

 atrium snowflake® + a b | e a u®

**Data Agility Unlocked:**  
The Data Strategy You  
Need to Succeed

# Your data is only as meaningful as you make it.

Without the proper accessibility, flexibility, and actionability, data is just another tool on the shelf. How can you move faster and keep pace with market changes if your data isn't agile?

Before we can answer that question, we have to look at the evolution of data agility. The age of "big data" has evolved to cloud data. Formerly cutting-edge systems are now seen as legacy, requiring a lot of resources to maintain. Businesses have started to streamline operations and replace legacy on-premise data lakes and warehouses with systems like Snowflake, a Data Cloud that doesn't require years of planning to scale. We now have the flexibility to land data for immediate accessibility, so that agile product teams can transform data iteratively for rapid value creation.

But in order to take advantage of cloud data platforms, the concept of data agility has to be at the forefront of your strategy and process. How quickly can you derive value from all your data? How quickly can you not only surface insights, but also take action?



*It's not about having mountains of data to hand over to your analysts and other business stakeholders — it's about making it easy and fast to find insights that are essential to the business and act on them.*

It is also about doing these things with strong roles and processes to manage governance, security, and accountability for quality.

## Data Agility Drives Innovation

Legacy systems can still be great at storing and processing huge amounts of data, but getting that data into a usable format tends to require someone who specializes in the exact technology. And that person, with their technological expertise, doesn't often understand the inner workings and needs of the business. A widening gap between IT and the business can stall and ultimately halt progress altogether.

Having a single system that provides an industry standard mechanism for data analysis allows business analysts to sift through the data directly and discover insights — without having to negotiate and translate business needs into IT actions. Knowledge sharing and continuity within one team also allows quicker iterations of discovery and shifts in analysis as business needs change.

### 2 TRENDS TO WATCH:

## Agile Data Teams and Self-Service Data Platforms

At Atrium, we've spent years working with businesses across to develop data strategies, implement analytics and data science initiatives, and to establish scalable architectures from data pipelines to shaping. Two clear patterns have emerged in the last few years.

1. The formation of agile data teams
2. The use of self-service data platforms

## Agile Data Teams

We're starting to see the enterprise data warehouse (aka EDW) and analytics space shift in the direction where business data owners, subject matter experts, and data architects and engineers operate in blended teams aligned to business functions. Part of the driver for this is an increase in demand for data insights, particularly as companies look to scale their business with insights vs. employee growth.

We've also seen a stronger sense of ownership from business leaders who want to be more hands-on in delivery. This is something businesses have seen in the past in the context of user experience, and the more recent takeaway is that it also cuts down the lead time for initial data acquisition and engineering through a closer pairing of skills.

## Self-Service Data Platforms

We've seen a lot of maturation in the cloud data storage space over the last few years, in particular, the ability of platforms like Snowflake and AWS to provide business access to data sets, which allow distributed data teams to shape analytic data without risk to the integrity of "golden datasets" (i.e., clean, validated, integrated datasets), which are typically more locked down.

By opening up accessibility to a wider cross section of the organization, businesses can expand their available resource pool for working on data.

It's time for businesses to ask the question:



*Is there a conversation internally around how to enable teams to better self-service so that they can engineer their own internal (or external) data products?*

## AGILE DATA IN ACTION: 3 CASE STUDIES

As we observed these patterns in action across our customers, we found varying levels of maturity in their agile data mindsets. Here's a closer look at the diverse approaches to data agility and subsequent results across three very different industries.

### Agile Data Case Study: Food Distribution

Our customer has been focused on improving operational agility in response to constantly fluctuating supply and demand across account segments, product lines, and geography. They've invested heavily in Snowflake and have been working to onboard sales, order, and supply chain data so that it can be consumed via self-service across departments. This shift has been hugely beneficial for them and has helped our team quickly deliver a set of account performance and product recommendation dashboards to their sales organization.

However, one of their lessons learned has been the need for stronger data architecture governance. They have a scenario in which their golden datasets live in Snowflake and various departments are enabled to customize their own dashboards in Tableau. This resulted in a lot of siloed Tableau customization.

