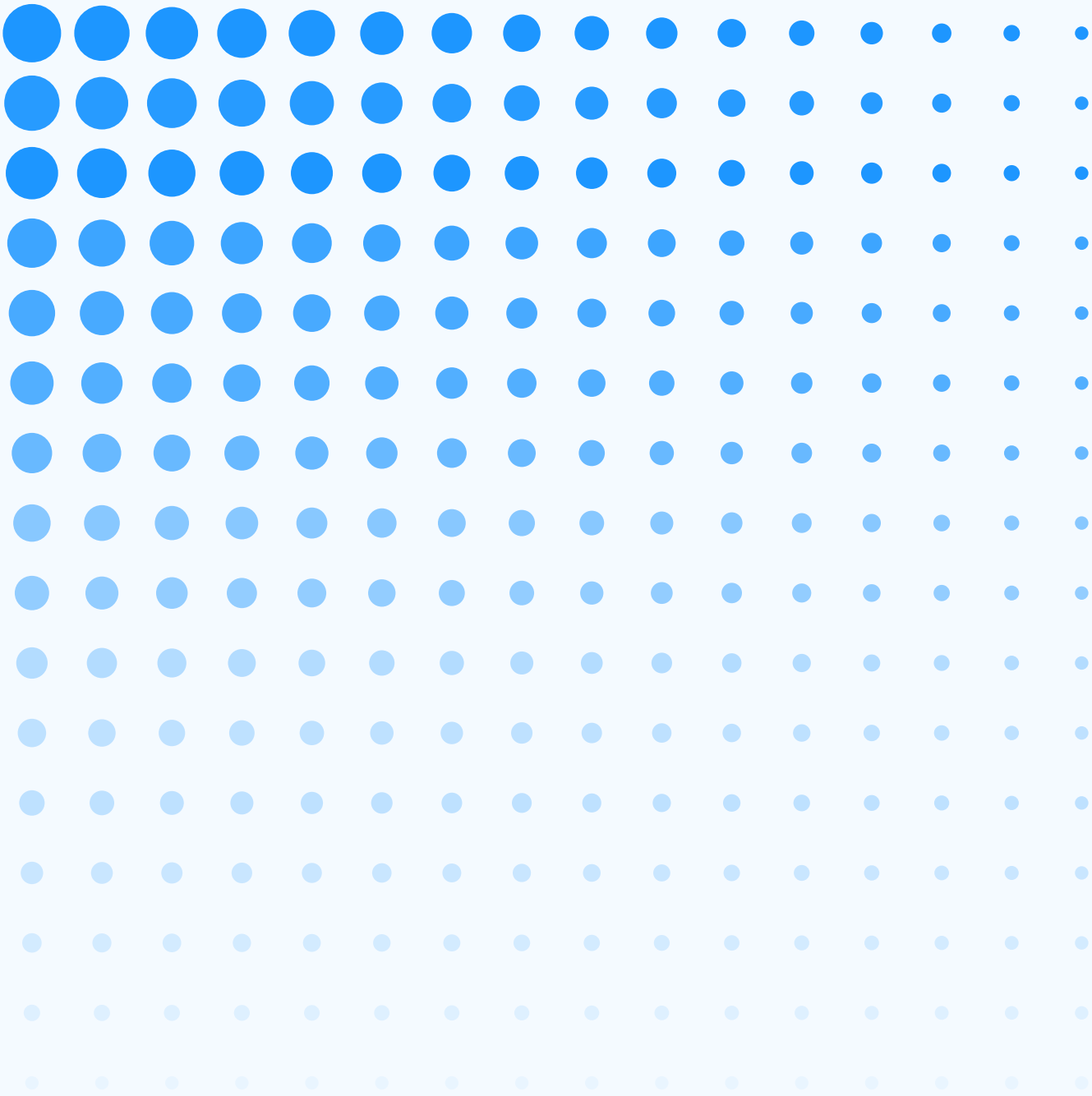
# Reverse ETL

## ENABLING SMARTER DECISIONS WITH OPERATIONAL ANALYTICS



integrate.io

# Table Of Contents

# 01

## INTRODUCTION

In order to beat competitors and better serve customers, businesses need fresh, cutting-edge insights supported by data-driven processes and workflows. With so much information on hand, however, obtaining these insights is practically impossible without a way to collect them in a centralized location.

