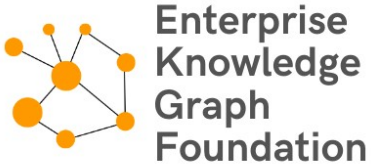
Knowledge Graph Conference (KGC)

2021 Knowledge Graph Industry Survey

Examination of the Practices and Priorities Driving Knowledge Technology



Organizations working to improve the function and value of data have made significant progress over the past few years. The driver of progress has been a more substantive adherence to the core principles of data management combined with maturity of semantic standards and knowledge technology. This survey, sponsored by the Knowledge Graph Conference (KGC), was designed to capture information about the size and nature of that progress and is part of an overall initiative to support the growing field of knowledge expression.

The results give reason for optimism. Adoption is growing across a broad spectrum of industry sectors. Companies are emerging from the era of experimentation and beginning to operationalize their pilot initiatives. Use cases are designed to facilitate data integration, semantic search and flexible analysis. Companies using knowledge technologies are beginning to understand that delivering business value is dependent on building foundational prerequisites of data harmonization and entity resolution. We applaud this recognition of the importance of establishing a controlled and predictable data infrastructure as the key to unraveling complex and interconnected environments.

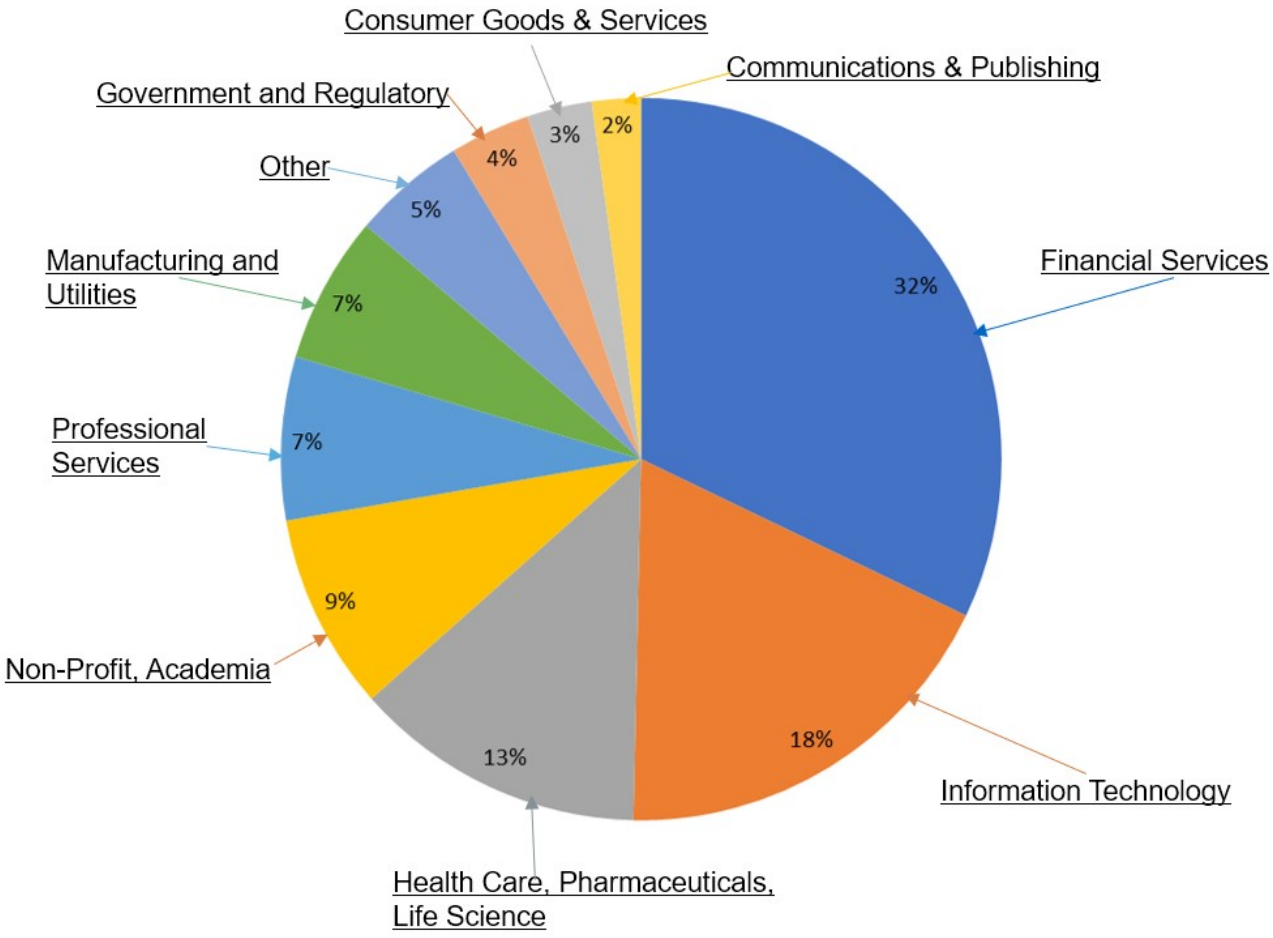
This industry survey was structured into three segments. The first paints a profile of these early adopters by industry type, experience and organizational function. The second defines the business contour of maturity, drivers, inhibitors and use cases. The third examines industry priorities, requirements for vendors and views on the future direction for knowledge technology. Individual responses to the survey were prioritized based on importance and given a weighted score to help readers make comparisons.

