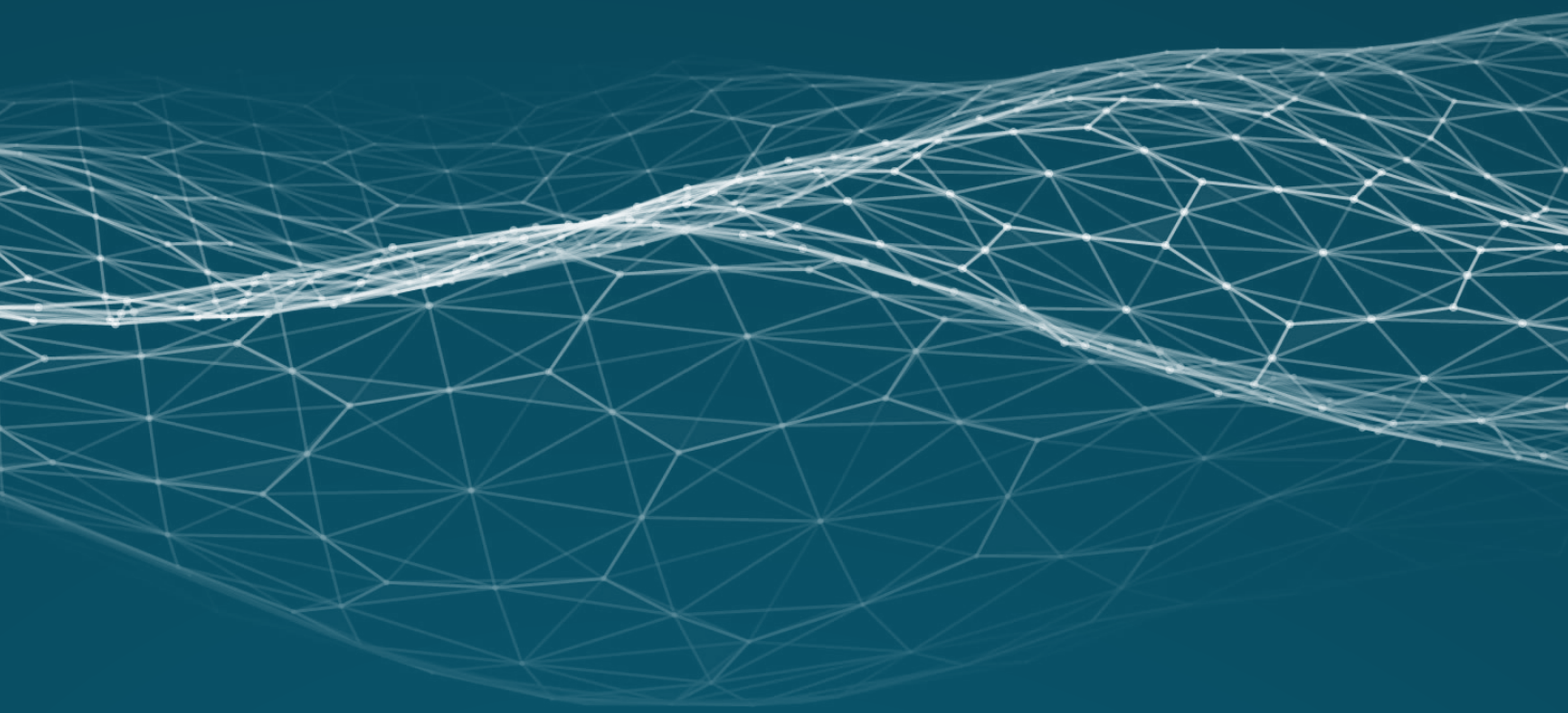
STARDOG



WHITEPAPER

Data Fabric

The next generation of data management

Build a data fabric to power collaborative, cross-functional projects and products. Escape reactive workflows with a resilient digital foundation, no rip-and-replace required.

Table Of Contents

Data Fabric: The Next Generation Of Data Management

The mandate for enterprise IT to deliver business value has never been stronger. 76% of executives believe **IT must be an active partner in developing business strategy**. To succeed in this daunting task, agility is key. However, enterprises are hampered by data strategies that leave teams flat-footed when the market shifts or new questions arise.

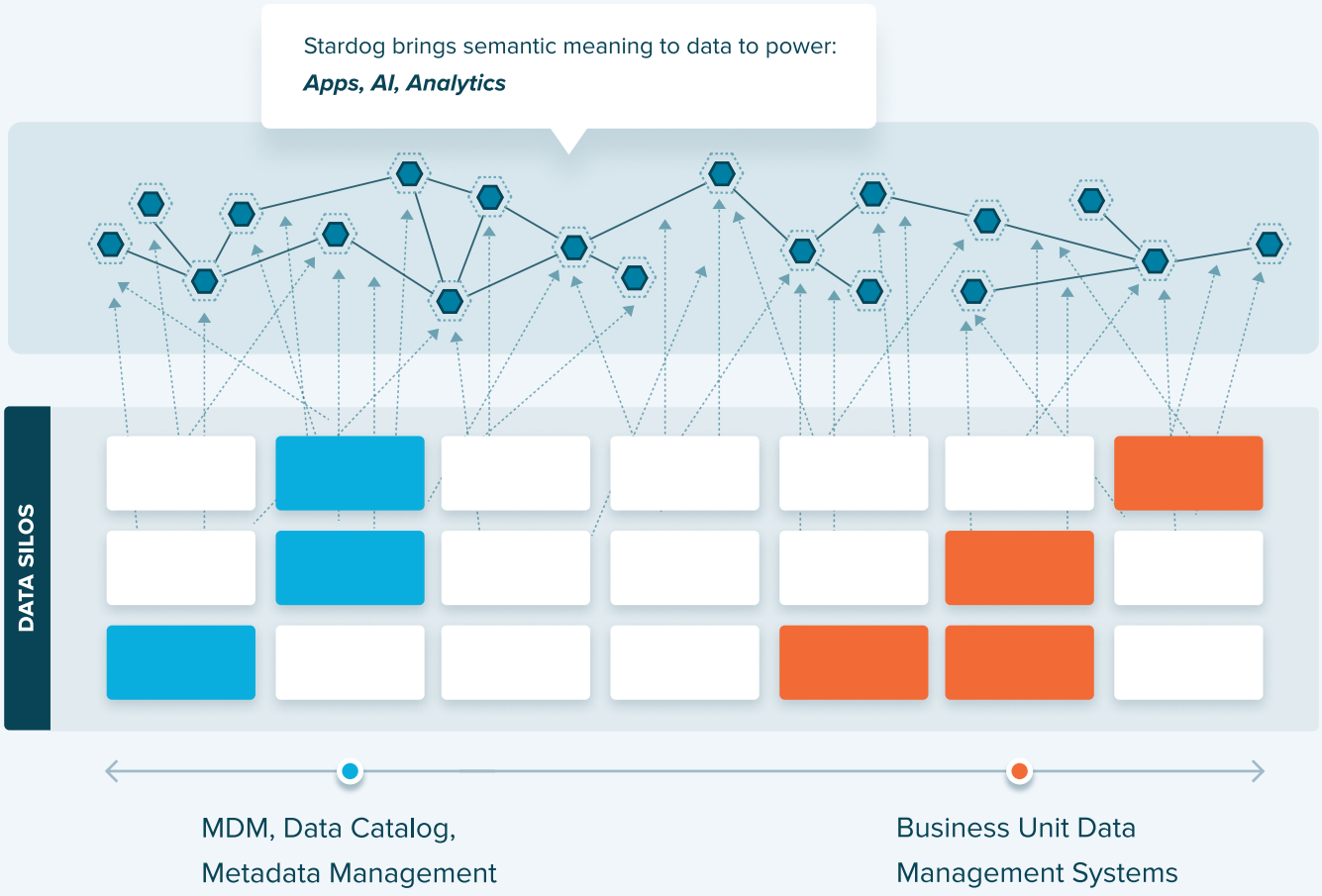
Structured data management systems worked acceptably well when the enterprise data landscape was itself predominantly structured. But the world is different now. The enterprise data landscape is increasingly hybrid, varied, and changing. The emergence of IoT, rise in unstructured data volume, increasing relevance of external data sources, and trend towards hybrid multi-cloud environments are obstacles to satisfying each new data request.