Enterprise data can take many forms, from detailed profiles of individual customers to financial predictions and forecasts. Assembling, cleansing, and organizing your enterprise data is the goal of the ETL process, creating a "single source of truth." Within this centralized data repository, users can run business intelligence (BI) and analytics workloads, giving decision-makers access to dashboards and key reports.

But this one-way flow from sources to the target repository is just one possibility for taking advantage of your business data. In many cases, it's just as valuable to reverse the direction: migrating information out of your data warehouse and into dedicated third-party systems for specialized queries and analytics.

When data travels in the opposite direction—using the data warehouse as the source and various third-party systems as the destination—it is known as "reverse ETL."

Using reverse ETL judiciously can help your employees get the information and insights they need to excel at their jobs on a day-to-day basis.

So how does reverse ETL work, exactly, and how can it specifically benefit your organization? In this white paper, we'll discuss everything you need to know about reverse ETL, including:

The definition of reverse ETL

The difference between "traditional" ETL processes and reverse ETL

The relationship between reverse ETL and operational analytics

The benefits, use cases, and applications of reverse ETL

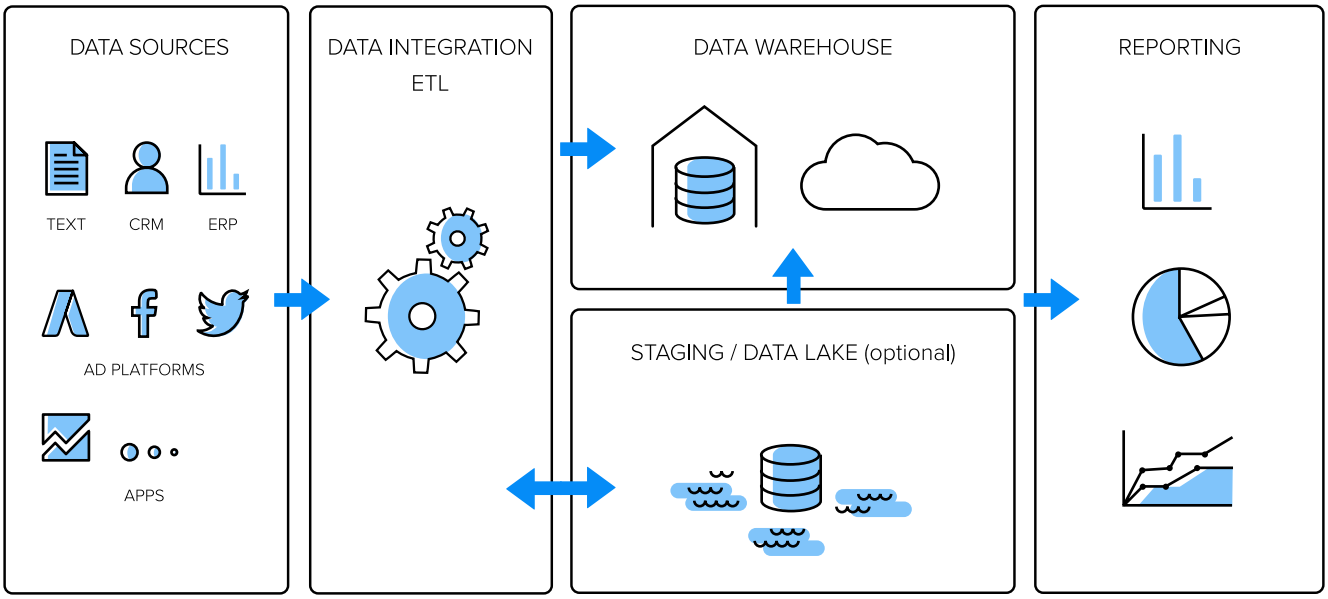
How to implement reverse ETL

How to select the right reverse ETL tool for your business

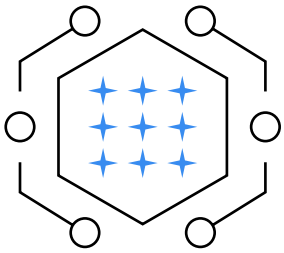
# 02

## WHAT IS “TRADITIONAL” ETL?

ETL (extract, transform, load) is the term for a three-step process that is fundamental to modern data management. The goal of ETL is to ingest data from various locations and collect it in a centralized repository for analytics and reporting. Since its emergence in the 1970s, ETL has been the backbone of organizations’ data integration efforts, ensuring that organizations always have access to the freshest, most accurate information.