The bottom line tells a story of progress in a moment of transformation. Overcoming organizational inertia after decades of silo-based data management is not a simple task. Our industry is still emerging from conducting pilots to operationalizing the value proposition of semantic technology. We are excited to see so many diverse companies cross that Rubicon and look forward to a groundswell of business adoption over the next few years.

Respectfully submitted

Michael Atkin	François Scharffe	Thomas Deely
Managing Director	Co-Founder	Co-Founder
Content Strategies LLC	Knowledge Graph Conference	Knowledge Graph Conference

Enterprise Knowledge Graph Survey Profile



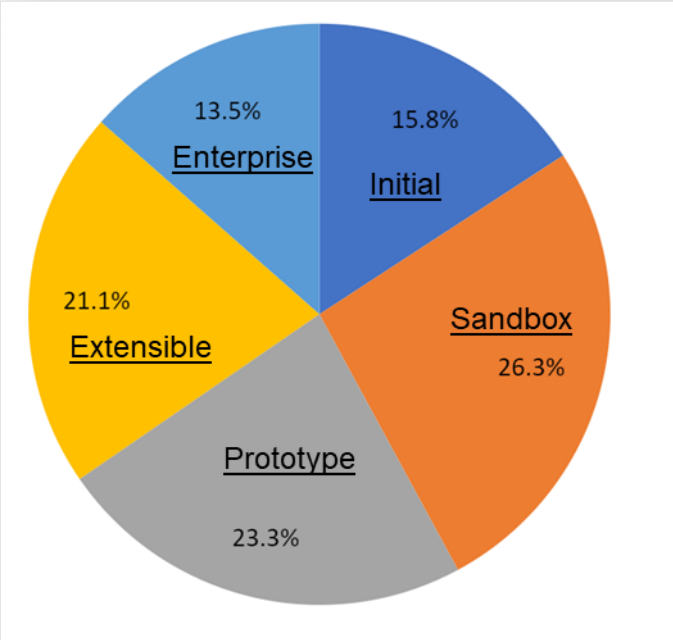
Analysis: Approximately 150 companies responded to the knowledge graph industry survey. Most are in information-intensive sectors characterized by interconnected lines of business where precision and nuance of data matter.

- 50% of participants have been working with knowledge graphs and related technologies for less than 3 years.
- 56% of responders are managing their knowledge graph on an enterprise-wide basis.
- 46% of participants work for companies with less than 100 employees (innovation and enablement drivers) while 35% work for companies with more than 5,000 employees (efficiency and regulatory drivers).

Most of the responders fall into the category of “engaged practitioners” who are already familiar with semantic standards and knowledge graph technology. The majority (about 60%) work in support functions including data management, ontology development, business architecture, data engineering and modeling. Just over 18% of responders are aligned with revenue generation including sales, marketing and other functions within a business unit.

Viewpoint: This profile meets expectations. Most industry sectors are still in the early stages of adoption of knowledge technologies. The responders were mostly the innovators who see the capability potential … academia who are doing the research … and vendors that are driving maturity. We can easily draw a parallel between the profile of the knowledge graph industry and that associated with the prioritization of “core data management” in the 2010s.

