What is the future of knowledge graphs in the era of ChatGPT and Large Language Models?

Atanas Kiryakov: Knowledge graphs will prosper in the ChatGPT era. To start with, large language models (LLM) will not replace databases. They are good for compressing information, but one cannot retrieve from such a model the same information that it got trained on. At the same time, most data management (DM) applications require 100% correct retrieval, 0% hallucination!

LLM will not replace knowledge graphs either. Many enterprise data and knowledge management tasks require strict agreement, with a firm deterministic contract, about the meaning of the data. When sales performance is analyzed and correlated with marketing data, for instance, it is critical to make sure that across the board there is good alignment regarding the categories of products, the regions, the suppliers and the relationships between them. Otherwise the results are nice diagrams that are useless or harmful.

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

It's important to realize that knowledge graphs can be used to fine tune and customize LLMs. They can also help with "oversight": validation, explanation and filtering of the answers of an LLM. From another view, LLMs can help make the process of developing and using knowledge graphs much easier. It becomes possible to build richer graphs with more context in less time and get better insights with less complexity. This is precisely what we need in order to deliver knowledge graphs to a much broader set of enterprises.

Let me share a few examples of how LLMs help knowledge graphs. Think of querying your graphs in natural language instead of SPARQL or GraphQL. And getting a free text summary of the results, instead of just a table. Think of enrichment of graphs using public LLM services via relation extraction from text. I am very optimistic!

Doug Kimball: We have already developed some of these features and will release them very soon, stay tuned.

LLMs can help make the process of developing and using knowledge graphs much easier. It becomes possible to build richer graphs with more context in less time and get better insights with less complexity. Share on X

How will LLM impact text analysis and content analytics?

Atanas Kiryakov: The future is not so bright for many in this field. This is the case with the so-called intelligent data processing (IDP), which uses a previous generation of machine learning. At the heart of such tools is the extraction of fields from forms or specific attributes from documents. LLMs do most of this better and with lower cost of customization.

Luckily, the text analysis that Ontotext does is focused on tasks that require complex domain knowledge and linking of documents to reference data or master data. That's something that LLMs cannot do. This is the kind of precision that gets lost in the vector embeddings that stay at the heart of the language models. We use other deep learning techniques for such tasks. Still, LLMs have a role to play – they can make our text analysis pipelines much more efficient for tasks like sentiment analysis, classification and event detection. We will demonstrate the results of such use of LLMs soon.

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

Back to the mainstream, how does Ontotext help deliver better analytics?

Doug Kimball: Using our knowledge graph, you can develop more complex analytics, such as data mining, Natural Language Processing (NLP) and Machine Learning (ML). With traditional data management systems, that can be difficult or in some cases can lead to more work than results. The rich semantics built into our knowledge graph allow you to gain new insights, detect patterns and identify relationships that other data management techniques can't deliver. Plus, because knowledge graphs can combine data from various sources, including structured and unstructured data, you get a more holistic view of the data. This enables more accurate and comprehensive analytics, which can help drive better decision-making.

Atanas Kiryakov: From a technical perspective, the advantages that Doug explained can be summarized in two major directions:

- Rich domain knowledge – you can think of it as Master Data++
- Better speed and cost for data preparation and enrichment

How is what Ontotext does different from what an Master Data Management vendor does?

Doug Kimball: Ontotext focuses on creating and managing semantic knowledge graphs (knowledge graphs), which integrate diverse datasets. This allows the knowledge graph to represent data in a way that captures the meaning and relationships between different entities, which among many things, allows for advanced reasoning and analytics. Using semantic technologies, knowledge graphs enable machines to understand the meaning behind the data, making it easier to extract insights and make connections between disparate data sources.

Master data management (MDM), on the other hand, is focused on ensuring data quality and consistency across different systems and applications. MDM creates a single, verified source of data that can be used throughout an organization, and enforce policies and standards to

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

maintain data quality. You don't have to do MDM in order to have a knowledge graph, but a knowledge graph will provide semantic understanding of the data delivered by MDM to add meaning and usability, as well as depth of insights.

In a way, MDM is like the letters in the alphabet and knowledge graphs are like sentences, because it makes sense of the letters. Long story short, MDM is mostly about data quality and precision (e.g., deduplication of people and addresses). Knowledge graphs come with a higher level of ambition; deeper understanding of the data, which allows for better interpretation, improved search and more insights (e.g., discovering relationships between people).

Atanas Kiryakov: I can add that a knowledge graph is a much broader concept than MDM. Knowledge graphs add more formal semantics (meaning) to the data in order to allow computers to interpret them in ways that are more advanced than frequency counting in tables and string look. Knowledge graphs use knowledge models with formal semantics (the magical ontologies) as semantic data schema in order to allow deeper interpretation for transactional data or documents.

Can you tell us more about the knowledge models?

