



# Smaller is Better:

Atomize Messages for more Flexible,  
Resilient Data Mapping and Modeling

---

Jared Stiff, CTO, Co-Founder

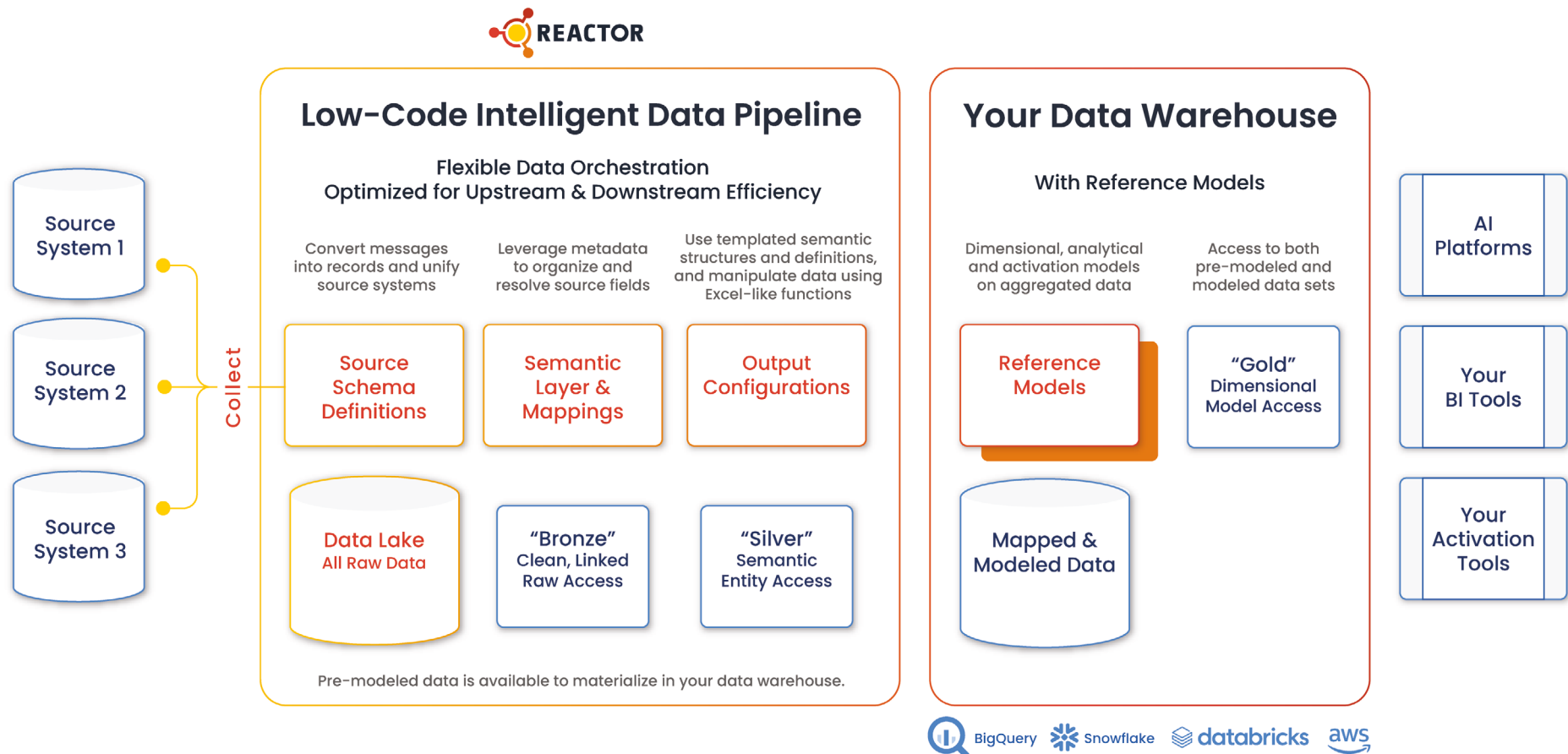
Rachel Workman, VP Value Engineering

Eric Best, CEO, Co-Founder

## Introduction: The cloud changes everything

Cloud data warehouses can transform the way you run your business, revealing the drivers and detractors of profitable growth. But cloud data warehouses can also become expensive dumping grounds for unusable data.

A useful and cost effective data infrastructure requires more than just a data warehouse filled with raw data, dependent upon brute-force data engineering to map and model data into useful business output.