The old data strategy centered around relational data systems is fundamentally broken. How can enterprises shift from a reactive to a responsive data strategy?

Enterprise data fabrics offer the new way forward. The data fabric weaves together data from internal silos and external sources and creates a network of information to power your business' applications, AI, and analytics. Quite simply, they support the full breadth of today's complex, connected enterprise.

Decrease time to insight by up to 90%

> **"Stardog enables you to browse through the data and all these relationships. It's a 10 to 1 savings. It's not only less overhead, it's much better job satisfaction and getting the knowledge in hand that you lacked before."**
>
> Program Data Integration Manager,
> Exploration Systems Division NASA



Key Principles Of Data Fabrics

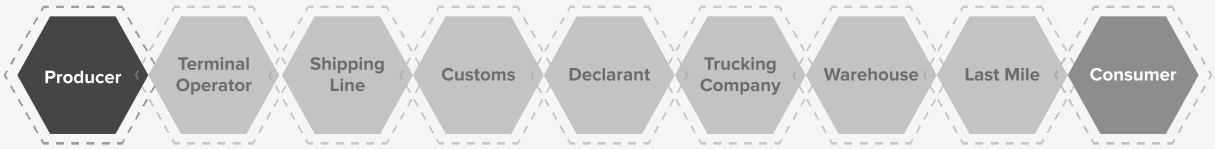
1. Data fabrics can answer unanticipated questions and adapt to new requirements.
2. Data fabrics bring meaning to data which leads to insight.
3. Data fabrics enable query across data silos and external sources, regardless of data structure.
4. Data fabrics modernize existing systems; no rip-and-replace required.
5. Data fabrics connect data at the compute layer, not the storage layer. This connects silos without creating another silo.

What New Transformations Can Data Fabrics Unlock?

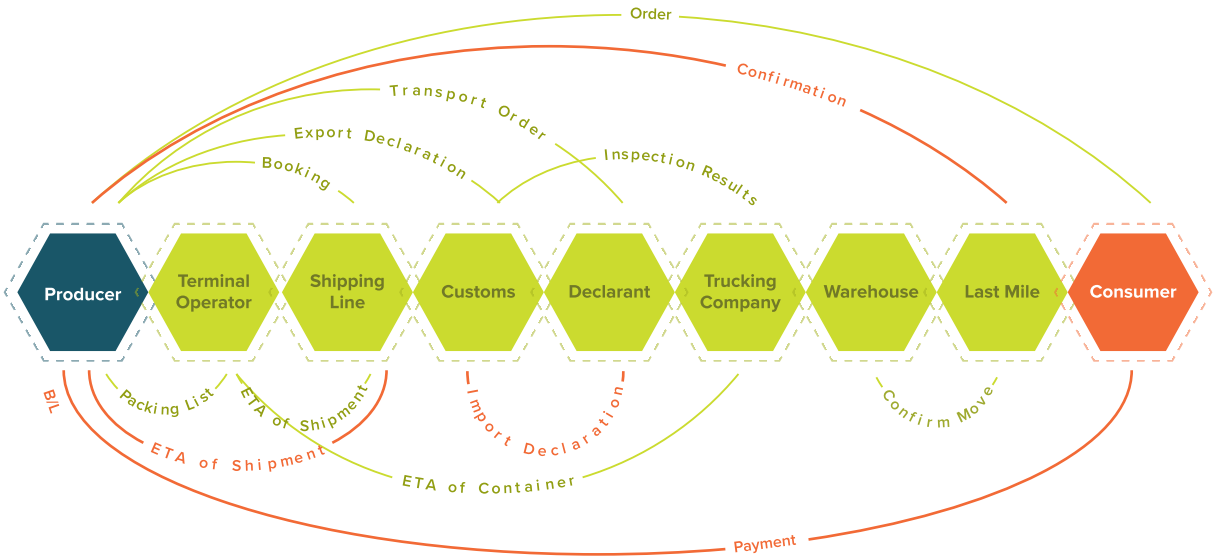
Data fabrics support cross-functional data connections that are key to creating and defending competitive advantage. Today, it is critical to go beyond business-line transformation and enable collaboration across the enterprise as well as with external partners.

Take supply chain for example. Traditional supply chain data systems are a relay race, operating with linear handoffs and siloed, peer-to-peer links between systems. When COVID-19 hit, supply chains globally collapsed. Some strain or even partial collapse was inevitable; but it was made much worse by bad data strategy that treated supply chain as a rigid system when, in reality, it's a complex network of actors who had to be fully in sync to adjust as needed.

THE OLD: ***A Structured Supply Chain***



THE NEW: ***A Digital Supply Network***



With a digital supply network powered by a data fabric, now enterprises can answer complex questions they were previously blind to. “Show me all the lots of raw materials and associated suppliers involved in the production of finished good lot 123.” Or, “How do COGS for product A compare between these two regions?” Or, “Which manufacturers supplied the raw ingredients involved in this customer complaint?”

But how exactly do data fabrics succeed where other approaches have failed?

1. First, data fabrics change the status quo by delivering meaning, not just data, across the enterprise. This meaning is woven together from many sources: data and metadata, internal and external sources, and cloud and on-premise systems. Meaning is captured within the data model, with all context on each data asset fully present and available, in machine-understandable form. With a data fabric, people and algorithms can make better decisions while also reducing the likelihood and risk of data misuse or misinterpretation.

2. Second, a data fabric delivers answers via powerful querying capabilities. A data fabric is not a static thing; rather, it’s a queryable data layer, allowing users to answer questions from across data silos. In a data fabric, query happens at the compute layer above the actual storage layer. It’s at this compute layer where the data fabric connects otherwise disconnected silos and systems. Data flows from source to app and back again, constantly enriching and improving upon the data fabric.

A data fabric is not static; rather, it’s a responsive, queryable data layer, allowing users to answer complex queries from across data silos.

