporting dashboard that uses data from several source systems

– Transactions that need real-time propagation to various receivers

– Events that need real time processing by numerous receivers

**4. Bi-directional Pattern**

Businesses no longer have to manually deal with various data irregularities, thanks to bidirectional sync data patterns. With high quality and real-time data accessibility, organizations can gain a competitive advantage, drive better decision-making, and improve operational efficiency.

Bidirectional sync is a data integration pattern that involves combining two data sets from two different systems using the bi-directional sync data integration pattern. Consequently, two independent data sets can coexist separately and simultaneously function as one data set.

Organizations with numerous systems and business processes running concurrently benefit from bidirectional sync data patterns.

**Why does it add value?**

Depending on the conditions that call for it, bi-directional sync can serve as both a facilitator and a rescuer. For example, bi-directional sync can be used to streamline your procedures if you have two or more separate, independent representations of the same reality.

Conversely, you can switch from a group of items that function well together using bi-directional sync. Still, you may not be the best at performing each of their individual functions in a group that you hand-pick and integrate using an enterprise integration platform.

**When is it useful?**

– Achieving a single view of your customers by granting everyone access to all the systems.

– When listing the data that must be made public for that customer object and system and identifying which systems are the owners is a more practical approach.

For instance, a salesperson should be aware of the delivery status but need not know whose warehouse it is. Similarly, the delivery person only has to see the name of the consumer receiving the delivery, not their purchase price. Through bi-directional synchronization, each of them can view the same consumer in real-time using the appropriate filter.

**4. Correlation Pattern**

Bi-directional synchronization is incorporated into the correlation data integration structure. The correlation data integration pattern first detects the points of intersection between two data sets. The item occurring in both systems is then bi-directionally synchronized.

It’s important to note that natural item existence in both systems is a prerequisite for bidirectional synchronization. However, since bidirectional synchronization is only used for the pertinent intersecting data, correlation eliminates the requirement for extraneous data storage.

**Why does it add value?**

When two groups or systems want to share data only if they both have a record that accurately represents the same thing or person, in reality, the correlation data integration pattern might be helpful. A hospital organization, for instance, operates two hospitals in the same city. You might want to share information between the two hospitals so that, if a patient visits either facility, you have a current record of the care they received there.

You can choose to establish two broadcast pattern integrations, one from Hospital A to Hospital B and one from Hospital B to Hospital A, to complete an integration of this kind. The data will be synchronized, but you will now need to manage two integration applications.

The correlation pattern is helpful because it only moves the items in both directions when it is necessary to do so, rather than constantly moving the entire dataset in both directions.

**When is it useful?**

– When having excess data is more expensive than helpful, and you only want to move data that is needed.

– When you need to share data between two groups or systems, but only if they both have a record that accurately represents the same thing or person.

– Weeding out the unnecessary data sets from each of the systems.

– Only sharing what is essential to save you time, money, and aggravation.

**5. Aggregation Pattern**

Data from several systems are received or taken and inserted into one outline using the aggregation data integration pattern. The aggregation data integration pattern preserves data integrity and offers a format-related integration solution.

Real-time accessibility is supported by the capability of processing data derived from several systems in a single application. Additionally, data replication is prevented, which is crucial for businesses with small data warehouses.

Aggregation integration data patterns are constructive for application programming interfaces that employ data from numerous systems for a single answer. In addition, enterprise data that is compliance-related is another exciting application.

**Why does it add value?**

You can harvest and process data from various systems in a single application using the aggregation pattern. This means that the data is up to date when needed, does not get replicated, and can be processed/merged to produce the dataset you want.

**When is it useful?**

– Beneficial when developing orchestration APIs to “modernize” legacy systems

– Creating an API that gathers data from several designs and then processes it into a single response.

– Creation of reports or dashboards that combine data from many methods to produce an experience.

– Combining pertinent data from various systems while ensuring that your compliance data is stored in a single system.

**Conclusion**

Data integration plays a more significant role in consolidating an organization’s existing data to give analytics apps and users a consistent snapshot of the operations. This is important as firms frequently wind up with several information systems and databases over time.

Data integration patterns provide distinct approaches to data management that can be used in various circumstances. By understanding the use cases and benefits of each pattern, you can use the right one for your needs, whether sharing data between two groups or taking data from several systems and inserting it into one.