

# Why JUST Data Lineage, Data Discovery or a Business Glossary Isn't Going to Cut it for Your Data Team

eBook

---



## Why JUST Data Lineage, JUST Data Discovery or JUST a Business Glossary Isn't Going to Cut It for Your Data Team

Data processing has evolved. The single-use ETL, database, and analytics tools that BI teams have depended on for the last decade are no longer capable of providing a complete picture of a company's data.

For enterprise organizations to remain competitive, the tools they use to find and understand their data must work together:



The average enterprise depends on multiple tools and capabilities to handle these processes, often created by different vendors and usually incompatible. Teams are often challenged to come up with [a way to locate and understand data](#) on a faster timetable but are held back by incompatible systems.

**So how can companies bring all these critical functions together and improve their BI & Analytics team's accuracy, productivity, and turnaround?**

### BI Teams Are Challenged With Data Overload

The speed at which the data analysis world has changed is staggering. Just five years ago, a model of how BI reports were produced looked like a [straight line with three significant points](#) of interaction. Those would be:

1. The amount of data waiting for the BI team to process
2. A response to that data, which companies would meet by scaling up the size and data processing capabilities of the BI team
3. The expectations of the decision-making business consumers using reports produced by BI

The single-use tools we still see in modern enterprise organizations could handle the flow of data through this straight line. Stakeholders got the information they requested when they needed it. The system worked for the amounts of data most organizations had.

However, throughout the 2010s the amount of data being created, stored and utilized by large companies exploded. That familiar straight line began to look more like a valley. The amount of data became a steep cliff in the beginning while business consumers' expectations at the other end rose just as sharply. BI professionals were stuck in the middle and found themselves without the communication needed between their tools to keep up with the demand.

# Modern BI Depends On Information Coming From Many Different Systems

Business Intelligence teams found themselves trying to meet an ever-increasing mountain of data while still providing key personnel with enough information to make solid business decisions. Increasing staff size was no longer an effective solution; even the most veteran BI professionals can only process so much data in a day.

Adding additional technologies was the next solution.

That fix led to the current situation in the BI industry: The number of tools depended on to generate and process data ballooned. Today, 25% of enterprise organizations have approximately 400 data systems managed by at least 11 BI systems. These tools are in place for a variety of functions, but in general, they are pursuing the three most important processes of BI.

These are:

- Data Lineage
- Data Discover
- Business Glossary

Each of these systems is vital to processing the mountain of data that face modern BI teams. Let's examine what each process does and why they are so important for enterprise organizations.



## Data Discovery

One of the most important ways that BI teams can locate data that is scattered throughout their BI environment is through the use of [data discovery](#).

At its core, data discovery is the process through which data is collected and categorized from the many points of origin that enterprise businesses have. The data is combined and placed into a system that enables analysis. The overall objective of data discovery is to create a useful repository of information from scattered data points.

This operation is vital to creating value from the data at-hand and must occur before any further data analysis. Unfortunately, data discovery is a time-consuming and resource-intensive process. Today's BI & Analytics teams strive for faster discovery methods and are turning to automation.



## Data Lineage

Another crucial capability BI & Analytics teams must have access to is accurate data lineage. [Data lineage](#) is an operation through which BI and Analytics teams map the flow of data through an organization. Effective data lineage allows an analyst the opportunity to visualize the complete data movement process from where a particular data item originated, through to its target and everything that happened to it in between.

Like data discovery, data lineage is an important capability for business intelligence decision-making. Over the last decade, it has become even more crucial to large organizations to help them comply with international data protection regulations, for example.



## Business Glossary

A [business glossary](#) is another extremely important capability that helps the BI & Analytics team function more effectively. It is an enterprise-wide tool that, while often managed by the BI team, is in fact a product of every department, team, and sector of a company. The goal of a business glossary is to gather all terms used across the organization and unify their meanings so that everyone can be on the same page when it comes to business terminology.