Positive next steps:

*We've been working with this customer to unwind their reporting tool customization and to better use Snowflake to house all shaped data.*

## Data-as-a-Service Case Study: Education

Our customer owns several for-profit online universities and is a couple of years into their data-as-a-service journey. Their strategy is to create predictive data services that any of their universities can use — from student enrollment scoring to student success and retention. Additionally, their strategy is to API enable their models and feature store, housed on Google Cloud Platform, so that they can be used across their 10+ Salesforce orgs and other student systems.

Their two biggest challenges are: resource constraints to get data moved onto their cloud platform so they can retire old databases, and establishing common data definitions with discoverability around datasets.

### Our solution:

- ✓ Create a data governance office to solve these issues
- ✓ Create “golden event streams” to support a data-driven workflow and digital experience around the student journey
- ✓ Prioritize work based on use case within the student journey, mapping back to common data elements that support clusters of use case

## Data-as-a-Product Case Study: Manufacturing

Of the three, this customer is furthest along in their agile data mindset. Their business has full accountability for engineered datasets and data products, and has started offering predictive part replacement as a new go-to-market product for airlines. They were experiencing a number of failures in a critical operating system, consisting of 50+ parts and assemblies. This is critical, as every hour on the ground costs airlines thousands of dollars.

### Our solution:



Airflow was implemented on the Heroku platform to facilitate the ongoing ETL and model productionization, with predictions and insights being surfaced through a series of Tableau dashboards



Discovery was expanded to capture critical components by incorporating in-flight messaging data, identifying the number of cycles (takeoffs/landings) which had the highest correlation for failure rates



Execute a part failure analysis pilot using multiple data sources including flight usage and repair order details to identify the components with the highest likelihood to fail

The insights our customer gained as a result led to process improvements for proactive parts replacement, driving significant ROI for them. It also helped improve their ability to manage service-level agreements, as well as better monetize service contracts with their customers. For future enhancements, they plan on expanding the service offering to additional parts and even provide a route stress index to their customers.

## Innovate Faster in a Self-Service Architecture

Let's get back to the trend of self-service data platforms. With enough effort and budget, AWS, Azure, and Google Cloud Platform can all enable some level of self-service capability for distributed data teams, though Snowflake consistently stands out as the clear frontrunner in their ability to balance data accessibility with security and governance. Why?

*Consumers of the data platform may be internal departments or business units, data science teams, or external customers and partners. They're not just accessing data via reporting tools — rather, they have table-level access to enrich and engineer data.*

Snowflake is unique in that it gives us the ability to administer security policies once on the data, and then allow that data to be cloned as a new database on the platform, so that the data can be reshaped and enriched across teams, but without risking security or integrity. This is because every clone database ultimately points to the original data set, and inherits the associated security and governance policies. This is a really important concept for enterprise data warehouses and enterprise architecture teams who are focused on establishing a single source of truth, and as a result have historically limited access to their enterprise data warehouse.

## Agile Data Architecture

The data architecture pattern we most commonly see segments data domains into three broad categories that are as follows — source data domains (i.e. “raw”), curated data domains (i.e., “aggregated”), and consumption data domains (i.e., “aggregated and combined”):

### SOURCE DATA DOMAINS

- Historical sources of truth
- Align to source data structures
- Set security and access policies

