ontotext

Today's enterprises are increasingly daunted by the realization that more data doesn't automatically equal deeper knowledge and better business decisions. The mere existence of the 175 zettabytes of data the International Data Corporation estimates the world will possess by 2025 doesn't matter if organizations can't leverage it effectively. Obviously, not all of that data is accessible to businesses, but what they can access is still overwhelming. From internal data to third-party data, from highly structured to video content, that data is diverse and complex. It's also usually stored in various systems that don't talk to each other. As a result, a big part of enterprise data remains practically invisible.

How enterprises choose to manage their data so they can take the most advantage of it depends on what they want to achieve. But whatever their business goals, in order to turn their invisible data into a valuable asset, they need to understand what they have and to be able to efficiently find what they need.

Enter metadata

Metadata describes data and includes information such as how old data is, where it was created, who owns it, and what concepts (or other data) it relates to. It enables us to make

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

sense of our data because it tells us what it is and how best to use it. As a result, leveraging metadata has become a core capability for businesses trying to extract value from their data.

Although many types of databases store metadata as well as data, semantic graph databases offer unique advantages. They smoothly integrate heterogeneous data from multiple sources and use semantic schema and semantic metadata to describe this data. The resulting knowledge graph provides enterprises with a single point of access to all types of valuable information related to their business. This information is interlinked and put into context, which makes it easy to find and analyze.

But to make these advantages tangible, let's look at them in the context of an investment bank that I'll call Fantastic FinServ. Fantastic FinServ is a fictional company, but it bears resemblance to many of the organizations that actually use knowledge graphs in this way. Like many companies in its space, Fantastic Finserv competes on knowledge. To win clients and keep them happy, its analysts must demonstrate comprehensive knowledge of their customers' industries and offer unique insights that can't be gleaned elsewhere.

Connecting the dots of data of all types

To begin with, Fantastic Finserv has to handle a wide variety of data. This includes traditional structured data such as:

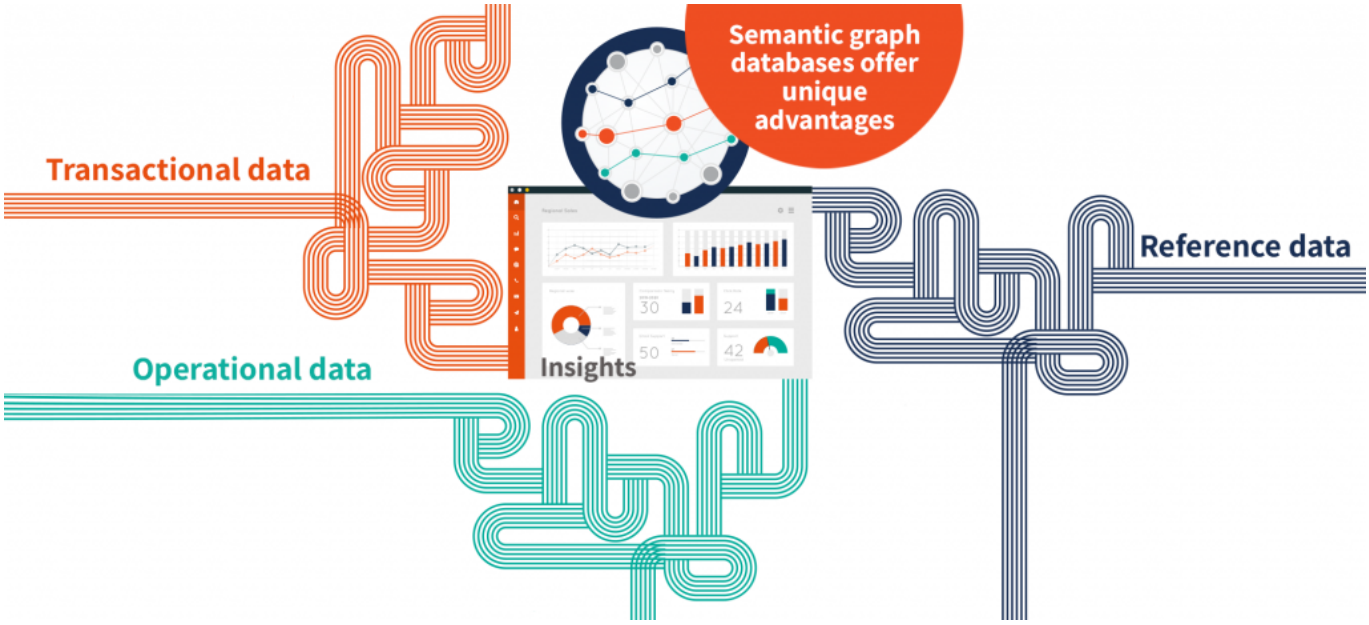
- **Reference data** – the data used to relate data to information outside of the organization. Think zipcodes, currencies, country codes, product lists, customer segments, etc.
- **Operational data** – the data generated by a Fantastic Finserv, itself. It includes information on employees, competitors, inventory, fleets, and anything else that has a day-to-day impact on running the business.
- **Transactional data** – a subcategory of operational data consisting of recorded business events. This captures information about orders, payments, customers, invoices, and so on.

