ATSCALE

Building an Enterprise Metrics Store

The role of a semantic layer in building an enterprise metrics store



Table of Contents

3	The Vision of an Enterprise Metrics Store
4	Decomposing the Value of Atscale
5	Making Data Analysis-Ready
6	How to Design Metrics
8	Protecting Data Pipelines
9	Managing Performance
10	The Role of Governance
11	Facilitating an Open Consumption Model
12	Building an Enterprise Metrics Store with AtScale
13	Take the Next Step



The Vision of an Enterprise **Metrics Store**

This eBook is the third in a series focused on implementing a semantic layer platform for modern cloud analytics. In our previous eBook, we discussed a modernizing approach to dimensional modeling for speed of thought analytics and bridging the gap between business and data science teams. Now we'll dive into delivering consistent and predictable analytics across the organization.

An enterprise metrics store - that is, a governed source of business metrics and analysis dimensions - is essential for accelerating the rate and guality of insight creation. Enterprises rely on metrics to drive business decisions, yet these are often defined differently across spreadsheets, dashboards, data pipelines, teams, and departments. This hinders the accuracy, timeliness, and confidence in insights generated from data.

In our previous eBook about speed of thought analytics and modernizing online analytical processing (OLAP) solutions, we discussed the challenges of managing multiple data sets and definitions that are stored within local OLAP cubes. Business teams often draw different conclusions from these disparate copies of data sets because they're siloed from the original data source. This can lead to misunderstandings and incorrect analyses.

By building a single source of truth for business metrics, organizations can break down these siloes, enabling more people within the organization to derive value from data. A semantic layer is foundational in deploying an enterprise metrics store strategy. Properly implemented, a semantic layer enables decentralized creation of new data products, while maintaining centralized governance and consistency.

The self-service analytics approach that an enterprise metrics store facilitates, in turn, can drive a greater ROI from data. That's because the overall value of data-driven business insights is directly related to the number of insights the organization can generate. More specifically, expanding the amount of business users that can access and analyze data will increase the frequency and velocity of insights.



www.atscale.com Building an Enterprise Metrics Store

Decomposing the Value of Data-Driven Insights in an Organization

Value of Insights(t)=(Avg Benefit/Insight - Avg Cost/Insight) * Number of Insights(t)

The overall business value generated by new datadriven insights over time.

Increase benefits by:

- Accessing broader range of enterprise data
- Incorporating 3rd party data
- Applying AI/ML to insight models
- Making insights more consumable with full range of visuallization and consumption
- Create and distribute • insights faster

Increase efficiency by:

- Reducing time to create •
- Reducing level of skill •
- Reducing data management • and compute costs
- Reusability of work ٠
- Reducing data leakage •

Increase the frequency and velocity:

- Giving more time to your • insight community
- Expanding your insight community



www.atscale.com Building an Enterprise Metrics Store

Making Data Analysis-Ready

Raw data presents a number of challenges for business teams. Since collecting and storing information is easier than ever, there are a lot of information silos throughout many organizations. These siloes could include:

- physical siloes (different data platforms)
- schema siloes (different data formats or structures)
- semantic silos (different data descriptions and definitions)

An important aspect of improving the ROI of data is simplifying the process of insight creation from these disparate data sources. When presented with analysisready data, or data that's optimized for analytics and reporting, data consumers can more easily generate business insights using the business intelligence tools they're familiar with.

In addition, analysis-ready data can incorporate blends of multiple raw data sources. For example, enterprises can combine data from different departments, such as marketing and sales data, to generate more valuable business insights. To make blended data sets analysis-ready, however, they need to be aligned with common dimensions like time, geography, or product. Blended data sets could also incorporate data from third parties, including data exchanges like AWS Data Exchange. This third-party data — like weather data from weather.com, foot traffic data from Foursquare, or stock prices from Bloomberg — can enrich an organization's internal data, leading to more valuable business insights. The key is being able to easily update data from disparate sources to make it analysis-ready. An enterprise metrics store with a dynamic foundation can make this possible.