Knowledge Graph Maturity



Key to Maturity Levels

Initial [Level 0] **–** No current knowledge graph initiatives

Sandbox [Level 1] **–** Pilots and POCs for limited use cases

Prototype [Level 1] – Operationalizing pilots and POCs for limited use cases

Extensible [Level 2] **–** Production ready for multiple (related) use cases; reusable architecture

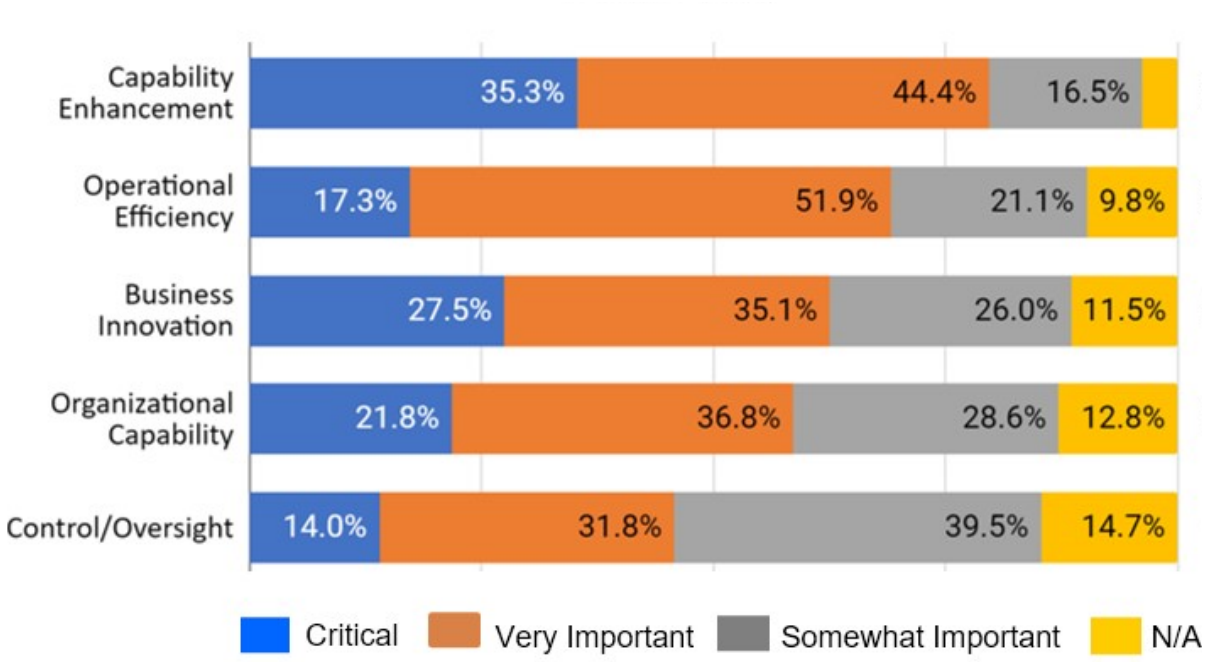
Enterprise [Level 3] **–** Multiple groups; scalable architecture; mission-critical applications

Analysis: Knowledge graph adoption is still "emerging" across most industry segments. Almost 50% of survey responders report they are experimenting with knowledge graph capability in the form of proofs-of-concept. At this stage of maturity, stakeholders are starting to recognize the business liabilities from incongruent data, driven by internal champions who are willing to take on the disruption challenge. Technology strategy is supporting the experimentation, with a limited infrastructure often characterized by mostly manual transformation and dedicated efforts to build initial knowledge graph components. The promising news is that some 23% of responders have moved beyond the "sandbox" (POC) stage and are operationalizing their pilot initiatives. Moving from POC to production is a notable step along the maturity curve.

Just under 35% of responders are implementing knowledge graph as part of the underlying "data infrastructure" for their firms. This means that stakeholders are beginning to adopt a data centric mindset focused on strategic business value. The technology architecture is based on expanded design principles. Use cases are defined and modeled to capture shared data relationships and we are beginning to see tailored operating models emerge to support the unique requirements of the knowledge graph. Responders at these stages are predominately big companies in defense, publishing, consumer services, finance, health care, manufacturing and life sciences.

Viewpoint: The criteria for advancing from one stage of maturity to the next are significant. The movement from "initial" to "pilot" requires commitment, funding and core data management capability (i.e., governance, inventory, managed glossaries, etc.). The movement from pilots to extensible platforms emphasizes content reusability with sustainable resource commitments. At this stage and beyond, we should begin to see the knowledge graph emerge as an authoritative source for data for related use cases.