Atanas Kiryakov: These conceptual models include many types of semantics/knowledge that go across multiple mainstream IT disciplines:

- **Entity relationship model,** like the one behind the schema of the relational databases, the UML and the object-oriented programming.
- **Named entities** such as people, organizations and locations. Those include the "business entities" that are central to MDM. While a typical MDM system will be focused mostly on customers, suppliers, their addresses and points of sale, the model used in a knowledge graph can include a much broader set of entities. Think of a knowledge graph that includes the knowledge about all geographic features on Earth and all organizations (above some threshold of importance) in a given region or domain.
- **Reference data** used to classify or categorize other data, including units of measure, codes, as well as controlled vocabularies of terms and taxonomies of topics.
- **Technical knowledge**, e.g., the configurations of all devices in a building, used to better interpret the sensor data and economize heating energy, or the parameters of all

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

actors in a national electricity grid – produces, wires, transformer stations, etc.
- **Product information** of the sort managed in Product Lifecycle Management (PLM) and Product Information Management (PIM) systems.
- **Scientific knowledge**, e.g., chemical compounds and drug names.

How is what Ontotext does different from a CMS?

Doug Kimball: A Content Management System (CMS) is mainly focused on managing, organizing and publishing content, while a knowledge graph is designed to capture the meaning behind the data. A CMS may include features that allow for tagging or categorizing content, but it does not have the advanced reasoning and analytics capabilities that a knowledge graph provides. A big advantage to a knowledge graph in support of a CMS is that knowledge graphs can work with both structured and unstructured data (such as the text or images you might have in a CMS) – allowing you to do more and gain more insights from your CMS system.

Atanas Kiryakov: A CMS typically contains modest metadata, describing the content: date, author, few keywords and one category from a taxonomy. When knowledge graphs are used for content management, there is rich semantic metadata that describes the content to enable easier discovery in a broad range of scenarios. Text analysis (NLP) is used for semantic annotation, which means enrichment of the content with tags such as:

- **Entities**, e.g., people, organizations, locations, products, etc.
- **Topic and categories** from a taxonomy, the same like in a mainstream CMS, but having more of those attached to each document.

So what are the benefits of a knowledge graph-based CMS for enterprises?

Atanas Kiryakov: A knowledge graph-based CMS allows for more precise hybrid queries, which combine full-text search, with structured queries and reasoning. The criteria for filtering can combine look-up in the row text of the document, with structured query about the entities and the topics mentioned in it. Think of querying a big collection of product documentation for "user guides for products of type X that use components manufactured by

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

suppliers in sanctions list Y, which will be distributed in the EU". Such query is not possible without a knowledge graph, because the following knowledge is not present in the document, but rather in databases and in other systems:

- Which component is produced by which manufacturer?
- Which manufacturers are part of what sanctions list?
- Which countries are part of the EU?

Doug Kimball: To wrap it up, knowledge graphs should be integrated with CMS to allow for better search: find a complete list of documents in less time and without false positives. Media, such as BBC and Financial Times, integrate knowledge graphs with their digital asset management systems (Digital Asset Management – DAM) in order to:

- Improve user engagement with more interactive exploration, e.g., generation of topical webpages, hyperlinking and personalized recommendations
- To automatically re-package content for the sake of b2b feed generation
- Better repurposing of existing materials

How does Ontotext help enterprises with data fabric projects?

Doug Kimball: Knowledge graphs help support data fabric projects as they can integrate data from disparate sources and provide a unified view of the data to ensure usability and access. Using our semantic knowledge graph, you add quality and consistency to your data being delivered through the data fabric, making it easier to analyze and report on. If you are looking to deliver enhanced analytics from your data fabric project, a knowledge graph is essential. Delivering better data mining, NLP and pattern identification – you can make better decisions.

Atanas Kiryakov: Technically, the data fabric paradigm employs rich metadata to streamline data discovery, integration and preparation. Knowledge graphs are all about using semantic metadata to serve these purposes. Data across different sources is aligned and enriched by using semantic models (ontologies) that have the necessary expressivity to capture different types of knowledge: database schema, taxonomies, master and reference

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

data, alongside technical and product information, when necessary. Describing and unifying the data from the different sources using these semantic models allows for unambiguous interpretation, easier updates, support and gradual development of the knowledge graphs.

How can Ontotext help enterprises planning to expand their AI spending?

Atanas Kiryakov: AI projects critically depend on data. Knowledge graphs help both analytics and AI in two fundamental ways:

They reduce time-to-insight and cost by reusing data preparation work. Data preparation must take place before data scientists can do their magic. It involves gathering, integration, normalization and unification. This can be a lot of work! And it's often not reused across projects. Knowledge graphs allow this to be done in a sustainable manner, building unified views that can be maintained, updated and extended as necessary over time.

They enrich the data to allow for better interpretation and richer insights. One cannot understand the trends in consumption unless the relevant additional data is linked, e.g., weather, traffic, market trends, etc. Knowledge graphs enable data enrichment in several ways:

- extraction of facts and relationships from unstructured data via NLP
- easier integration of 3rd party data
- discovery and materialization of multi-hop relationships, which can serve as input for ML
- applying the most advanced ML techniques directly on the graph structures, resulting in link prediction, which is more powerful than the ML techniques working on tables

Technically, knowledge graphs use ontologies and taxonomies to capture the meaning behind the data. This helps AI models to better interpret the relationships between entities and concepts in a particular domain, allowing for more accurate and effective decision-making. This is critical for many applications, including fraud detection, risk assessment, and recommender systems.

Doug Kimball: All in all, the best way to get more competitive insights from your data is to gain advantage in the speed and cost for data preparation and enrichment. This way you can

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

get more insights, quicker and at a lower cost. This is how knowledge graphs make your AI projects more efficient and productive.

Do you want to benefit from Ontotext's semantic approach towards data and content management?

Book a Discovery Call

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