3. Third, data fabrics weave together existing data management systems, enriching all connected apps. They are the next step forward in the maturation of the data management space. Data lakes once held the promise of centralizing an enterprise's assets, but failed to make the data usable. Data lakes fail precisely because they tried to connect data at the storage layer, not at the compute layer, based on data location rather than based on data meaning. Physical colocation of information does not by itself accomplish data connection or provide meaning. An older generation of storage-based integration systems, the data warehouse, is in fact even less capable than data lakes since they only admit structured data to begin with, leaving the semistructured and unstructured data silos completely disconnected. Lately companies have turned to data catalogs to try to address the bewildering diversity of their data landscapes. However, cataloging alone doesn't lead to connected enterprises.

These previous solutions failed in part due to hybrid, varied, and changing data, but also due to organizational pushback. Data fabrics, however, are built for collaboration. By leveraging and connecting these existing assets, data fabrics are driving a new breed of cross-functional data management projects.

Modernize Existing Investments

While previous technologies such as data lakes, data catalogs, and data integration platforms have promised to end data silos, the truth is, data silos are inevitable! They exist for very good reasons. They allow for local control and governance when it is important to a particular part of your business. Some data must be stored apart from other data to comply with legal regulation or simply for legacy business reasons. Or data is just too essential to business operations to bear the risk of consolidating, eliminating, or modernizing it.

1. Required for local control & governance

2. Mandated by regulation

3. Optimized for a business unit

Data silos are the result of enterprise data that is:

- **Diverse**, and it's only becoming more diverse as unstructured data growth rates skyrocket.
- **Distributed across multiple systems in different places,** particularly as hybrid and multi-cloud computing become necessary.
- **Controlled by division business leaders** who may have competing interests.
- **Enabled by vendors** who want to lock you in to their solution.

Whereas previous data management solutions have focused on eliminating silos through mastering, migration, consolidation, or governance; data fabrics offer a practical alternative to fighting data silos. Rather than working against data silos, a data fabric leverages these data silos without requiring further copies of data.

Instead of replacing legacy technologies, a data fabric works alongside existing investments and improves their utility. This is because a data fabric is not a single solution, it is actually an architecture design that operates at the compute layer and focuses on connecting data wherever it resides. and, thus, actually improving upon existing data storage assets like data lakes, data catalogs, warehouses, and other data integration platforms like MDM.

We can start to see now how “data fabric” actually works as a description of what’s really going on: just like an ordinary fabric, which conforms to whatever it lays over, an enterprise data fabric lays over existing data assets and connects to them via individual threads, and weaves these sources together into a unified layer. By doing so, data fabrics actually compound the business value of existing investments.

The key ingredient to this transformation? **A knowledge graph.**

> **"In a data fabric approach, one of the most important components is the development of a dynamic, composable and highly emergent *knowledge graph* that reflects everything that happens to your data. This core concept in the data fabric enables the other capabilities that allow for dynamic integration and data use case orchestration."**
>
> Gartner "How to Activate Metadata to Enable a Composable Data Fabric," Mark Beyer, Ehtisham Zaidi, 16 July 2020

Knowledge graphs are able to represent everything that happens to enterprise data because they serve as a universal format for data, regardless of its source structure or location or format. A knowledge graph replaces the current laborious process for integrating enterprise data, which typically involves extraction, translation, modeling, and mapping between various applications. The custom code required for modeling and mapping quickly becomes unwieldy at large scale, slowing the pace of innovation and insight.

In contrast, a knowledge graph creates a reusable network of knowledge to power your business. It easily represents data of various structures and supports multiple schemas. Furthermore, it creates the semantic understanding of enterprise and third-party data that provides critical access to business insight. This serves as the core of the data fabric, enriching and accelerating existing investments.

Stardog's Enterprise Knowledge Graph platform is uniquely able to deliver a data fabric architecture without requiring rip-and-replace or building yet another data silo. After implementing Stardog, enterprises achieve 50-90% improvement in time to insight, dramatically cutting down on previous data preparation timelines.

IN PRACTICE

A leading global pharmaceutical company evaluated several different tech stacks before realizing they needed a more proactive data strategy to support the company's R&D goals. While they had a data lake in place, their data scientists were still mired in searching for data needed for critical drug discovery analysis.

30% of active ingredients under evaluation were sourced from external collaborations, and they have limited control over the quality of this data. They needed a flexible solution that could relate their internal experimental results to external and publicly available studies. They also needed to be able to evaluate the many-to-many relationships within their R&D data, such as, "Find a set of compounds which are creating a similar effect," or "Find compounds which have been tested in similar conditions and similar treatments."

By implementing Stardog atop their data lake, they created a company-wide data fabric that provides a consolidated, one-stop shop for 90% of their R&D data. Their data fabric brings data access directly to data scientists and accelerates drug target identification and drug repurposing efforts, helping deliver innovative new drugs to market faster.

> **"For us it was a natural choice to deviate from the pure data lake technologies to a more sophisticated model."**
>
> Head of IT Research Computational Biology and Translational Science

Learn How Stardog Works

In this section we answer some common questions we hear about data fabrics:

- How is graph different from relational data? Why do I need to change?
- How is semantic meaning generated?
- How exactly is data “enriched” and how does this impact analytics outcomes?

In the next section, Connecting the Enterprise, we’ll cover practical requirements of implementation, including building a team, socializing your data fabric, and developing a data model. Skip ahead to **page 21** to get straight to work!

SEMANTIC GRAPH

The future of data management will be based on semantic graph

Semantic graph is the beating heart of the data fabric, responsible for creating meaning from data silos. However, this isn’t its only contribution. Semantic graph uniquely supports your ability to:

1. Connect all the data that matters
2. Answer unanticipated questions quickly
3. Support multiple use cases with the same data

WHAT IT DOES:

Semantic graphs create meaning by mapping entities, their metadata, and their relationships in an evolving information network. Semantic graph, also called RDF graph, is the only way to represent data that is natively stored in other structures while maintaining all relevant metadata and context.

Connect all the data that matters

In order to create business value within the enterprise, you must be able to connect all the data that matters. Some of this data will be stored in tables, but also in PDFs, webpages, emails, and other semistructured and unstructured sources. Only semantic graph is able to represent data that is natively stored in other structures and connect all relevant metadata and context.

With Stardog, different data dialects and structures embedded in legacy systems can be represented in the standard language of RDF. This allows for queries across relational databases, NoSQL databases, documents, and even geospatial data—seamlessly.

Regarding Resource Description Framework [RDF], Gartner provides "it is a simple data model with a standard syntax that can represent information of any form. The true power of the RDF, however, is its ability to explicitly and unambiguously capture meaning — or semantics — in the data itself. Information architects can use the RDF to create vocabularies that define and describe every element in a particular domain. These vocabularies are shared and accessible across systems, so data can leave home without fear of being misunderstood."

> Gartner, How to Use Semantics to Drive the Business Value of Your Data,"
> Guido De Simoni, 2 April 2020

The key to understanding how semantic graph integrates data is to know that it links or connects related data, rather than transforming it. Each data object is assigned a unique ID, to which all related information is linked. This unique process allows data owners to maintain control of source data while enabling enterprise-wide collaboration.