### CURATED DATA DOMAINS

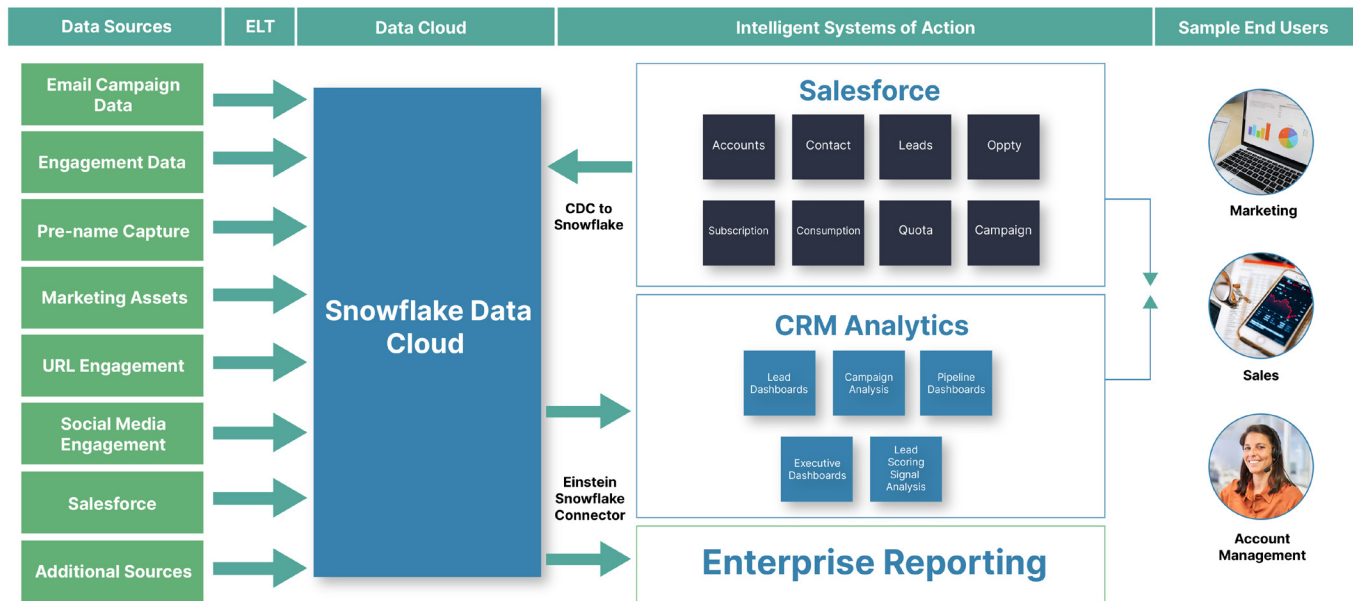
- Combined and shaped for broad usage
- Inherited security and access policies

### CONSUMPTION DATA DOMAINS

- Tailored for use cases such as machine learning feature stores and purpose-built analytics
- Inherited security and access policies

Before you can dive into the data deep end, consider the graphic below as a helpful guide to understand these different domains and the mechanisms in place to support end users, with our example spanning marketing, sales, and account management. “Self-service” architecture doesn’t begin that way; only through the careful mapping and integration of these different tools and mechanisms can end users reap the benefits on their own terms.

## The Road to Self-Service Architecture



Source data domains most closely align to the source data structures as they land in Snowflake, and security policies are often set here. With curated data domains, these databases represent data that has been shaped and curated for broad consumption, such as customer 360 datasets. With consumption data domains, we have databases that have been formed for specific use cases or clusters of use cases. These might be machine learning feature stores or data shaped for departmental analytics, such as sales forecasting or demand planning.

With so much to consider, a good rule is to start with a smaller set of use cases. We help our customers not only as they think about adopting these concepts, but also as they work through tactical and large-scale data strategy planning.

## Partner with Certified Snowflake Experts

The right partner, like the right data strategy, can help set you on the path to success... and help you stay there in the long term. Our team of data experts and certified SnowPros at Atrium can help you take advantage of Snowflake’s Data Cloud, unlock the value of your data, and take action on timely insights in your other everyday systems such as Salesforce, Tableau, and Slack. Reach out to us to discuss your data strategy needs and find out how we can enable data agility for you and your team.

## About Atrium

Atrium is a consulting services partner that helps advise, implement, optimize, and manage analytics, AI, and machine learning solutions. We combine industry expertise, leading cloud platforms, and data science to help our customers unlock the power of their data and solve their most important challenges.

[Learn more at atrium.ai](https://atrium.ai)



 [info@atrium.ai](mailto:info@atrium.ai)