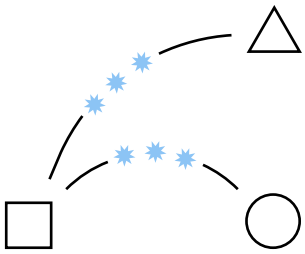
The three steps of the ETL process are as follows:



## Extract

First, data is extracted from one or more source locations. The sources of information for ETL may include:

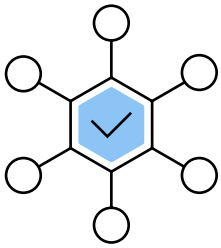
- Relational and non-relational databases
- Flat files (e.g. XML, JSON, CSV, Microsoft Excel spreadsheets, and others)
- Text files
- SaaS applications and platforms
- APIs (application programming interfaces)
- Third-party websites
- Analytics and monitoring tools
- System logs and metadata



## Transform

Second, the extracted data is transformed in various ways, preparing it for storage in a centralized location. There are many possible goals during the transformation stage, depending on the contents of the source data and the format of the target repository. Potential data transformations include:

- Rearranging unstructured data into a structured format
- Limiting the data you've extracted to just a few fields
- Sorting the data so that all the columns are in a certain order
- Joining multiple tables together
- Cleaning the data to eliminate duplicate and out-of-date records



## Load

Finally, the transformed data loads into a centralized target location. The most common target is *a data warehouse*: a data management system that stores large amounts of structured information. A data warehouse is usually optimized for business intelligence and analytics. The target of ETL may also be a data lake: a centralized repository that can store both structured and unstructured information.

This three-step ETL process—extract, transform, and load—has been the dominant form of data integration for decades. However, it's not the only data integration option available. For example, ELT (extract, load, transform) flips the order of the final two stages, loading raw data into a data lake before transforming it. During ELT data is transformed inside the data lake on the fly, as deemed necessary for analytics and reporting. This approach offers faster loading speeds and greater flexibility; however, ELT is a newer technology that doesn't yet have the institutional support of ETL.

Yet another data integration technique is known as ETLT (extract, transform, load, transform). This method attempts to offer the best of both worlds from both ETL and ELT. The first transformation stage in ETLT consists of "light" transformations, such as data cleansing or data masking to conceal sensitive information. The second transformation stage, performed inside the data lake, consists of more complex actions, such as integrating multiple data sources.

# 03

## WHAT IS REVERSE ETL?

As the name suggests, reverse ETL inverts the source and target of the traditional ETL process. During reverse ETL, the standard ETL targets (e.g. data warehouses and data lakes) are instead treated as sources, while the standard ETL sources are treated as targets.

In her article "Reverse ETL—A Primer," Astasia Myers, investor at Redpoint Ventures, defines reverse ETL as "the process of moving data from a data warehouse into third-party systems to make data operational." These destinations may be SaaS (software as a service) platforms, databases, cloud storage, or any other systems used by sales, marketing, and customer support professionals.

For example, the targets of reverse ETL are often third-party SaaS applications that serve as data sources in standard ETL. These targets may include CRM (customer relationship management) and ERP (enterprise resource planning) systems that are specialized for different purposes and collect different kinds of information.

More specifically, CRM platforms collect all manner of data on your customers and potential customers—from their first interaction with your business to their most recent purchase, website visit, or call to customer support. ERP systems, meanwhile, collect data for functions such as human resources, finance, accounting, logistics, and manufacturing (depending on your choice of solution).