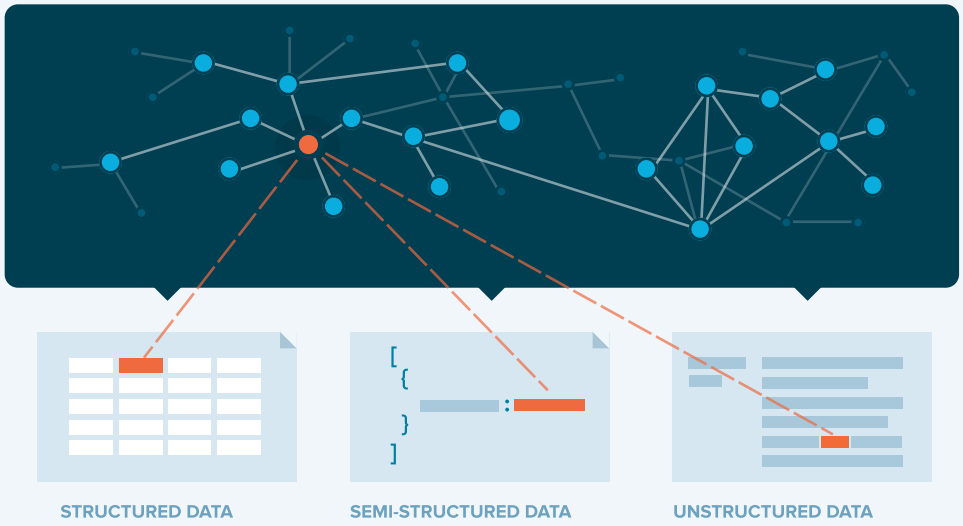


Figure. Data and metadata from varied sources is unified within Stardog, creating a network of knowledge to power your organization.

At this point, people typically start to worry about scale. But in fact, the largest information integration projects on the planet already use this model. Look no further than your web browser to see this in action. The Web contains a world of information, created by different contributors, and accessible through a single browser. Google Search is also powered by a Knowledge Graph, a network of **500 billion facts about five billion entities**. Both Google and the Web are proof for this model of large-scale, complex, and decentralized information integration.

IN PRACTICE

A well-known publisher of business news and financial data used Stardog to deliver personalized news insights at scale to their corporate clients. By unifying structured and unstructured data across internal and external sources, they delivered an innovative new product to market.

READ MORE →

Answer unanticipated questions quickly

What is it about this network of meaning that leads to data agility? It has everything to do with semantic graph's flexible data model.

> **"Formally transitioning from a relational model to that of linked data was a huge strategic benefit to the bank. We are now able to design and link domain models across organizations and silos."**
>
> Executive Director
> Top 5 US Bank

Semantic graph operates in stark contrast to relational data. Finding connections between different relational databases requires time-intensive data modeling and query operations. Each new question produces a new dataset with its own schema. That's not sustainable for the rate of new and unanticipated questions that the business wants to ask of its data. Today, data and analytics leaders need to be able to quickly support iterative question and answer cycles from the business and easily dig into new territory in their data.

Instead of rows and columns and tables and keys, semantic graph organizes information using nodes and edges to represent for entities and the relationships between those entities. This graph data model is fundamentally simpler than the relational model, yet it's also far more expressive and powerful, easier to modify, and endlessly extensible.

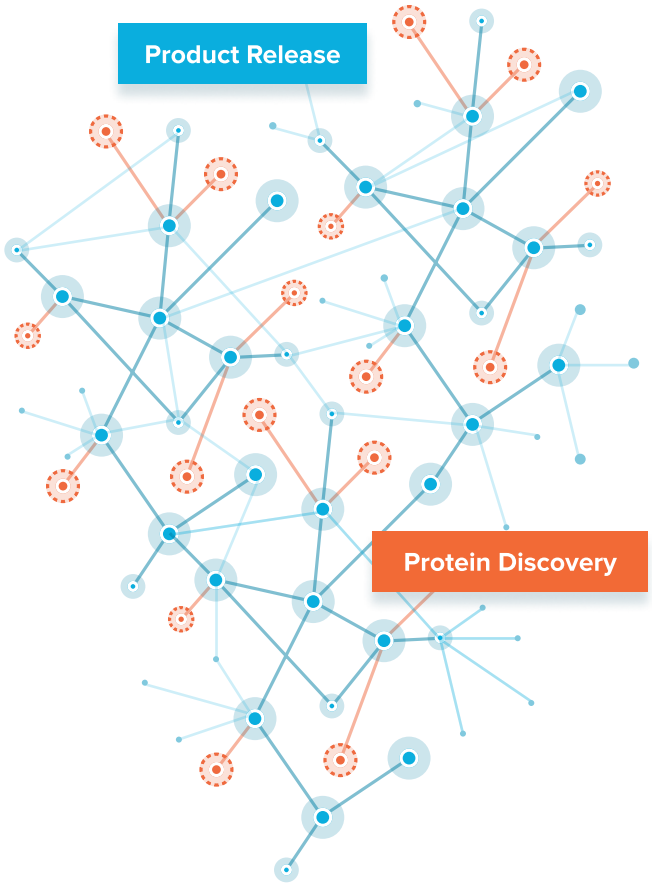
The model actually exists at the compute layer, not at the storage layer, which means you can modify the schema at any time by adding new nodes and edges, you don't have to struggle at a point in time to come up with a single shared data model covering all current and future enterprise data needs. It also means that the enterprise can have many different, even mutually incompatible schemas, that all apply discretely to the common pool of connected data. And that means you never have to force-fit emerging data sources and use cases to adhere to standardized rules from an already outdated perspective. The result? The same data can be reused for new questions, without starting from scratch.

Support multiple use cases with the same data

We just made a point worth diving more deeply into. What happens if you have many schemas, as is typically the case in enterprises? Can there be one schema to rule them all? In an ideal world, different use cases, organizations, lines of business, and applications would all see things in precisely the same way. Since this is not an ideal world, however, more often than not, that's just impossible.

Stardog fully supports data reusability so that different use cases, orgs, lines of business, and apps can share and reuse connected data without stepping on each other's toes or, just as crucially, without requiring a single schema to rule all the others. Stardog calls this capability schema multi-tenancy, and it supports customers deploying multiple use cases from the same data fabric.

Ultimately, this is the key to how Stardog customers achieve dramatic reduction in time to insight: they are able to leverage previous work and, just as importantly, avoid time-consuming, winner-takes-all fights about the one and only one schema for the business.



IN PRACTICE

A global bank with nearly $900 billion in total assets uses Stardog for both IT asset management and risk management. With the banks people, IT assets, and controls modeled within Stardog, the bank supports dynamic inquiries from various analyst groups. In particular, when a risk event occurs, operational risk analysts use Stardog to quickly assess the event against over 25,000 controls – measures instituted to prevent risk events – to identify what control should have prevented the risk and how to manage these risks in the future.

Historically, relevant data was stored across 15 separate applications, forcing analysts to run ad-hoc reports in Excel. This was not only time-consuming but also made it impossible for analysts to know if they had captured all data related to a particular incident. When analysts made decisions with incomplete information, they left the bank exposed to future risk events and possible financial loss.

The bank implemented a broad, reusable data fabric to identify relationships across the various applications involved in risk management, including incident reporting, control registries and IT asset management systems. Now, analysts can can traverse the linked information in Stardog to uncover dependencies within the data and identify root causes of particular incidents. Furthermore, they can proactively ask “what if?” questions to predict the impact of theoretical risk scenarios, creating a more proactive risk strategy that allows them to triage and mitigate potential risks.