Knowledge Graph Adoption Drivers

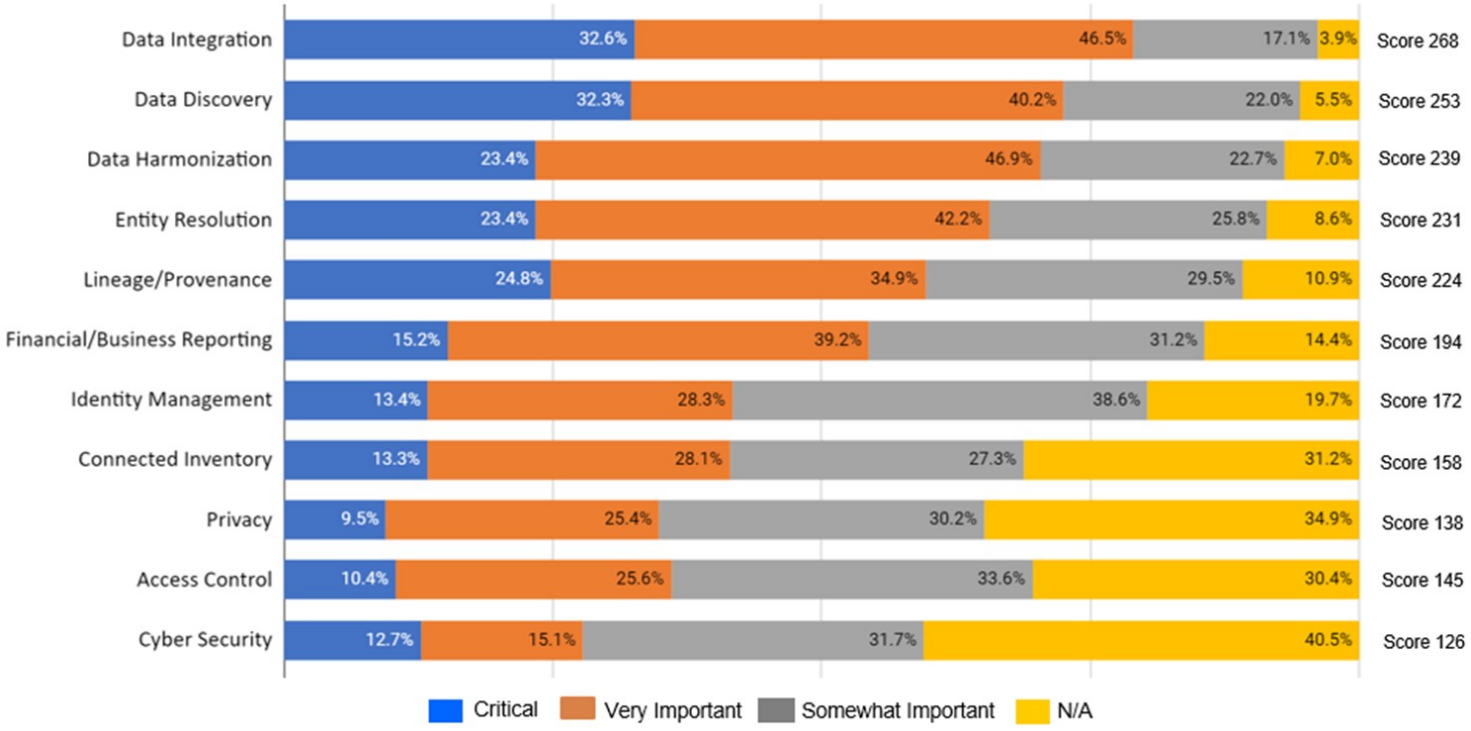


- **Capability Enhancement** [score 281] **-** analytical flexibility, data science, improved data quality
- **Operational Efficiency** [score 235] **-** automation, cost containment, digitization, resiliency
- **Business Innovation** [score 234] **-** client 360, improved customer experience, product and market insight
- **Organizational Capability** [score 223] **-** no better way to address “insolvable” problem
- **Control/Oversight** [score 187] **-** regulatory compliance and simplification of governance

Analysis: Business innovation and capability enhancement are the clear drivers and considered critical to success of knowledge graph initiatives. Most firms are looking to knowledge technologies to support data science and enhance analytical objectives. There is no question that the rationale for adoption is rooted in improved customer service, better product development and flexible scenario-based analytics. The other side of the adoption story focuses on the value proposition associated with operational efficiency and the alignment of processes across business and functional silos. The financial industry (in particular) has been propelled by regulatory pressure to achieve a “control environment” across interconnected functions for transactions reporting, privacy management and risk aggregation.