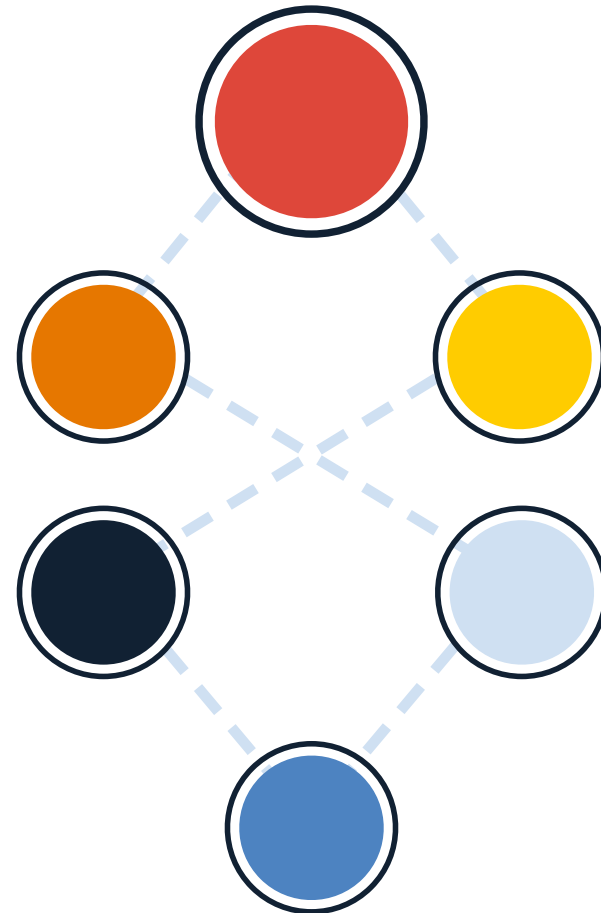


In the **Marvel Cinematic Universe**, mythical Infinity Stones grant their owners great powers. One stone in particular, the Reality Stone grants its holder the power to manipulate matter and defy the laws of physics.

Back in the real world, **data engineers and analysts and the data consumers they serve face the reality of complex data whizzing from system to system**, and complex data modeling work taking place within modern cloud data warehouses like Snowflake and Google BigQuery.

In reality, point-to-point data integration tools lift-and-shift data bundled into messages from one system to another. But extracted, transformed, and loaded (ETL or ELT) data usually lands in disconnected, unjoined tables. It requires a massive lift from engineering to get to entities and models that are useful and business-ready for analysts and end users. Even with the modern cloud and modern data stacks, we still spend massive amounts of time giving data the right business context so it's useful for consumers and customers of the data.

Wouldn't it be great if data teams had access to a Reality Stone to grant them data manipulation superpowers? There is a **real-world corollary for matter manipulation in the data world**. It's an emerging concept, and it's called "[data atomization](#)."



## Data Atomization in the Real World

Data atomization is the process of breaking down a larger dataset into smaller, more granular pieces or "atoms" of data. Each atom typically contains only a single data point or fact, along with relevant contextual information, such as when and where the data was collected.

Atomic data stores and atomic data warehouses are data management structures designed to store and manage granular, atomic-level data. Atomic data stores are typically used to store individual data points. In contrast, atomic data warehouses are used to store and analyze relationships across and between large collections of atomic data points.

**These atomic structures are important in building analytical data pipelines because they enable efficient data processing and analysis.** By storing data at the atomic level, analysts can easily filter, sort, and aggregate the data in various ways, depending on their specific research questions and needs. This allows for more targeted and precise analysis and can help uncover insights and trends that might not be apparent when analyzing the data in its original form, ingested from system silos.

For example, ingested messages from Shopify, NetSuite, and ShipStation contain customer billing and shipping address data attached to order origination, order management, and shipping system messages respectively. All of these systems generate address data, but to unify, rationalize and harmonize that address data without atomization, a data engineering team would have to manually parse and isolate the address data – dealing with all of the context and complexity of each system's message format and schema. With atomization, however, data contained in each message is stripped from the message itself

(its data "container"), and the resulting, atomized customer address data is much easier to access, label, map and use. With atomization, the customer address information from ANY system can be rendered using common definitions and mappings, regardless of which source system generated that data.

**Atomization and associated data abstraction have the potential to create HUGE efficiency gains for data teams and consumers, as data can be organized and defined independent of the scope and purpose of incoming (or outgoing) messages.** Teams have a flexible data structure to use within the data warehouse because the source system inputs aren't wrapped up in fixed-schema messages or tables, they've been broken down into the smallest possible components to maximize their flexibility. It's faster for teams to gain insights on atomized data because the data has been cataloged, often in real-time and in a way that was designed to enable retail business use cases.

Atomization also offers improved data pipeline resiliency because source system changes are handled before they can impact (i.e., break) downstream analytical and behavioral models in the data warehouse.

Another benefit of atomization is the lower operational processing loads – meaning your Snowflake or BigQuery data warehouse bill! – as a result of breaking down and standardizing the data earlier in the process.

**With atomization, data teams can stop worrying about the peculiarities of each source system and instead spend their time improving common models that actually deliver value for business customers.**

## Data Atomization in Retail and Direct-to-Consumer (DTC) Commerce

In the retail industry, atomic data stores and warehouses can be used to manage a wide range of data, including sales, customer, inventory, and marketing data. Here are five examples of how atomic data structures can be applied in the retail industry:

### Sales Data:

Retailers can use an atomic data store to manage individual sales transactions. This data can be used to analyze sales trends by product, store location, or customer segment. An atomic data warehouse can be used to store a larger collection of sales data, enabling retailers to analyze sales performance over time and identify patterns and trends.