LEARN MORE →

VIRTUALIZATION

Virtualization is a critical capability for scalable data fabric design

Data virtualization is a cost-effective data integration technique because it eliminates the expense of replicating, moving, and storing data multiple times. Virtualization connects source data directly, cutting down on what would be an otherwise complex and cumbersome ETL system, migrating data from dozens or even hundreds of systems and external vendors into a single repository. Copying data for each new analysis leads to human error and data drift. It leads to uncertainty about which data sources are trusted, current, or canonical. Data virtualization provides access to live source data and it means you're guaranteed to always get the most up to date data every time you ask a question.

While data virtualization has skyrocketed in popularity in recent years, every standalone data virtualization platform is based on a relational data model. These systems are only as powerful as the relational model itself, which means they cannot easily connect semistructured or unstructured data. They can only virtualize data that can be neatly fitted into tables, rows, and columns.

WHAT IT DOES:

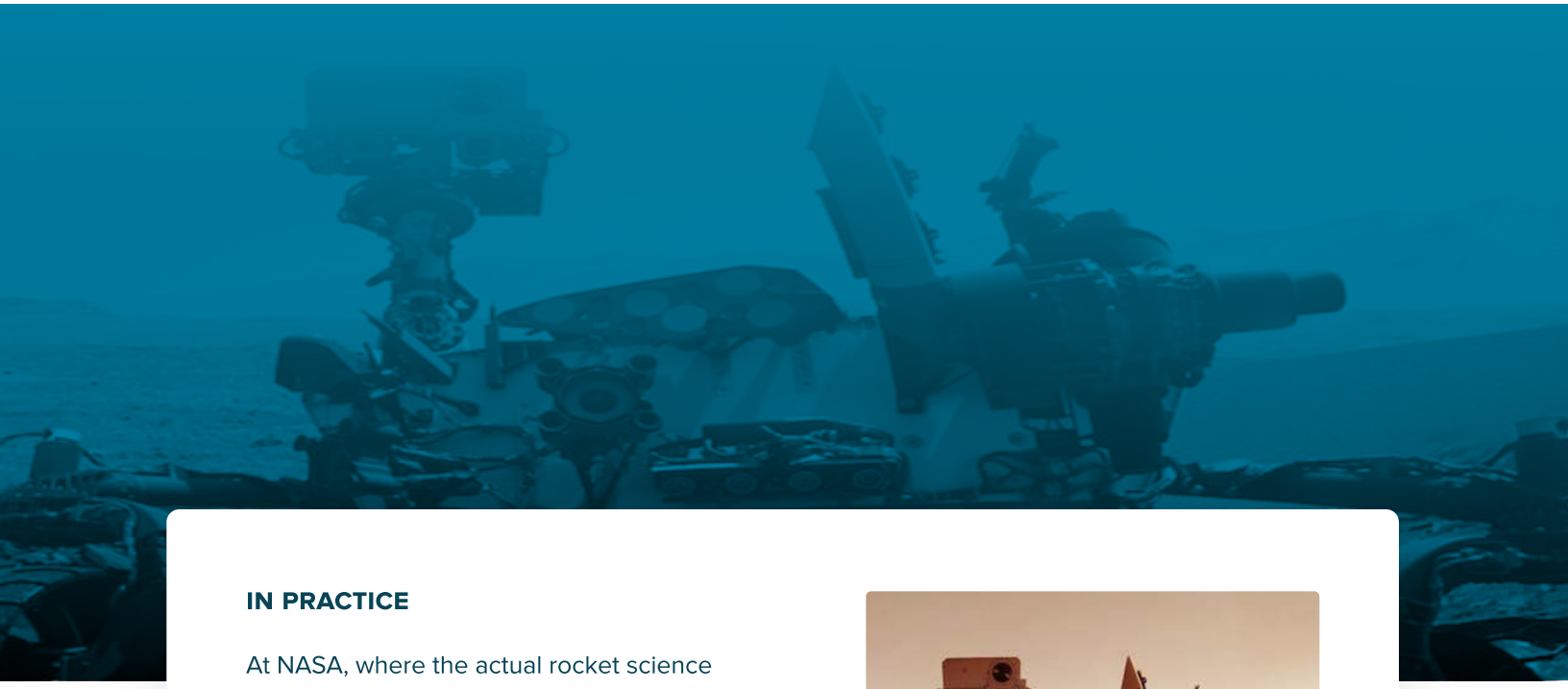
Data virtualization connects the enterprise without requiring moving or copying source data, saving time and money and reducing error from duplicated data. A data fabric based on Stardog exists at the compute rather than at the storage layer precisely because of Stardog's unique, patented data virtualization technology.

Because these data virtualization platforms don't have the power of semantic graph, they suffer from exactly the same rigidity as other relational systems. While they can protect data lakes from accidental edits, they cannot integrate data that is of diverse structures, is externally sourced, suffers from frequently changing schemas, has conflicting definitions, or has uneven properties.

Stardog's Virtual Graph capability is the most mature and powerful graph-based virtualization solution on the market. Virtual Graphs connect data across data silos, even without copying that data into Stardog. Further, they provide a direct access line for external data sources. Lastly, they offer a reliable scale-out mechanism. Stardog can also virtualize other Stardog instances as well as other graph systems, including SPARQL endpoints.

This gives the users ability to scale out their data fabric by using multiple Stardog installations, with each clustered instance storing up to 150 billion data points.

Since not all data can be virtualized, whether due to regulation or internal policy, Stardog offers both graph virtualization and graph storage in a completely seamless blend. Use both in combination to support the needs of different data owners while still feeding your data fabric with all relevant enterprise data.



IN PRACTICE

At NASA, where the actual rocket science happens, Stardog is powering the design, test, and manufacturing lifecylce of complex systems like Space Launch System, the biggest rocket in the world. NASA uses Stardog's unique blend of graph-based storage and virtualization to connect data silos across NASA centers, vendor sites, and even across international borders.

READ THE CASE STUDY →

INFERENCE

Realize the full potential of your enterprise data

Stardog's Inference Engine associates related information stored in disparate sources, and then uses this rich web of relationships to discover new relationships within your data. By expressing all the implied relationships and connections between your data sources, you create a richer, more accurate view of your data.

Inference creates new relationships by interpreting connected data against your business logic in the data model. A knowledge graph's data model is often called an ontology or vocabulary and lays out common relationships between entities. This allows companies to describe complex domains, such as medicine, in which multiple facts, modeling constructs, and business rules interact with each other to imply new connections.

Some examples of inference in action: linking people to infrastructure via the applications they use, inferring new controls based on the similarity of new incidents to past incidents, and inferring links between investigators and therapeutic areas based on the conditions being investigated in studies they're working on — the list goes on.

WHAT IT DOES:

Inference analyzes enterprise data and infers new relationships, data values, and properties from your data, increasing its value exponentially. Like with any other network, the value of connected enterprise data grows exponentially with the number of connections, which is exactly what Stardog's Inference Engine capability creates automatically in the data fabric.