Viewpoint: Demonstrating business value is dependent on building foundational prerequisites. Content integration, entity resolution, quality validation, mapping and lineage traceability are not the primary drivers, but they are essential components of efficient operations. Don’t ignore the tangible benefits associated with reducing data transformation and mitigating the tyranny of content reconciliation.

Knowledge Graph Use Cases

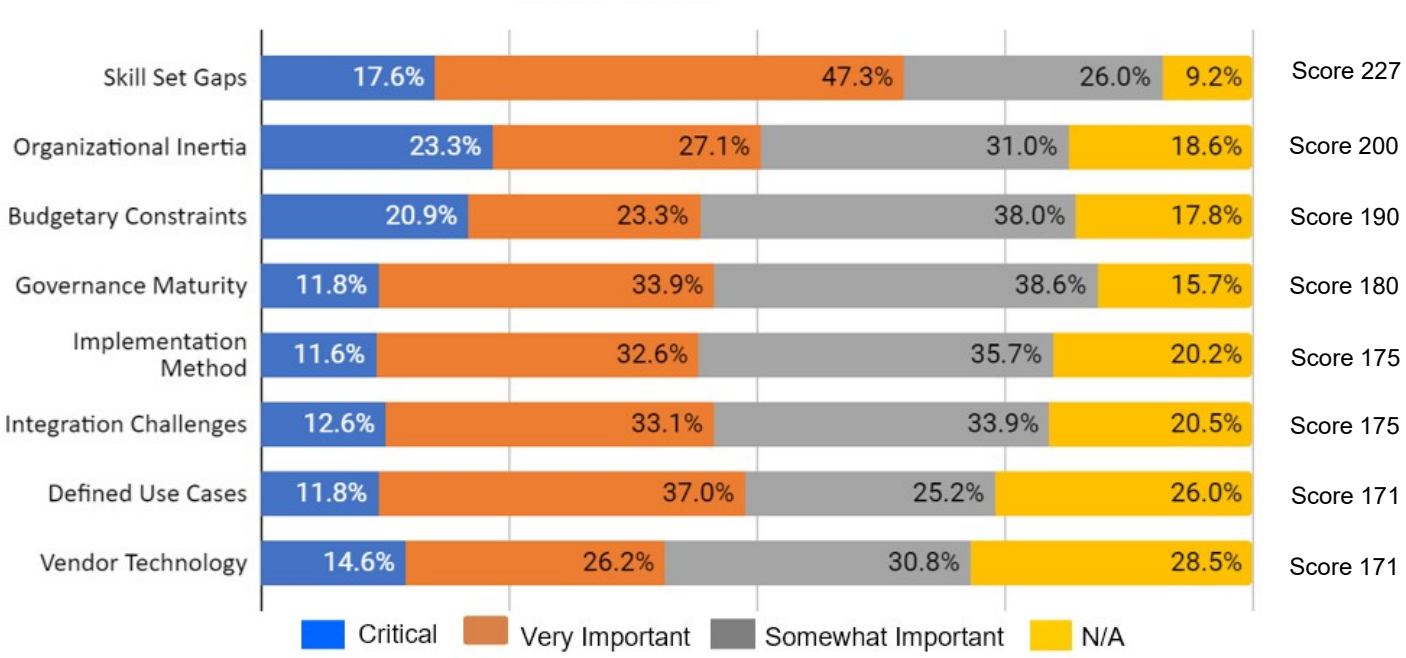


Analysis: This question illustrates the relationship between the adoption goal of "business enablement" and the operational goal of "data integration." The use cases deemed the most critical by responders are those related to establishing a controlled and predictable data environment – including data integration, content harmonization, entity resolution, lineage traceability and business provenance. These are foundational capabilities and prerequisites for achieving business value.

Data interoperability and the goal of establishing a shared language across processes and applications are tailor-made for knowledge graphs and are essential for integrating data into operational environments. Business focus use cases such as data discovery, semantic search, business reporting and cybersecurity represent just under 40% of existing applications but are the keys to driving adoption.

Viewpoint: We anticipate progress coming in three stages. The first stage represents baseline activities associated with the harmonization of data across repositories. The second step results in interim (but essential) deliverables associated with building the data catalog and establishing a connected inventory of people, processes, technology and data. Targeted business value (customer profiling, product engineering, flexible query, integrated views, etc.) will be constructed from these foundational components.

