Author: Atanas Kiryakov

In 2023, data leaders and enthusiasts were enamored of — and often distracted by — initiatives such as generative AI and cloud migration.

The generative AI buzz and interest in cloud migration shouldn't be ignored, but as with any technology that requires data strategy, it's critical that data and analytics professionals be crystal clear about their priorities and confident in the projects that will positively impact their business and goals.

As companies in almost every market segment attempt to continuously enhance and modernize data management practices to drive greater business outcomes, organizations will be watching numerous trends emerge this year. These will include developing a better understanding of AI, recognizing the role semantic metadata plays in data fabrics, and the rapid acceleration and adoption of knowledge graphs — which will be driven by large language models (LLMs) and the convergence of labeled property graphs (LPGs) and resource description frameworks (RDFs).

I expect to see the following data and knowledge management trends emerge in 2024.

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

Trend #1: Organizations will (finally) manage the hype around AI

As the deafening noise around generative AI reaches a crescendo, organizations will be forced to temper the hype and foster a realistic and responsible approach to this disruptive technology. Whether it's an AI crisis around the shortage of GPUs, climate effects of training large language models (LLMs), or concerns around privacy, ethics, bias, and/or governance, the challenges will worsen before they get better, leading many to wonder if it's worth applying generative AI in the first place.

Although corporate pressures may prompt organizations to "do something with AI," being data-driven must come first and remain top priority. After all, ensuring foundational data is organized, shareable, and interconnected is just as critical as asking whether generative AI models are trusted, reliable, deterministic, explainable, ethical, and free from bias.

Before deploying generative AI solutions to production, organizations must be sure to protect their intellectual property and plan for potential liability issues. This is because although generative AI can replace people in some cases, there is no professional liability insurance for LLMs. This means that business processes that involve generative AI will still require extensive "humans-in-the-loop" involvement which can offset any efficiency gains.

In 2024, expect to see vendors accelerate enhancements to their product offerings by adding new interfaces focused on meeting the generative AI market trend. However, organizations need to be aware that these may be nothing more than bolted-on Band-Aids. Addressing challenges such as data quality and ensuring unified, semantically consistent access to accurate, trustworthy data will require setting a clear data strategy as well as taking a realistic, business-driven approach. Without this, organizations will continue to pay a "bad data tax" as AI/ML models will struggle to get past a proof of concept and ultimately fail to deliver on the hype.

Trend #2: Knowledge graph adoption accelerates as

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

LLMs and technology converge

A key factor slowing down knowledge graph (KG) adoption is the extensive (and expensive) process of developing the necessary domain models. LLMs can optimize several tasks, such as updating taxonomies, classifying entities, and extracting new properties and relationships from unstructured data. Done correctly, LLMs could lower information extraction costs because the proper tools and methodology can manage the quality of text analysis pipelines and bootstrap or evolve KGs at a fraction of the effort currently required. LLMs will also make it easier to consume KGs by applying natural language querying and summarization.

Labeled property graphs and resource description frameworks will also help propel knowledge graph adoption because each is a powerful data model with strong synergies when combined. Although RDFs and LPGs are optimized for different things, data managers and technology vendors are realizing that together they provide a comprehensive and flexible approach to data modeling and integration. The combination of these graph technology stacks will enable enterprises to create better data management practices, where data analytics, reference data and metadata management, and data sharing and reuse are handled in an efficient and future-proof manner. Once an effective graph foundation is built, it can be reused and repurposed across organizations to deliver enterprise-level results instead of being limited to disconnected KG implementations.

As innovative and emerging technologies such as digital twins, IoT, AI, and ML gain further mind share, managing data will become even more important. By using LPG and RDF capabilities together, organizations can represent complex data relationships between AI and ML models as well as track IoT data to support these new use cases. Additionally, with both the scale and diversity of data increasing, this combination will also address the need for better performance.

As a result, expect knowledge graph adoption to continue to grow in 2024 as businesses look to connect, process, analyze, and query the large volume of data sets currently in use.

Trend #3: Data fabric comes of age and employs

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

semantic metadata

Good decisions rely on shared data, especially the right data at the right time. Sometimes, the challenge is that the data itself often raises more questions than it answers. This trend will continue to worsen before it improves, as disjointed data ecosystems with disparate tools, platforms, and disconnected data silos become increasingly challenging for enterprises. This is why the concept of a data fabric has emerged as a method to better manage and share data.

Data fabric's holistic goal is the culmination of data management tools designed to manage data from identification, access, cleaning, and enrichment to transformation, governance, and analysis. That is a tall order and will take several years to mature before adoption happens across enterprises.

Current solutions were not fully developed to deliver all the promises of a data fabric. In the coming year, organizations will incorporate knowledge graphs and artificial intelligence for metadata management to improve today's offerings, and these will be a key criterion for making them more effective. Semantic metadata will enable decentralized data management, following the data mesh paradigm. It will also provide formal context about the meaning of data elements that are governed independently, serving different business functions and embodying different business logic and assumptions. Additionally, these solutions will evolve and incorporate self-learning metadata analytics, driving data utilization pattern identifications to optimize, automate, and access domain-specific data through data products.

Data security, access, governance, and bias issues continue to routinely impact daily business, and with generative AI getting so much attention, organizations will look to leverage a data fabric powered by semantic technologies to lower cost of ownership and operating costs while improving data sharing and trust.

What Can We Expect?

In 2024, we stand on the verge of extraordinary technological advancement. Keeping these trends in mind, and embracing the ever-changing technological environment, successful

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

businesses will apply a data strategy that is driven by a business results mindset. More important, they will apply this to strategically position themselves for success by using data as a foundational asset to deliver growth, innovation, and competitive advantage.

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