It is crucial in ensuring company-wide consistency but is widely known to be very difficult, if not impossible to create and maintain, which is why most organizations never actually get around to building one. The standard process of creating a business glossary requires the participation of key members from every single department, and all terms must be debated and defined to eliminate overlap. This can be a project that takes months; eating up time and resources better spent preparing the analysis a company needs.

## Automated BI Intelligence Improves Results All Around

While each of these three functions are vital to running a productive and efficient BI department, they all require a great investment of time to be successful. Overcommitting to one process creates negligence in another and the overall quality of BI suffers as the team struggles to keep up with the overwhelming amount of information waiting for processing.

The key to tackling the increasing amount of data BI teams face lies in **automation**. Specifically, by leveraging machine learning technology, organizations can complete these processes within seconds instead of weeks, or sometimes even months.

Let's dive into some specific cases of how automation improves the effectiveness of BI.

## Automated Data Discovery

Fast and accurate data discovery and metadata analysis ensure BI teams can deliver faster and more accurately to the business, enabling them to make better business decisions.

The improvements that [automated data discovery](#) provide are visible in the case of [Mimun Yashir](#). This leading consumer credit company's BI team was crippled by a change in the zip code system of their home country, Israel. The process of finding every field in which a zip code was used and trying to create an impact analysis for the change with manual data discovery was far more time consuming than it should have been.

After this experience, Mimun Yashir partnered with Octopai to begin using automated data discovery tools for future projects and to inform impact analysis reports. Their six-person BI team now utilizes Octopai 60 times per week, performing the work of two additional full-time employees. After two years with Octopai, Mimun Yashir reported an 85% increase in efficiency for their BI department.

Automation is the key component to enabling stronger BI. Tasks that once required the focus of the entire team to finish even once are now repeatable. This routine data discovery keeps an organization's data map current, aiding in faster report generation and easier error detection.

BI teams no longer struggle with an insurmountable wall of data and can instead focus on creating quality reports for business consumers that need the information the team generates.

## Automated Data Lineage

With automated data lineage, BI teams are able to increase accuracy and save significant time and effort as they no longer need to manually map their data to understand where it originated, what changes it went through, and where it's going.

One customer, Big Ass Fans, an industrial ceiling fans and lighting system manufacturer, took advantage of exactly this. The company's BI team needed to streamline metadata management and eliminate redundancies in its reporting while preparing for a massive system migration.

Big Ass Fan's BI analysts were unable to follow data points back to their origin in a timely manner, which was undermining the authority and value these reports had for consumers down the line. Octopai's automated data lineage and discovery allowed the completion of data mapping projects in seconds, which left the BI team confident in their ability to streamline their migration from legacy data warehousing systems to Snowflake.

The benefits of automated data lineage have a measurable effect on the costs required to run a strong BI team. Difficult and time-consuming projects, such as [system migrations](#), can often take months. But they become simple tasks when BI teams can quickly identify and sort what needs to be migrated and what can stay. In turn, this can lead to a significant reduction in the volume of

reports and processes that need to be moved over to new systems.

This is especially evident in another example, where one customer conducted 704 data lineage searches in a single day using Octopai. Each search saved more than half an hour of manual mapping work, so the team eliminated the need for more than 350 work hours, at minimum, in one day.

Due to the labor-intensive nature of performing data lineage tasks by hand, automation doesn't just make teams move faster, it also delivers measurable economic benefits to the organization. If you consider a BI analyst will take home roughly \$15,000 per month and can only output roughly 180 hours per month, that customer saved more than [\\$29,000 in one day's](#) worth of searches.

Automated data lineage delivers higher quality results. By automatically updating all lineage information for any data in the system, BI teams can create compelling illustrations of how data gets from one point to another and [perform better impact analysis](#).

