



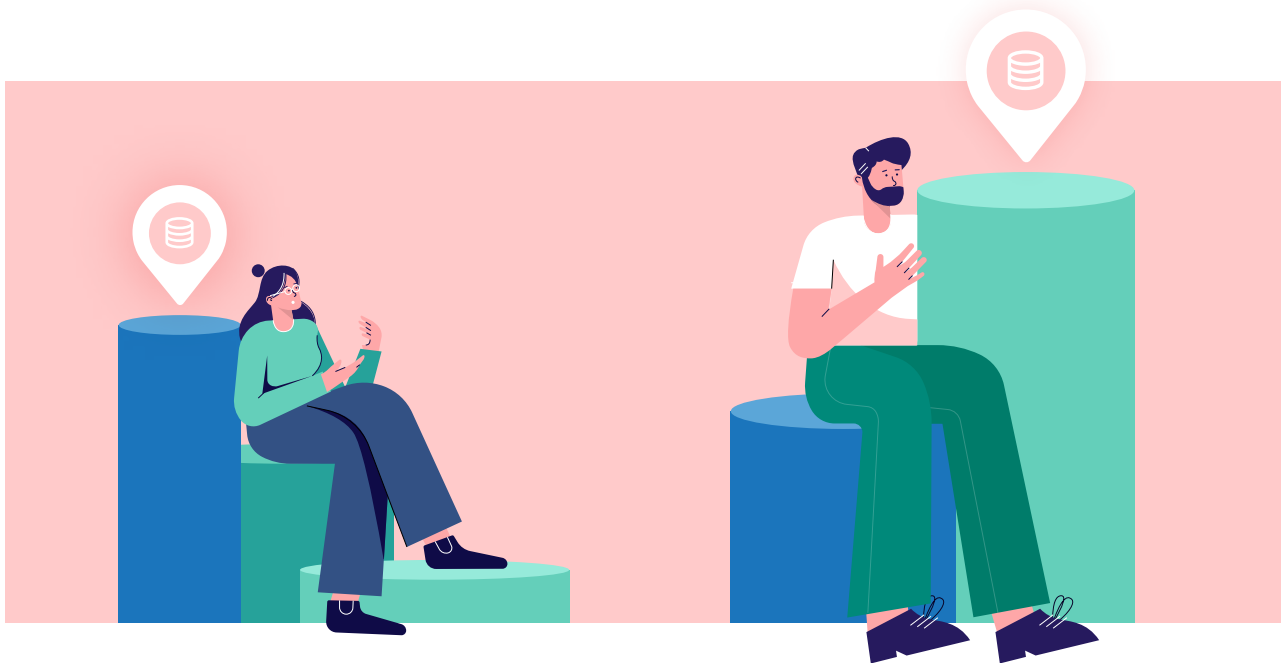
white paper

# How data silos are harming your business (and what you can do about it)

Pages: 6 | Reading time: 15min

[www.cluedin.com](http://www.cluedin.com)

# How data silos are harming your business (and what you can do about it)



## What is a data silo?

A data silo is a collection of data held by one group or department that is not easily discoverable or usable by others in the same organization. Every department needs data to operate, whether it be Finance, Marketing, Sales, HR, or Operations, and in many companies, each department will store its data in a different location. These different repositories are commonly referred to as data or information silos.

## Why do data silos occur?

Data silos occur naturally over time and tend to mirror the structure of the organization. As each department collects and stores its own data for its own purposes, it creates its own data silo. These silos grow over time as data volume and types increase. In a lot of companies, departments are used to working in their own spheres with their own processes and methods. Even when departments need to collaborate closely – such as Finance and Operations – company culture may dictate that they still keep their data separate. It comes as no surprise then that the higher the degree of separation in company culture, the less motivated employees are to share their data, and the higher the likelihood of well-entrenched data silos.