Knowledge Graph Inhibitors



Analysis: Making the transition from conventional data processing environments to knowledge graph is a cultural challenge, not a technical one. This question (combined with our ongoing conversations with the industry) reinforces our conviction that the most important challenges facing the knowledge graph community are those related to crafting the "business narrative." This means overcoming organizational inertia (i.e., convincing executive leadership to lead); defining implementation requirements in concrete business terms; expressing transition expectations using empirical metrics; and having pragmatic discussions on what is really required to cross the knowledge graph minefield.

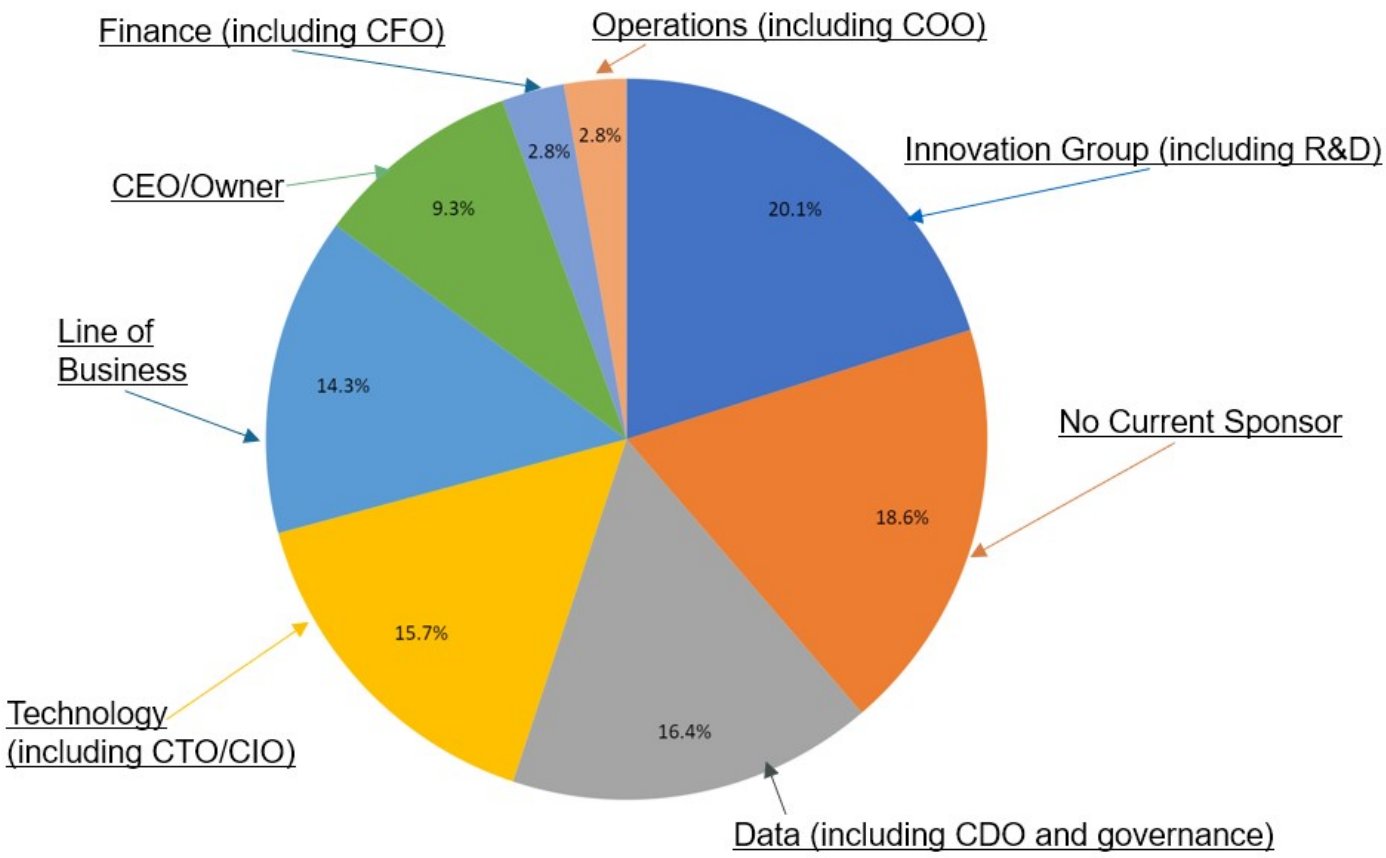
Beyond the business case challenge looms a growing skills gap concern. This is less of a problem for early adopters, as the majority of POCs and pilots are being implemented by small specialist teams. This skills gap issue becomes much more important as we expand adoption across related use cases. The transition from "technology innovation" to "operational deployment" points to the need for education on how technology and governance professionals will make the leap to a new way of operating.