CRM, ERP, and other SaaS software offer valuable data for the standard ETL process—but they also have powerful BI and analytics functionality of their own. For this reason, it can sometimes be smarter to run queries and generate reports from within these third-party applications, rather than performing them within the data warehouse itself.

In the reverse ETL paradigm, your data warehouse or data lake is the “single source of truth” that stores the most accurate, up-to-date information, rather than running analytics workloads itself. Rather, this repository serves as a centralized hub that propagates information to the places and people that need it.

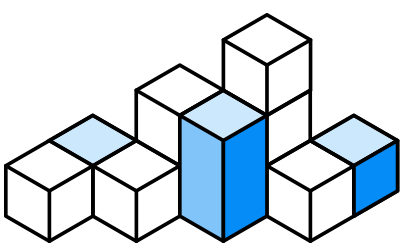
# 04

## ETL VS. REVERSE ETL

The mechanics of standard ETL and reverse ETL are largely the same: data is extracted from sources, transformed as necessary, and loaded into a target location. By simply flipping the role of source and target, however, reverse ETL offers a radical new paradigm for data integration.

Below are the most important differences between ETL and reverse ETL:

- **Sources and destinations**: ETL transfers data from one or more source locations into a single centralized repository, such as a data warehouse or data lake. Reverse ETL transfers data out of this centralized repository into one or more operational systems.
- **Users**: ETL is intended for use by IT professionals and data analysts who are skilled at running business intelligence workloads inside the data warehouse. Reverse ETL "democratizes" analytics by making data available to the users of these operational systems, many of whom are less technically skilled.
- **Applications**: Finally, the use cases of ETL and reverse ETL are different. In standard ETL, data scientists and business analysts access enterprise data from within a centralized repository, using it to create dashboards and reports. In reverse ETL, business users (e.g. members of your sales, marketing, and customer support teams) can access fresh, accurate, up-to-date enterprise data from within the specialized systems that they use.



# 05

## OPERATIONAL ANALYTICS AND OPERATIONAL SYSTEMS: WHY THE NEED FOR REVERSE ETL?

"Reverse ETL" sounds like a fancy, revolutionary idea, but the underlying concept of reverse ETL has actually been around for some time. The difference is that it's been operating under a different name: "operational analytics."

In traditional BI and analytics workflows, data and insights are presented in the form of dashboards and reports for consumption by key decision-makers. This information typically takes a broad, overarching view of your organization, helping inform long-term strategic business decisions.

For example: Which markets are the best options for expansion? How well is the new product line performing in comparison to existing offerings? What are the best ways to reach and acquire new customers? These questions, and more, are the domain of traditional analytics.

The term "operational analytics," on the other hand, indicates that your data is being operationalized, i.e. used for day-to-day business activities by the teams who need it. In order to excel at their jobs, sales, marketing, customer support, and product development teams all need access to the latest information and cutting-edge insights.

That's exactly the goal of operational analytics: leveraging your enterprise data across the entire organization for the good of all your employees—rather than keeping it locked up in a single centralized location, accessible only to people with the right technical skills.

As such, operational analytics facilitate tactical, rather than strategic, decision-making. Below are just a few examples of operational analytics in action:

- Product managers using a mobile app's metrics and KPIs (e.g. customer retention rate) to determine which features or bug fixes to prioritize.
- An e-commerce store running A/B tests to discover which website redesign results in lower cart abandonment rates.
- Marketing teams deciding which customers to target with their next campaign by using their "customer 360" views of the company's clients and prospective clients.

If operational analytics is the key to these scenarios and more, then reverse ETL is the means to achieve operational analytics. Reverse ETL pushes data out of your data warehouse or data lake and into the operational systems that your employees actually use on a daily basis.

# 06

## 5 USE CASES OF REVERSE ETL

We've defined reverse ETL and discussed the relationship between reverse ETL and operational analytics. Next, we'll look at some common use cases and applications of reverse ETL within a typical business.