Compounding the problem is the fact that the technology and tools used by different departments have facilitated the creation of data silos. With the Finance team using accounting software, Operations using an ERP system, Sales and Marketing using a CRM, and the Product team using its own Product Information Management platform it's easy to see how data silos are created and grow over time. These often legacy systems were not designed to share information, and they each store and manage data using different formats and methods.

## Why are data silos a problem?

Historically data silos weren't considered a problem; it was perfectly natural for each department to accumulate, store and use its own data. All that changed with the advent of Big Data and Cloud computing, as organizations came to understand that in order to truly realize the potential of their data and take advantage of the scale and efficiencies of the Cloud, their segregated piles of data simply weren't going to cut it.

When data is siloed, it creates barriers to information sharing and collaboration. It prevents companies from getting a holistic view of their customers, products, employees, and more. Invariably, data quality also suffers due to inconsistencies in format, overlaps, and different cleaning regimes.

Siloed data is typically difficult to find and use in a timely manner and is often of poor quality. All of this means that it can't be used effectively to fuel business intelligence and analysis, and is inadequate for data science purposes such as Machine Learning and Artificial Intelligence. Any business embarking on Digital Transformation initiatives or that is attempting to pivot to a more data-driven operating model will struggle to do so if its data remains siloed.

In addition, siloed data is itself a risk as it is almost impossible to properly govern data that is buried and isolated. This obstructs regulatory compliance efforts and increases the likelihood that sensitive data will be mishandled.

# How data silos are slowly destroying your business

Businesses today operate in an environment that is constantly changing. Increased competition, economic and political disruption, and cost pressures are all contributing to the need to not only do more with less but to prioritize only those projects which will have a direct impact on either the bottom or top line. All of this is motivating organizations to extract more value and insight from their data, and to use it to maximize process and operational efficiencies.

Organizations with a technical debt of siloed data will struggle to achieve this. Here's why:

## 1. Data silos inhibit a 360 view

Data silos make it very difficult to build a comprehensive view of any data object – be it a customer, an employee, a location, or a product. Without that view, it is hard to ascertain how you can serve your customers better or identify enterprise-wide inefficiencies. How can you hope to foster customer loyalty and referrals if you don't know what they like most about your product or service, and what needs attention?

## 2. Data silos threaten trust and confidence in data

Imagine a scenario where you are presenting recommendations concerning the preferred location for a new manufacturing facility to the leadership team. Your recommendations are based on data gathered concerning where your employees are located, proximity to major transport routes, external suppliers and customers, transportation costs, rents, local labour costs, and so on. Now imagine that once you have made your recommendations, someone in the room challenges one or more of your assertions with their own data from their own system which you didn't know existed. When one point of reference is undermined, it raises doubt about the rest, and the whole proposal is damaged.

When the same information is stored by different departments in different systems, inconsistencies will invariably occur, and it can be a long and arduous process to establish which set is most reliable.

## 3. Data silos waste resources and drain productivity

When similar data is stored in different systems and used by different teams, resources suffer. Data storage may be relatively cheap, but human time and expertise are not. By keeping your data in its silos, the chances are that multiple users will be cleaning, deduplicating, and enriching similar data. They will also be creating their own separate sources of truth based on their own departmental data, which may or may not be consistent with the version created by other teams.

#### 4. Data silos hinder collaboration

A culture of separation encourages data silos, and in turn data silos encourage separation culture. One of the cornerstones of successful collaboration in the workplace is the sharing of data and information, and data-driven organizations make delivering an enterprise-wide view of data a priority.

## How to break down your data silos

Eliminating data silos is both a cultural and a technological challenge. Here's how to address both:

#### 1. Cultural change management

As with most cultural changes, it starts at the top and means leading by example. Senior leadership teams need to communicate the benefits of data sharing and reliability and link these benefits to the wider goals of the company. They also need to demonstrate their commitment by sponsoring data-driven projects that champion the sharing of information and data and demonstrating the results.