How to Design Metrics

Metrics design includes defining quantitative measures and the analysis dimension that properly describe business activity. An enterprise metrics store presents a managed set of metrics and dimensions to data consumers. Designing metrics can be simple (choosing the right way to define revenue or product quantity) or complex (defining a gross margin calculation across product lines). Metrics can also be designed for timerelative measures, such as parallel period and period-to-date growth. The best way to define a particular metric is best left to subject matter experts.

Since there are varying complexities for metrics, there are also different techniques and tools that data scientists or business analysts can use to design them. Codebased approaches to metric and feature design offer greater flexibility, while lowcode and no-code approaches using visual modeling simplify the process when the need is less complex.

During the metric design process, IT teams are important subject matter experts, but ultimately, this process should be accessible to users of varying technical ability. By simplifying the process for publishing new metrics, this creates a pipeline for delivering analysis-ready data to end users. This also supports data lineage analysis, enhancing the visibility organizations have into how changes to the underlying raw data or data structures will impact metrics or features. With an enterprise metrics store, organizations can more easily design and publish new metrics to onboard data faster.



How to Design Metrics





7

Protecting **Data Pipelines**



Once metrics are designed and published to enterprise metrics stores, it is crucial to manage data pipelines and guard from disrupting downsteam data consumers. Data pipelines may be physical or virtual.

Physical pipelines are ETL (extract, transform, load) scripts that data engineers use to take raw data from one or more data sources and turn it into analysis-ready data for end users. This data is refreshed on-demand or on a regular basis, but often reguires data engineering resources to complete.

Virtual pipelines are managed by virtual transformations, which is how a semantic layer like AtScale can transform raw data into analysis-ready data without any data movement or wrangling. The virtualization engine translates data consumer queries into queries appropriate for the raw data source, applying the transformations at runtime without the need to extract and load the data.

Both approaches have their place, but physical pipelines typically require data engineers to build and maintain. Virtual pipelines, like those defined by AtScale, are automatically created with each new metric that's published. In the upcoming sections, we'll dive deeper into some of the advantages of virtual pipelines for building an enterprise metrics store.



www.atscale.com Building an Enterprise Metrics Store

Managing Performance

Ensuring user adoption of enterprise metric stores relies on consistently high performance.

If queries of metric stores are slow or inconsistent, data consumers will resort to other solutions, such as local data extracts. This will inevitably lead to metrics sprawl, shadow IT, and reporting discrepancies, further hindering performance and consistency in the long run.

Physical transformations using structures like OLAP cubes can enhance performance, but they require the support of data engineers to implement and maintain. However to these new data structures become disconnected from the original data source.

Virtual approaches to data transformations avoid many of the challenges of physical transformations, but they still need to use optimization techniques to achieve consistently high performance. Since the transformations occur in real time, slow performance can become immediately apparent to end-users during each query. AtScale implements a range of techniques to enhance query performance using aggregates, which we've discussed in our previous eBook about speed of thought analytics.

BREAKING DOWN DATA SILOS WITH TYSON FOODS

Tyson Foods – the global food giant – wanted to deliver selfservice analytics to its 144,000 employees, but its data was fragmented across a diverse set of platforms. The company needed a unified analytics strategy, enabling analysts to easily access data and blend data sets for smarter decision making.

By building a centralized enterprise metrics store using AtScale's semantic layer, Tyson was able to unify its disparate data sources into a governed data model that was analysis-ready. Using predefined calculations, metrics, and dimensional hierarchies, the organization was able to cut down on redundant work and maintain centralized data. Decoupling the model that data consumers work with from the underlying raw data also allowed Tyson to migrate to the cloud without disrupting end users. This shift to more modern infrastructure dramatically improved analytics performance. In fact, Tyson now runs over 21,000 queries each day in less than ten seconds each, and it takes less than fifteen minutes for raw data to become analytics ready.

Providing business analysts with trusted building blocks of data using AtScale also dramatically reduced errors and conflicting analysis. This also formed the cornerstone of self-service analytics at the company, leading to a more data-driven approach to decision-making.