Stardog supports multiple inference schemas or data models at the same time applied to the underlying data fabric. By offering this support, Stardog can support multiple applications that require different interpretations of the same data! This is possible because Stardog connects data at the computation rather than storage layer. By contrast, other data integration approaches, including data lakes and data warehouses, connect data based on storage, in which case only one schema can be applied to that data. Which is one reason enterprises have to continually create new data silos for every new challenge or problem!

Stardog's innovative Inference Engine goes even further. Not only does it infer new data connections, but it can explain any new connection it creates. In contrast to black box recommendation systems, which cannot provide any explanation or rationale for their results, Stardog's Inference Engine can explain all inferences and results in terms of data, schema, and business rules. So users can review how Stardog arrived at an answer and the business logic referenced to do so. This explanatory transparency is not only critical for providing trusted results and accountability within an organization, but also necessary for certain legal and regulatory requirements.

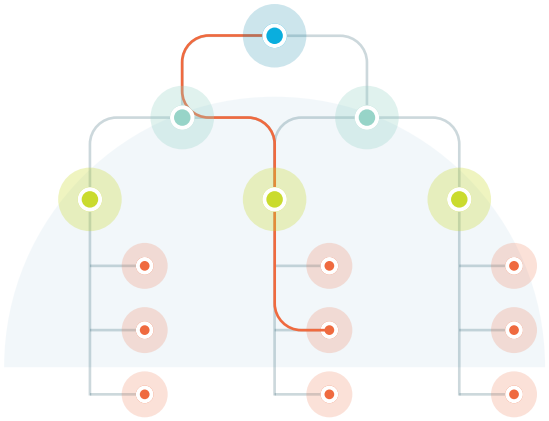


Figure. Ontologies are data models that show how various concepts relate to one another.

IN PRACTICE

NIH's Models for Infectious Disease Studies (MIDAS) Digital Commons facilitates collaborative epidemiological research to respond to disease outbreaks. Associating related research through an ontology allows researchers to search across a number of attributes, including pathogen type, host data, and disease forecasters. Users can now query over 700 mapped data sets, 62 indexed software applications, and over 200 data-related websites in 28 different formats.

READ THE CASE STUDY →

Connecting the Enterprise

While a knowledge graph is the key ingredient of the data fabric, it is not the only thing you need to be successful. Stardog has led dozens of companies through connecting their enterprise and can advise on every step of the process.

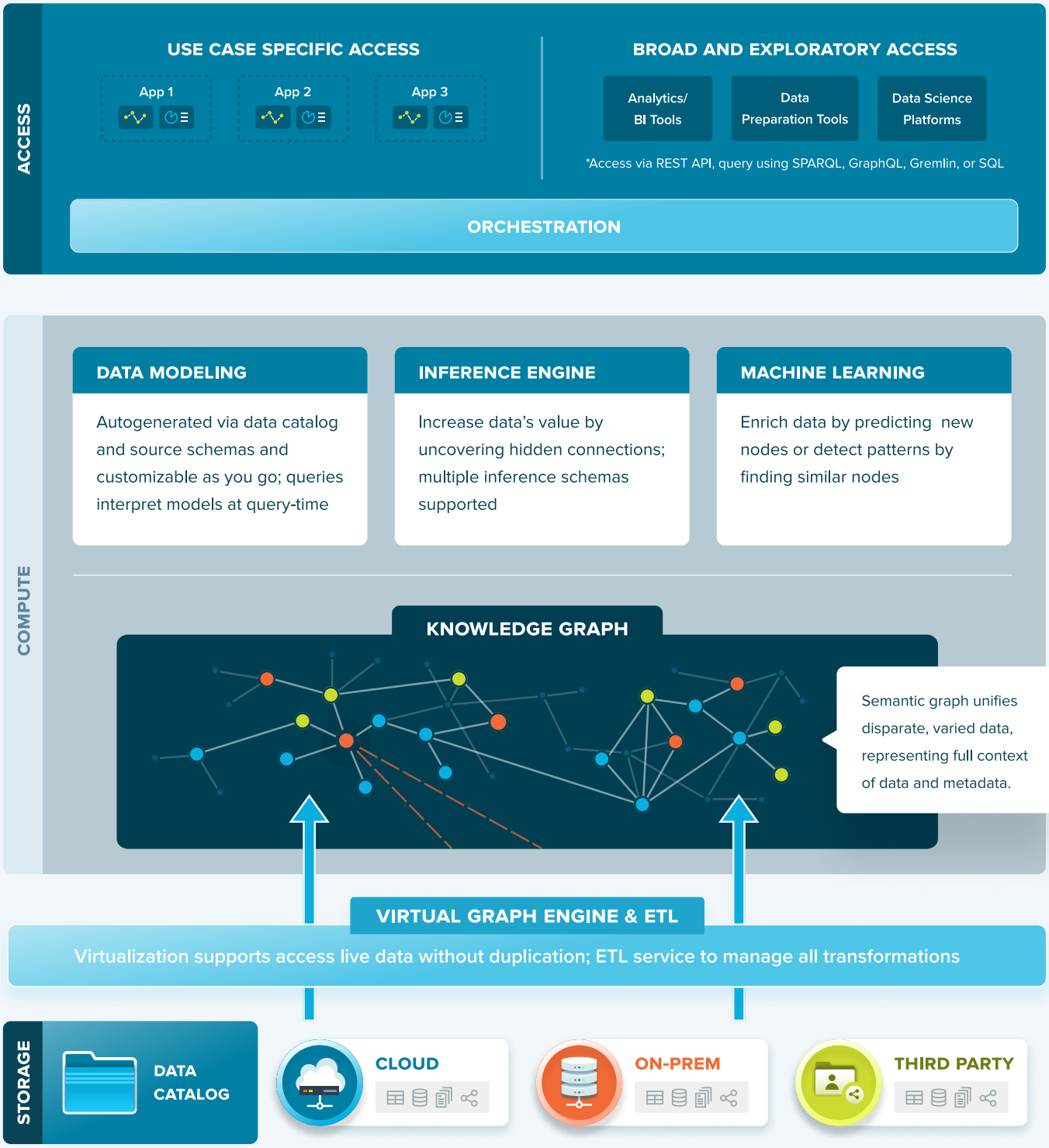
In this section, learn how to best use Stardog alongside your existing data management investments and how to successfully get started with your data fabric deployment.

The data fabric ecosystem

A successful data fabric requires leveraging and connecting existing source systems. Stardog's Virtual Graphs connect to existing data catalogs, data lakes, databases, and other data management platforms, offering comprehensive support of the most important enterprise data sources.

For data fabric deployments, Stardog recommends leveraging work completed in data catalogs to accelerate data discovery and semantic enrichment within Stardog. Using the data catalog as an input, Stardog builds a data map of your enterprise data assets. This data map accelerates data fabric creation through partially automated learning and auto-mapping of existing sources.

Stardog provides a seamless end-to-end user experience for creating and managing your knowledge graph as the key part of your data fabric deployment. Stardog Designer is a no-code, visual environment for data engineers and analysts to connect, map, model, and publish data. Stardog Explorer is an intuitive web browser tool for anyone to browse the connections in their knowledge graph. Stardog Studio is a feature-rich integrated development environment that makes working with Stardog quick and easy.