### 1. Customer Relationship Management (CRM) Software

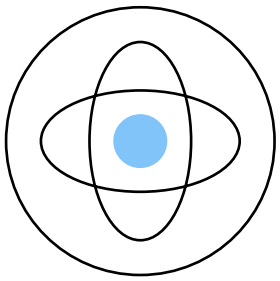
One of the biggest use cases of reverse ETL is for customer data, transferring it to CRM systems such as Salesforce. Users of this CRM software can then take advantage of this fresh information to enable intelligent, up-to-the-minute decision-making.

Moving your enterprise data out of the data warehouse can help you take advantage of your CRM software's specialized functionality. For example, reverse ETL and operational analytics can help you quickly and easily identify individual customers and customer segments to focus on—like sorting your customers by lifetime value to find the biggest spenders, or launching a "win-back campaign" to re-engage customers who haven't made a purchase in a while.

### 2. Enterprise Resource Planning (ERP) Software

ERP systems are already packed with a wide range of functionality, from finance and accounting to logistics, human resources, manufacturing, and more. Reverse ETL and operational analytics can also facilitate smarter use of your ERP system, leveraging its own analytical capabilities to deliver new ideas and insights.

There are numerous use cases of reverse ETL and operational analytics for ERP software. For example, enterprise data can help ERP users manage your product inventory, placing new orders in a timely manner to account for historical spikes and declines in consumer demand.



## 3. Customer Support

Your customer support team is on the front lines of your business, serving as a direct liaison between your clients and the rest of your organization. Reverse ETL is a key technique to deliver the data and insights they need to excel at their jobs.

Having access to product usage data, for example, could help your customer service agents redesign a better support workflow, helping address the most common issues and questions. Meanwhile, customer profiles could help these agents prioritize clients' roles in the support queue—perhaps by serving important or repeat customers first.

## 4. Business Intelligence (BI) and Analytics

Although data warehouses offer powerful analytical capabilities, you might sometimes prefer to export your information into a dedicated BI and analytics platform. Reverse ETL can also move the data in your data warehouse into your BI and analytics platform of choice, such as Tableau, Qlik, or Microsoft Power BI.

Once the data is inside your BI system, you can run queries and generate dashboards, reports, and visualizations on an ad hoc basis. Real-time BI and analytics solutions give users the power to view and manipulate data to make better-informed tactical decisions.

## 5. Data Governance

Last but not least, adopting an enterprise-grade reverse ETL solution also helps you improve your data governance. In enterprise IT, "data governance" refers to the framework of policies and regulations that govern your organization's data management.

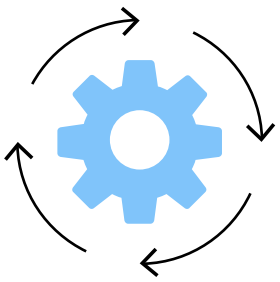
A single, unified reverse ETL platform is much easier to manage and monitor than multiple API connectors, each of which requires individual maintenance. Implementing reverse ETL can save your IT team a great deal of time, making them available to perform more important duties.

# 07

## HOW TO IMPLEMENT REVERSE ETLT

We've talked about what reverse ETL is, as well as its various benefits and use cases. The next question is: If you want to implement reverse ETL within your own organization, what are your options?

Although the concept of reverse ETL has been around for decades, it's only recently that it has been explicitly defined and incorporated into the ETL family. Before the availability of reverse ETL solutions, each organization would have to manually create its own API (application programming interface) connectors between the data warehouse and their third-party operational systems. What's more, this API creation process would have to be repeated for each destination system—including Salesforce, Marketo, HubSpot, and others.



This approach has multiple disadvantages:

- Manually creating API connectors from scratch can take several days to a week or more, even if you have a technically qualified in-house team.
- Writing connectors is challenging due to "brittle" API endpoints that cannot handle real-time data transfer.
- Your team needs to maintain these connectors over time to account for any changes in the underlying technologies at either end.

Rather than creating your own connectors, it's a much wiser decision to choose a purpose-built ETL (or reverse ETL) solution. ETL tools come with built-in, "plug and play" integrations that dramatically simplify constructing data integration pipelines. The good news is that many existing ETL tools can also operate in the reverse direction, pumping information from your data warehouse to the target third-party systems.