Structured data comes in rows and columns and is neatly stored in tables. For decades it's been the bread and butter of enterprise data teams. Most legacy approaches to data management, like relational databases and data warehouses, focus on structured data. But in

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext

the modern world, more and more data comes from unstructured sources, and here's where knowledge graphs gain the upper hand.



What's the big deal about unstructured and semi-structured data?

Unstructured data might seem like an oxymoron, but it just means data that doesn't come in tables. Semi-structured data consists of formats such as JSON or XML that have rigidly defined structures but aren't made of columns and rows. Videos, pictures, and written documents are considered fully unstructured data because the information in them can't be automatically extracted without advanced tools like machine learning.

Because of the greater difficulty of dealing with semi-structured and unstructured data, enterprises have historically neglected data in those formats. But 80-90% of all data is unstructured causing those businesses to miss out on potential insights. The flexible and dynamic structure of knowledge graphs makes those sources easier to use and allows organizations to manage them just as they would structured data assets.

As a result, Fantastic Finserv can incorporate and manage additional sources of information

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext

that its competitors leave on the table. Emails, documents, news articles, even recordings of speeches can all be linked and made queryable in its knowledge graph. This gives the company a unique base of knowledge from which to draw, differentiating it from competitors who all rely on the same tabular data subscriptions. Especially when working in developing economies, this ability to make systematic use of non-traditional data sources gives Fantastic Finserv the edge.

Knowledge (metadata) layer

In addition to simply being able to deal with data in all formats and from multiple sources, knowledge graphs add an extra layer. This layer transforms the information from data to knowledge by putting it into context and relating it to one another using metadata. At its core, a knowledge graph is like a concept map. It stores data as objects that connect to one another through highly defined relationships. Through those definitions, which are a form of metadata, a knowledge graph begins to surface emergent information.

For instance, say Fantastic Finserv has a client that wants information on start-ups in Southeast Asia as it considers an acquisition in that region. That question seems straightforward, but what is a “start-up?” Different data sources might have conflicting definitions and the client another altogether.

Using a knowledge graph allows Fantastic Finserv to build its own definition of a start-up directly into the graph and link it with how other organizations understand the concept. Then, it can query the graph for information about prospects. The report it delivers to its client will be different from what other banks would have been able to provide because the robust definitions inherent in the graph will allow it to identify companies as start-ups that were never explicitly referred to as such in the data layer. That knowledge is instead captured in the metadata layer.

Applications

The robust and flexible nature of semantic metadata makes a semantic knowledge graph an obvious fit to serve as a metadata hub within an organization. Such a hub, which centralizes the data about a company’s data in a single, searchable location, has tremendous benefits

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext

for the business.

The Fantastic Finserv metadata hub acts like a data catalog, allowing its analysts to rapidly find, and *find out about* different data assets the company may possess using descriptions of assets that are stored in the graph. They spend less time searching for the right data and more time performing actual analysis, improving the quality and depth of their reports. Analysts can also use the knowledge graph as a data fabric to create new data products.

In a data fabric, data is automatically ingested or virtualized from a wide array of source systems. Each of those systems has its own approach to handling data – different labels, schema, format, and data types. If that sounds like metadata, that's because it is. A knowledge graph can store all of that metadata about the data coming from each system, allowing the data itself to be disentangled from the design decisions of software developers.

As a result, the knowledge graph can effectively virtualize queries across the entire fabric because it knows how to speak the language of each source system. Thanks to the metadata hub, a Fantastic Finserv analyst can define and query a data product that pulls comprehensive results from multiple platforms, allowing for synthesis of information that simply isn't possible with traditional approaches.

Conclusion

A metadata hub uses the power of data about the data to help businesses make better use of the information they already possess but is out of their reach. It empowers them to understand the data assets they have and to easily find what they need.

For a company like Fantastic FinServ, utilizing a knowledge graph as a metadata hub provides a competitive advantage. In knowledge-based industries, there's no substitute for meaning-centered approaches to data. Being able to combine and analyze metadata to capture knowledge can mean the difference between success and failure. As we move from the data revolution into a new metadata revolution, those capabilities will become more important than ever.

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext

Want to utilize a knowledge graph as an enterprise-wide metadata hub?

Book a Discovery Call

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com