ACCESS
USE CASE SPECIFIC ACCESS
App 1
App 2
App 3
BROAD AND EXPLORATORY ACCESS
Analytics/ BI Tools
Data Preparation Tools
Data Science Platforms
*Access via REST API, query using SPARQL, GraphQL, Gremlin, or SQL
ORCHESTRATION
COMPUTE
DATA MODELING
Autogenerated via data catalog and source schemas and customizable as you go; queries interpret models at query-time
INFERENCE ENGINE
Increase data's value by uncovering hidden connections; multiple inference schemas supported
MACHINE LEARNING
Enrich data by predicting new nodes or detect patterns by finding similar nodes
KNOWLEDGE GRAPH
Semantic graph unifies disparate, varied data, representing full context of data and metadata.
VIRTUAL GRAPH ENGINE & ETL
Virtualization supports access live data without duplication; ETL service to manage all transformations
STORAGE
DATA CATALOG
CLOUD
ON-PREM
THIRD PARTY

Creating an enterprise data model

A common question regarding deploying a data fabric is how to develop an enterprise-wide data model. Many think this is a prerequisite to the initiative, and the undertaking may strike you as potentially expensive and time-consuming.

In fact, you only need to define as many concepts as needed for your initial use case. Identify a critical business problem to spearhead the broader data fabric initiative. Approach your data fabric with an MVP mindset and do strictly the minimal work to accomplish the first significant tranche of business value.

A key premise of Stardog's platform is that data modeling is reusable. When things change, simply write a new modular rule to amend the model and proceed with accessing your connected data. Due to this reusable data modeling principle, the business value derived from Stardog compounds over time.

IN PRACTICE

At one global pharma, in just 6 months of time with one back-end engineers and two front-end developers, a total of 13 enterprise data sources were modeled and unified. The resulting application was accessed by over 1,000 internal users.

Get a head start on data model development

There are many public data models that Stardog can read, helping customers to accelerate their data model development. A public data model may account for about 80% of modeling required for your project, with the remaining 20% customized based on your proprietary data or unique internal operations. Our team can help advise on publicly available data models that can suit your use case.

Stardog has also committed to the development of additional public data models through the **Cloud Information Model (CIM)**. CIM aims to provide ready-to-use data models for predefined domains that are not tied to any application or vendor.

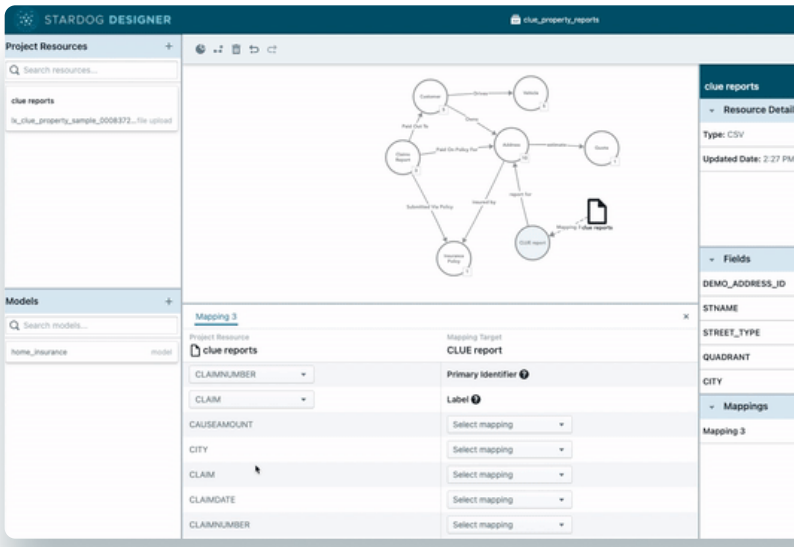


Figure. Mapping data from a CSV file into a model

Once you are ready to start modeling, Stardog Designer makes the process of building a model to meet your first use case even easier. For users new to graph data modeling, it allows you to visually create a semantic data model through a virtual whiteboard-like experience. Create a project, add classes, and link them together through relationships. Easily connect to data virtually from data lakes, such as Databricks, or warehouses with just a few clicks. Map in new data from CSVs and data source connections, all while referencing your model. Publish your model and mappings to flat files for your version control system or directly to your Stardog server for use in Explorer and Studio.

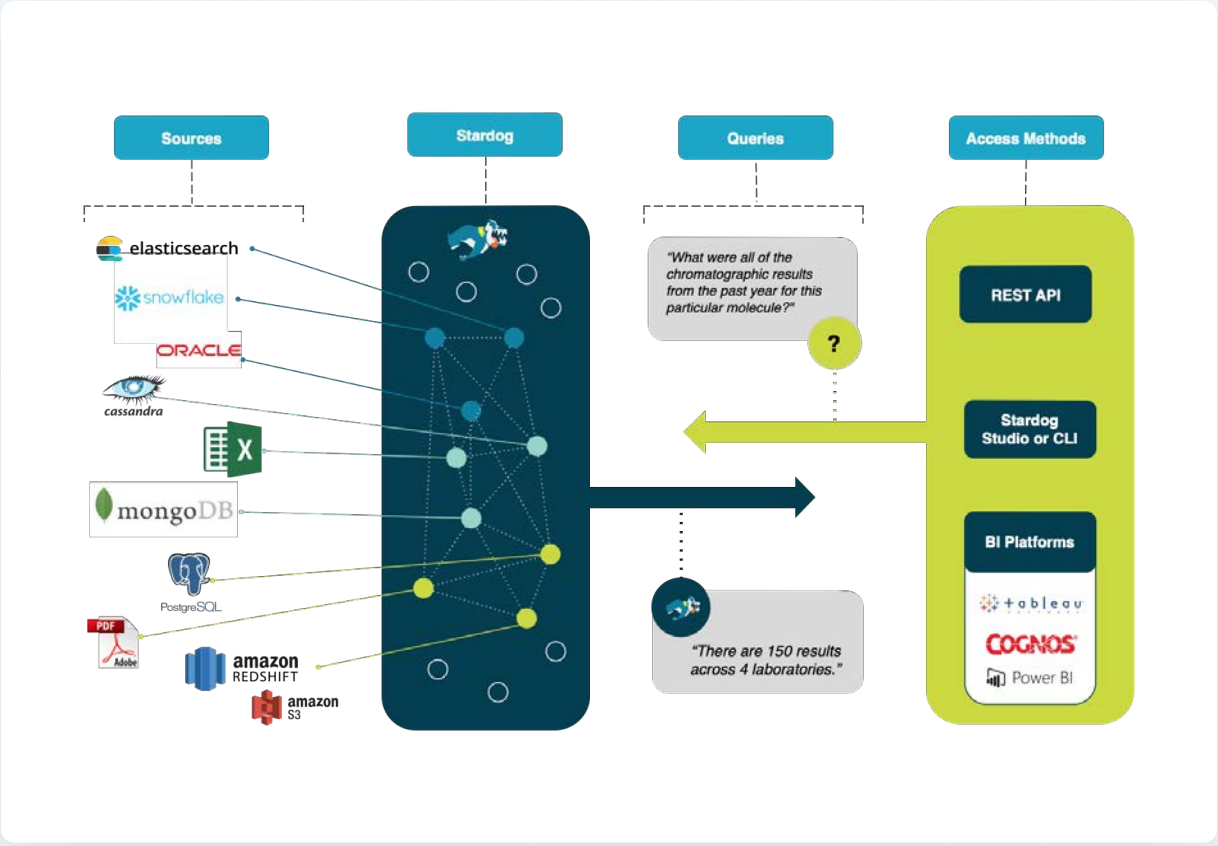
Share access to socialize your data fabric

With an MVP in hand, it's important to socialize your data fabric. A data fabric would be useless if the business meaning is locked away from the business. This lack of access has traditionally occurred due to two reasons:

- literal lack of access to the data; data is trapped in source systems or within IT only
- inability to access due to skill gap, ie lack of workers skilled in manipulating graph data

Stardog continually strives to make data more broadly available and usable. Stardog offers direct connections to popular business intelligence platforms via our **BI/SQL server**, which converts graph data back into SQL to make it available through all major SQL variants. You can use the BI/SQL Server to connect Stardog to any platform that runs on SQL. Or, you can use our supported Connectors to BI platforms including Tableau, Power BI, cumul.io, Apache Superset, Siren, IBM Cognos, Metabase, and RapidMinder.