### Customer Data:

Retailers can use an atomic data store to manage individual customer interactions, such as purchases, returns, and website visits. This data can be used to analyze customer behavior and preferences and to develop targeted marketing campaigns. An atomic data warehouse can be used to store a larger collection of customer data, enabling retailers to track customer behavior over time and identify trends in customer loyalty and engagement.

### Marketing Data:

Retailers can use an atomic data store to manage individual marketing interactions, such as email opens, website clicks, and social media engagements. This data can be used to analyze marketing performance by channel, audience, and campaign. An atomic data warehouse can be used to store a larger collection of marketing data, enabling retailers to analyze marketing trends over time and optimize their marketing strategies.

### Inventory Data:

Retailers can use an atomic data store to manage individual inventory transactions, such as stock levels and replenishment orders. This data can be used to analyze inventory performance by product, store location, or supplier. An atomic data warehouse can be used to store a larger collection of inventory data, enabling retailers to analyze inventory trends over time and optimize their supply chain operations.

### Operational Data:

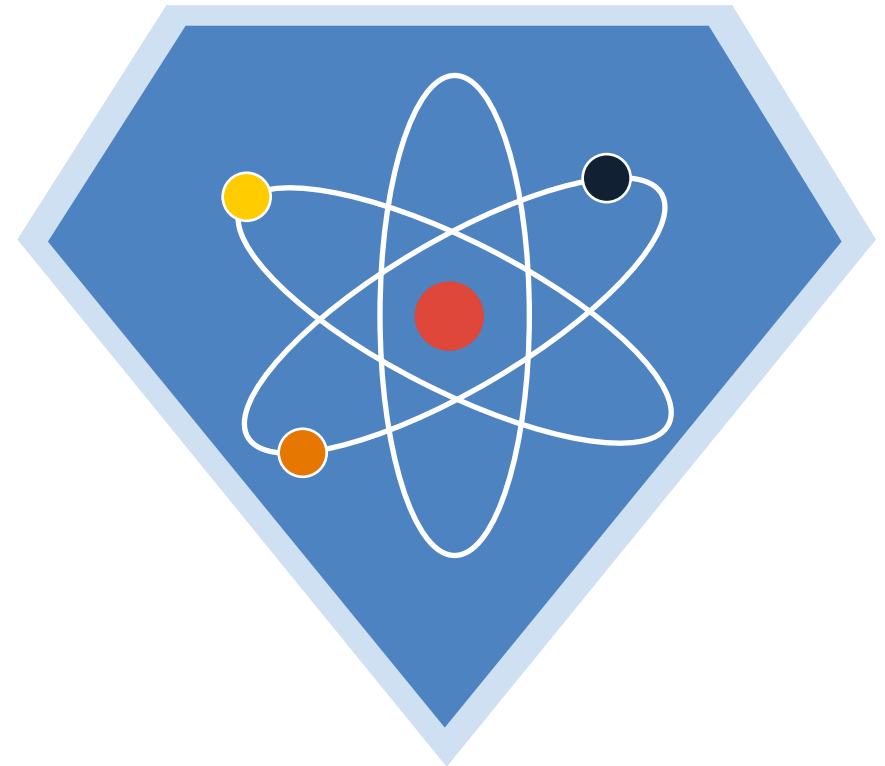
Retailers can use an atomic data store to manage individual operational transactions, such as employee schedules, payroll, and vendor invoices. This data can be used to analyze operational performance by store location, department, or vendor. An atomic data warehouse can be used to store a larger collection of operational data, enabling retailers to analyze operational trends over time and optimize their business operations.

## Atomize for Flexibility

With data atomization, retail data practitioners can separate, label, define and use data without worrying about the peculiarities and complexities of system end-points and message structures.

**Atomize your data before your data warehouse analytical modeling to build more efficient, observable, flexible, and useful models for your retail business decision-makers.**

Even in the real world, and without the need for the Reality Stone, you can manipulate data like an MCU superhero!





Put Your Data to Work.

## Future Proof your Data Stack with Atomization

Atomize your data before your data warehouse analytical modeling to build more efficient, observable, flexible, and useful models for your business decision-makers. Your data practitioners can separate, label, define and use data without worrying about the peculiarities and complexities of system end-points and message structures.

Find out more about all **nine characteristics of a Future-Proof Cloud Data infrastructure** in our comprehensive [exclusive ebook](#).

Contact Reactor to learn more and get started today!



Get the Full E-Book



**Reactor** provides the fastest, most efficient path to useful, business-ready data for generative AI, analytics and activation. Built for retailers of any size or complexity, Reactor transforms your unique data infrastructure into an easy-to-use, no-code environment that's accessible to everyone — no engineering degree required. Reactor onboards and ingests data from business critical systems and applications, landing clean, well-defined data modeled directly in your data warehouse.

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

1-888-417-6863

[Grow@reactordata.com](mailto:Grow@reactordata.com)