Viewpoint: Overcoming organizational inertia is a serious obstacle. Inhibitors to adoption are certainly not due to technology maturity. Knowledge graph technologies work as advertised but most entities operate in fragmented technology environments. Data is still managed in silos. We transform the meaning of data to fit the requirements of proprietary software. Approaches to software architecture are continually reinvented on an application-by-application basis – and then hard coded into schemas that are rigid and inflexible. Countless time and money is spent reinventing, mapping, moving and transforming data. And as organizations grow and applications multiply, the sheer number of physical elements in systems continues to increase. The more redundant systems the organization has, the more bespoke data elements exist and the more technical debt accrues.

We must learn how to tell a better story. We know that unconnected data using 50-year-old technology is a serious liability. We know that semantic technology was designed specifically for interconnected data. We know that knowledge technology stimulates cross-departmental communication. And we know that it is a prerequisite to achieving the goal of AI-powered applications and intelligent search. We are standing at the new precipice of content interoperability – with a solvability message that doesn't require us to "rip and replace" our existing Infrastructure. It is the organizational concerns that are the dominant inhibitors to knowledge graph adoption.

Knowledge Graph Initiative

Primary Sponsor

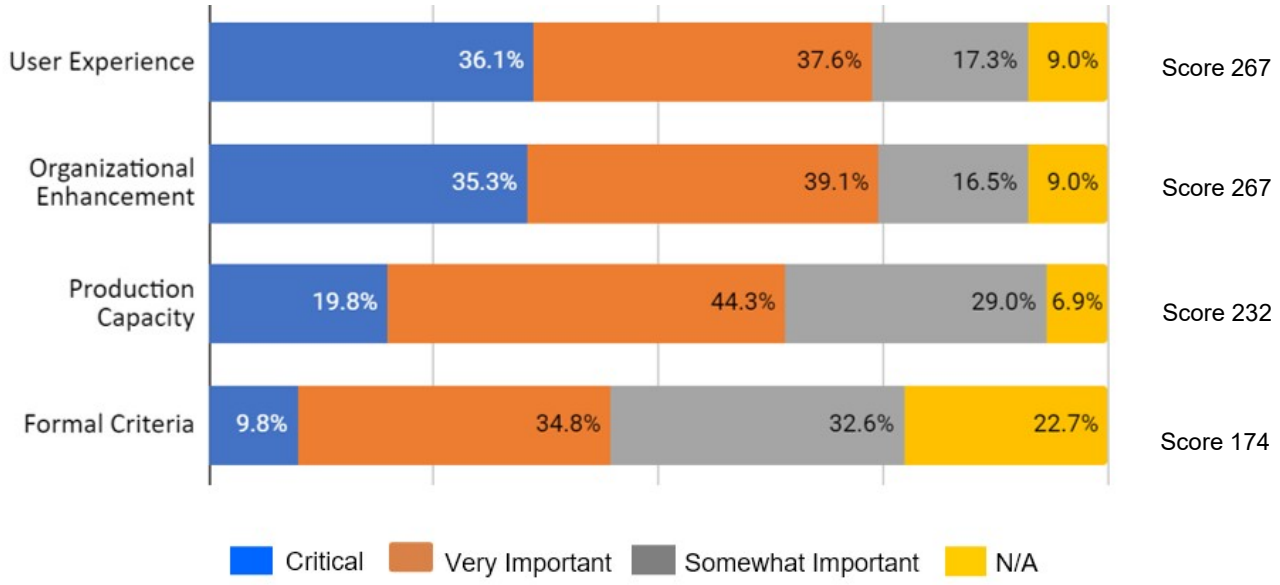


Analysis: Knowledge graph technology is still viewed as an “experiment” by most organizations and has not yet been fully integrated into the business culture of organizations. The principal sponsors (71%) are from innovation groups, research and development, technology or data management.