Stardog improves upon the capabilities of these BI platforms. For example, as visualizations are created in Power BI, Virtual Graph queries would run behind the scenes allowing users to analyze data from multiple data sources as if all the data is stored in one data source. Similarly, if you had a dozen different point-of-sale data sources, you can write a Stardog rule defining the relevant columns as geographic coordinates so that Tableau can automatically display all twelve sources on one map.

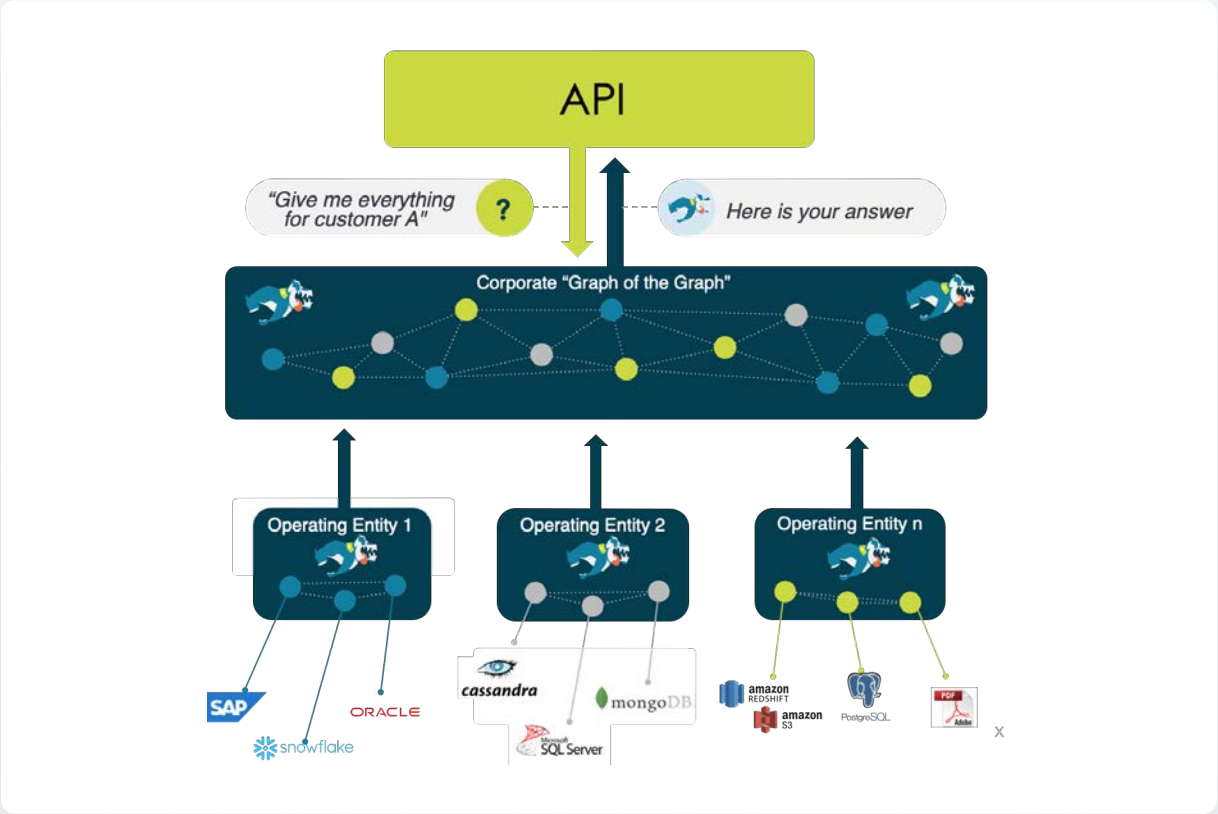


Stardog Access Method	Used By
REST API	**Applications** can access data from Stardog directly via a REST API endpoint.
BI/SQL Server	**Business analysts** gain direct access to the breadth of unified data directly in their BI platform of choice. Stardog's BI/SQL Server allows any BI platform that operates on any of the major MySQL variants to query Stardog.
CLI	**System administrators** have the option to access Stardog directly via the Command Line.
Stardog Designer	**Data engineers and business analysts** can visually connect, map, model, and publish data using Stardog Designer's virtual whiteboard-like experience.
Stardog Explorer	**Business analysts, data modelers, or anyone** can use Stardog Explorer to quickly see, understand, and search using the relationships in connected data.
Stardog Studio	**System administrators and data modelers** can use our IDE, Stardog Studio, to query, visualize data, explore data models, and evaluate data provenance. Stardog supports SPARQL, GraphQL, and Gremlin query languages.
Python extension	**Data scientists** can access the unified data via Stardog's python extension pystardog. Data scientists can also use Stardog's **built-in machine learning** to train models directly on the virtualized data.

Build a compliant global data fabric

Use Stardog's **data quality constraints** to manage overall data quality and ensure conformance with defined rules. Constraints also support measuring the quality of the data, performing verification after an integration, and assisting in planning future improvement measures.

As your data fabric grows, Stardog grows with you. Stardog also has the ability to query other Stardog instances, which is key for compliance with data movement regulations. For organizations with distributed environments, it is still possible to query across operating entities without copying any data. Set each operating entity up with their own Stardog instance and Stardog can easily execute a global query across all virtualized data.



Get started today

Stardog makes it easy to get started with your data fabric. In addition to our platform detailed above, we have the team and expertise to take you from MVP to global deployment! Contact us to learn more about our customers who have successfully reduced time to insight 50-90%.

Build your data fabric with Stardog

Say "yes" to every data request by creating a flexible, reusable data layer for answering complex queries across data silos.

Stardog, the leading Enterprise Knowledge Graph platform, turns data into knowledge to power more effective digital transformations. Industry leaders including BNY Mellon, Bosch, and NASA use Stardog to create a flexible data layer that can support countless applications. Stardog has been recognized by *Fast Company* as one of the world's Most Innovative Companies, by *Database Trends and Applications* as one of the 100 companies that matter most in data management, and by *KMWorld* as one of the 100 companies that matter most in knowledge management. Stardog is a privately held, venture-backed company headquartered in Arlington, VA.

Contact us: sales@stardog.com

Learn more: stardog.com