# 08

## A BUYER'S GUIDE FOR REVERSE ETL TOOLS

In nearly all cases, going with a pre-built ETL tool is a much better choice than trying to build data integration pipelines by hand. Of course, not all ETL and reverse ETL tools are created equal—so how can you decide which one is best for your purposes?

Below, we'll offer some of the most important criteria you should consider when evaluating reverse ETL tools.

Connectors and integrations: As discussed above, offering pre-built connectors is a crucial concern for reverse ETL tools. Does the tool you want to use easily integrate with your existing technology stack? If not, can you build your own connectors to fill in the gaps, or can you change your tech stack to compensate? In particular, if your business or IT needs may change in the future, look for a flexible reverse ETL solution with a diverse offering of connectors and integrations.

**Ease of use**: Some reverse ETL tools are intended for anyone to use, with low-code or no-code functionality that does all the work behind the scenes. Other tools, however, are for technically literate users who already have some data integration experience. Whichever scenario applies to you, make sure you choose a tool that fits your situation.

**Customer support**: If you're less experienced with ETL and reverse ETL, consider not only ease of use but also the strength of the tool's customer support. When you have a question, issue, or problem, how can you get the help you need in a timely and professional manner? Ask about the tool's customer support options, which may include phone, email, live chat, user forums, FAQs, manuals, or knowledge bases.

**Pricing**: ETL tools range from cost-effective to very expensive, depending on the features and functionality they offer. But while budget is one concern, different ETL tools also have different pricing models. For example, some tools base their pricing on usage (e.g. the number of connectors, or the number of CPU hours used), while others separate their offerings into tiers, charging a flat fee for the members of each tier.

**Data security**: Much of the data handled during ETL and reverse ETL is personal, sensitive, or confidential. For example, customer profiles may contain PII (personally identifiable information), while ERP systems may contain sensitive data about your company's internal financial reporting. Your choice of reverse ETL tool should offer data protection techniques such as data masking and encryption, especially if you need to comply with data regulations such as GDPR or HIPAA.

# 09

## HOW INTEGRATE.IO CAN HELP WITH REVERSE ETL

From CRM and ERP to better customer support and smarter data-driven decision-making, the potential advantages and use cases of reverse ETL are nearly limitless. But how can you actually go about implementing reverse ETL for your own purposes?

With small, one-off reverse ETL jobs—and for businesses with the right technical chops—building your own API connector can be a feasible goal. But what if you need to implement reverse ETL on a large scale within your organization, or you don't have the necessary technical skills required for the task?

That's where ETL and data integration tools like Integrate.io come in. Integrate.io is a powerful, feature-rich, automated ETL and data integration platform that helps you link your third-party sources and systems with your cloud data warehouse or data lake.

With over 100 pre-built connectors and integrations, Integrate.io makes it easy for any organization to build ETL and reverse ETL pipelines, as best fits your needs. The connectors supported by Integrate.io include:

- Databases (MySQL, PostgreSQL, Microsoft SQL Server, MongoDB, Oracle, IBM Db2, and others)
- Data warehouses (Amazon Redshift, Google BigQuery, Azure Synapse Analytics, Snowflake, and others)
- Cloud storage (Amazon S3, Google Cloud Storage, Azure Blob Storage, HDFS, SFTP/FTPS, and others)
- SaaS platforms (Asana, Crunchbase, GitHub, Eventbrite, Google Drive, Slack, Twilio, Stripe, and others)
- Analytics platforms (Google Analytics, Mixpanel, Buffer, Segment, and others)
- Advertising platforms (Facebook, YouTube, Outbrain, Google AdWords, and others)
- Logging tools (CloudTrail, Loggly, and others)
- Business intelligence tools (Looker, QlikView, Chartio, and others)

What's more, the Integrate.io platform comes with a user-friendly, no-code, drag-and-drop interface, so that anyone at any technical level can build robust data pipelines and move your data where it needs to go. By taking advantage of Integrate.io's rich suite of third-party integrations, it's never been easier to implement your own customized reverse ETL solution.

Ready to learn more about how Integrate.io can help with ETL and reverse ETL? Get in touch with our team of data experts today for a chat about your business goals and requirements, or to start your 14-day pilot of the Integrate.io platform.