And remember those regulatory bodies that require data lineage processes to keep companies in compliance? They can request an audit at any time. If companies cannot produce those reports within the requested time frame – a challenge for many manual BI teams – they are liable for heavy fines.

Automated data lineage helps companies prepare for these challenges and then exceed expectations, making it a vital component of modern business intelligence.

## Automated Business Glossary

BI & Analytics team productivity can be increased even more so by adding a third capability: the automated business glossary.

The initial time investment in completing a business glossary is one of the biggest hurdles in revitalizing a BI team that has fallen behind. So how can these teams put together a better business glossary in less time? Again, the solution lies in automation. An automated business glossary scans all the metadata resources throughout the BI environment (reporting, database, and ETL), isolating overlapping terms, and creating a unified resource that serves the entire organization.

[Menora Mivtachim](#) is one of the largest Israeli insurance providers. As a result, they had a complex metadata landscape that was being managed by an array of tools including Oracle, DataStage ETL, Cognos, and Qlik Sense. However, the difficulty of communicating across a shared data catalog left the BI team with a lack of control over their data assets.

To corral these tools, the team needed to bring all these pieces of software together. Key to this was Octopai's central data automation platform and the business glossary it could help establish. Without automation, this process could take weeks. Within five minutes of work with Octopai, Menora Mivtachim had a business glossary that could define new items to implement in the data warehouse.

The glossary contains a complete compilation of all terms used by the company, assigning them values and definitions that now will inform both data discovery, data lineage, and users. An automated business glossary will continue to monitor those resources, adding new terms as necessary while preventing redundancies and errors. This creates a single source of truth trusted even in the most expansive enterprise organizations.

## The Winning Team: Automated Data Discovery + Data Lineage + Business Glossary *Working Together To Provide The Whole Story About The Data*

The benefits of each of these functions are vital to the business intelligence of a company. However, when these functions operate in isolated technology stacks, they cannot overcome the valley of data overload and consumer expectations challenging BI teams today.

If tools from different vendors cannot communicate with one another, as is most often the case, the available information exists on isolated islands of data. Despite having the information they need, BI teams now must spend their time translating information from one format to the other. This process can be just as tedious and time-consuming as trying to build a business manual or perform data lineage manually.

To escape the hole in which so many BI & Analytics teams find themselves, they need to not only automate data discovery, data lineage, and their business glossary, but it is also critical that these capabilities **work together**.

This is what an automated BI Intelligence platform such as [Octopai](#) provides. With Octopai, these tools now have a way to share information, and BI teams have a way to maximize their understanding of what the data means. Mountains of data become manageable, and the expectations of the key stakeholders who need actionable BI are met.

The benefits of an automated and unified system supporting data discovery, data lineage, and a business glossary will bring to your organization will impact all elements of BI & Analytics.

Some of the immediate advantages include:

- The ability to visualize the entire data supply chain
- A new understanding of how to maximize efficiency
- A faster way to deliver analytics throughout the organization

Octopai requires zero installation and there are no professional services required to get up and running. BI & Analytics teams can literally start using Octopai within a couple of hours once their metadata is uploaded to the platform.

Contrast that with weeks and sometimes months (not to mention the cost) it would take to revamp an entire BI technology stack. Octopai will get your BI team the results they need, **now**.





## Get Smarter About The Way You Handle Your Data With Automated BI Intelligence

Investing in just data discovery, only using data lineage, or relying solely on the construction of a business glossary is not enough for modern data teams. Utilizing just one of these aspects of business intelligence will keep analysts trapped in a valley of high expectations and data overload. The way out of that valley is an automated BI platform that brings all three of these capabilities together.

Modern BI challenges require tomorrow's BI solutions. Investing in an automated BI intelligence platform gives data teams complete visibility into their data flow so they can be smarter about the way they handle their data and ultimately make better business decisions. Octopai's unified platform, comprising automated data lineage, discovery, and business glossary, forms the winning team that will make business intelligence teams more intelligent.