#### 2. Data integration

Data integration is the most effective technological means of preventing data silos. This can be done in a number of ways:

##### Scripting

Traditionally, IT teams have been tasked with writing scripts in SQL, Python, or other scripting languages to move data from silos into a data warehouse or other data management system. There are a number of disadvantages to this method. As the number and types of data sources grow, so does the complexity, and any changes will require a return trip to the IT department to update scripts. Manually-coded integrations can quickly become a time and cost burden for IT experts and business users alike, as business users are constantly waiting for their request to make it to the top of the list of IT jobs.

##### ETL systems

ETL (extract, transform, and load) tools automate the process of moving data from various sources to their target systems. These tools extract data from sources, transform it into a common format for analysis, and load the result into a data warehouse or other data management system. ETL tools can be either on-premises or Cloud-based, with Cloud providers typically making the ETL process easier and faster.

One of the major downsides of an ETL system is that you will need to have completed data modeling upfront, in order to tell the system what you want your data to look like. Often this means that by the time the data modeling exercise is complete (think anything from three months to a year), the data has changed, is out of date and the business has moved on.

### **ELT systems**

ELT (extract, load, and transform) tools are similar to ETL tools in that both are processes used to standardize and move data from one place to another. The key difference is where the transformation takes place. In the traditional ETL process, data is first extracted from its source, then transformed within an intermediary system like a Data Warehouse, and finally loaded into a destination system like a Data Lake or Master Data Management platform.

ELT focuses on moving the transformation process closer to the final destination. One of the primary benefits of ELT is that it eliminates the need for a separate transformation system and enables the data to be transformed in real-time. This can be especially useful for companies that need to analyze large amounts of data in real-time.

### **Master Data Management**

Master Data Management systems are designed to help organizations create a Master or “Golden” Record which is essentially a single version of the truth as it relates to a particular domain – be that a customer, location, product, employee, or supplier. Creating this single, trusted view involves a number of steps such as data ingestion, cleaning, integration, and enrichment. Old-school Master Data Management systems have taken an ETL-led approach and demanded that the data is already “transformed” and ready to be neatly slotted into a data model that you have pre-defined. More advanced, or augmented platforms, automate the transformation process and accept data in its raw state, allowing the natural data model to emerge using technologies such as Graph. This approach is far more scalable and removes the need for domain experts to sit together and map systems, and for data engineers to spend time building complex pipelines.

## **3. Enable governed self-service access**

When data is integrated and available as an enterprise-wide resource, IT professionals, domain experts, data stewards, and business users alike can all be given access to the data via a data governance framework. By removing gatekeepers and bottlenecks, business users with the correct permissions and access rights can easily access the data they need. Master Data Management systems enable this by taking a no/low-code approach and empowering business users to wrangle with the data themselves, without having to rely on IT.

## Shattering data silos for good

Data silos have a negative impact on productivity and collaboration, as well as undermining business intelligence and insights-based projects. In a world where businesses are constantly searching for both the marginal and major projects that will set them apart and continue to fuel growth, throttling your data assets in this way is completely counterintuitive.

Breaking down data silos is both a cultural and a technological undertaking, although traditional methods of integrating and sharing data using IT tools have often brought their own brand of problems due to their heavy reliance on IT teams to manage them. True data democratization is the complete opposite of segregated data, but can only be achieved when business users are empowered to access the data easily themselves and can use it at the speed with which the business needs to move.

Modern Master Data Management solutions offer data integration as part of the package, because they understand that as your company grows so will the number of sources and types of data you acquire. It simply isn't efficient or scalable to expect data stewards and engineers to manually integrate a data estate that is constantly expanding. Nor can traditional ETL tools do it at the speed with which your organization needs to move.

Of all the available ways to integrate data, the best approach is the one that eradicates the problem entirely. CluedIn does precisely this, by automating the entire process and delivering data that is ready for insight in the shortest possible time and at the lowest possible cost, allowing you to focus on the tasks that will generate the most value from your data.

**Learn more about CluedIn's  
unique approach to Master  
Data Management**

[Learn More](#)