9

The Role of Governance

When it comes to managing the way data is accessed and analyzed, there are two key considerations: data governance and analytics governance. Data governance is focused on who can access what data using security and access controls, while analytics governance focuses on how that data gets used. Analytics governance includes the proper definition of metrics and hierarchical relationships for the fields used with dimensions (like geography, time, product, family). This helps data consumers better understand the data and ensures consistency across consumption tools.

With analytics governance, it's possible to radically simplify data definitions, reducing the chance for misinterpreting key business metrics or analytics dimensions. These simple data definitions can form an enterprise metrics store, empowering individuals to interact with the data independently and with the confidence that they're using it in the right way. Enforcing the access and use of data is crucial for maintaining security and reducing risk, while also facilitating broader data usage throughout the organization. By enforcing these analytics governance policies at the time of analysis, organizations can drive consistency and predictability across a wide range of consumption tools. This standardizes the definition of data assets and ensures data consumers are presented with simple, business-oriented information.



Facilitating an Open Consumption Model

Analysts and decision makers need to consume metrics from a range of tools, depending on the use case.

For example, BI tools like Tableau or Power BI are optimized for dashboarding and interactive analysis whereas Excel is often used for financial modeling and ad hoc analysis.

It's common for organizations to use many different tools for data analytics, yet each tool has its own data structures, dialect, and more.

Headless BI is a new concept aligned to the idea of a metrics store, where data consumers can access the same metrics from different BI tools or "heads." The headless approach decouples the data layer from the consumption layer, meaning every data consumption tool shares the same semantic model.

Maintaining an enterprise metrics store enables organizations to standardize analytics using a single, governed semantic layer. This open consumption model eliminates local semantic layers defined in different BI "heads" as well as local copies of data sets. Ultimately, this gives organizations the agility to adopt new tools and data platforms as technologies evolve without losing clarity or meaning of data.







Building an Enterprise Metrics Store with AtScale

AtScale is a semantic layer platform that can streamline the process of building an enterprise metrics store. Here are some of the essential capabilities AtScale provides:



SEMANTIC MODELING

With a point-and-click interface or advanced scripting, AtScale enables a broader set of users to create semantic models and publish a centralized set of data definitions.



DATA PREP VIRTUALIZATION

Once a new model is published, users can immediately interact with data through the AtScale model. AtScale automatically generates SQL that runs natively on the underlying cloud platform.



MULTI-DIMENSIONAL CALCULATION ENGINE

AtScale's multidimensional calculation engine can deliver high speed access to calculated metrics at any level of aggregation which highly simplifies query design.



PERFORMANCE OPTIMIZATION

AtScale automatically aggregates data definitions and orchestates data structure creation, optimizing query performance.



ANALYTICS

GOVERNANCE

By applying real-time

semantic governance

during gueries, AtScale

drives consistency and

integrates with access

management systems

to enforce security and

predictability across

the organization.

AtScale also

reduce risk.



DATA DISCOVERY

Integrations with data catalogues supports data discovery, data lineage analysis, and other capabilities for protecting data pipelines. This ensures data consumers know where to access and leverage relevant analysis ready data.

SEMANTIC





ACCESSIBILITY

AtScale leverages live connections between data platforms and the consumption layer, enabling business analysts and data scientists to access the data they need from the tool of their choice.

Take the Next Step

About AtScale

AtScale enables smarter decision-making by accelerating the flow of data-driven insights. The company's semantic layer platform simplifies, accelerates, and extends business intelligence and data science capabilities for enterprise customers across all industries. With AtScale, customers are empowered to democratize data, implement self-service BI and build a more agile analytics infrastructure for better, more impactful decision making. For more information, please visit www.atscale.com and follow us on LinkedIn, Twitter or Facebook.

REQUEST A DEMO →

CONTACT US →

Read The Practical Guide to Using a Semantic Layer LEARN MORE \rightarrow

Download The Complete Buyers Guide to a Semantic Layer

Review How does using a semantic layer impact cloud data warehouse performance? (benchmark reports)

Get advice from fellow data & analytics leaders on how to scale smarter data-driven decision-making

Stay current on analytics strategies with short articles and webinars on topics in analytics strategies