Viewpoint: We have not yet penetrated the cognitive screen required to overcome organizational inertia and capture the attention of business. Perhaps we should stop talking about how it works. It is not as important as making it work.

Knowledge Graph Initiative

Success Criteria



User Experience – ease of adoption, compelling use case

Organizational Enhancement – competitive advantage and enhanced operational capabilities

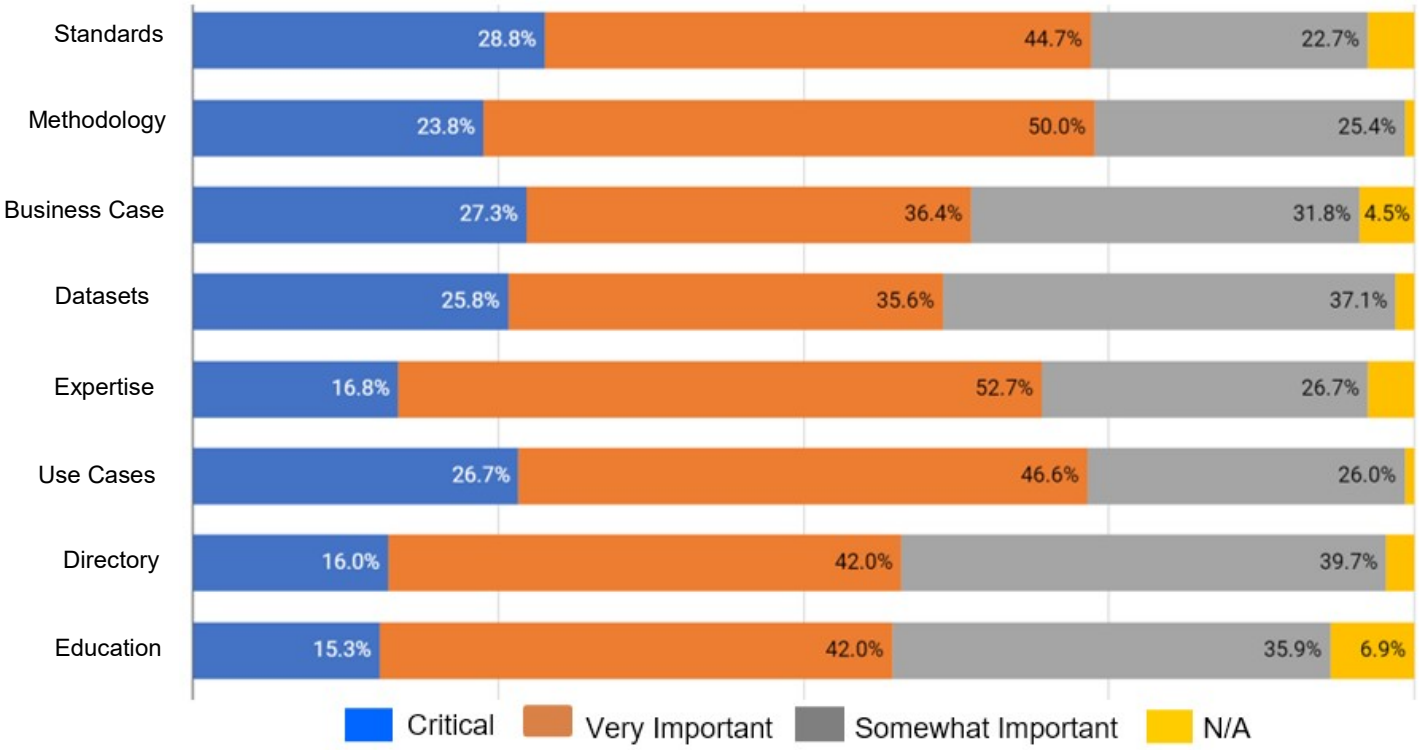
Production Capacity – testing, integration, scalability

Formal Criteria – TCO, ROI, other KPIs

Analysis: The goals are laudable. We are emphasizing ease of adoption, competitive advantage and enhancement of organizational objectives as the preferred methods of measuring success of knowledge graph initiatives.

Viewpoint: The measurement criteria for evaluating the success of these initiatives is fuzzy and hard to demonstrate. The challenge is that formal approaches that provide empirical evidence of return on investment are still desired by business as the way to measure success. We must be thinking about how to translate some of our foundational capabilities (i.e., identity resolution, locking down meaning, structural quality validation and concept reuse) into concrete business terms.

Knowledge Graph Priorities

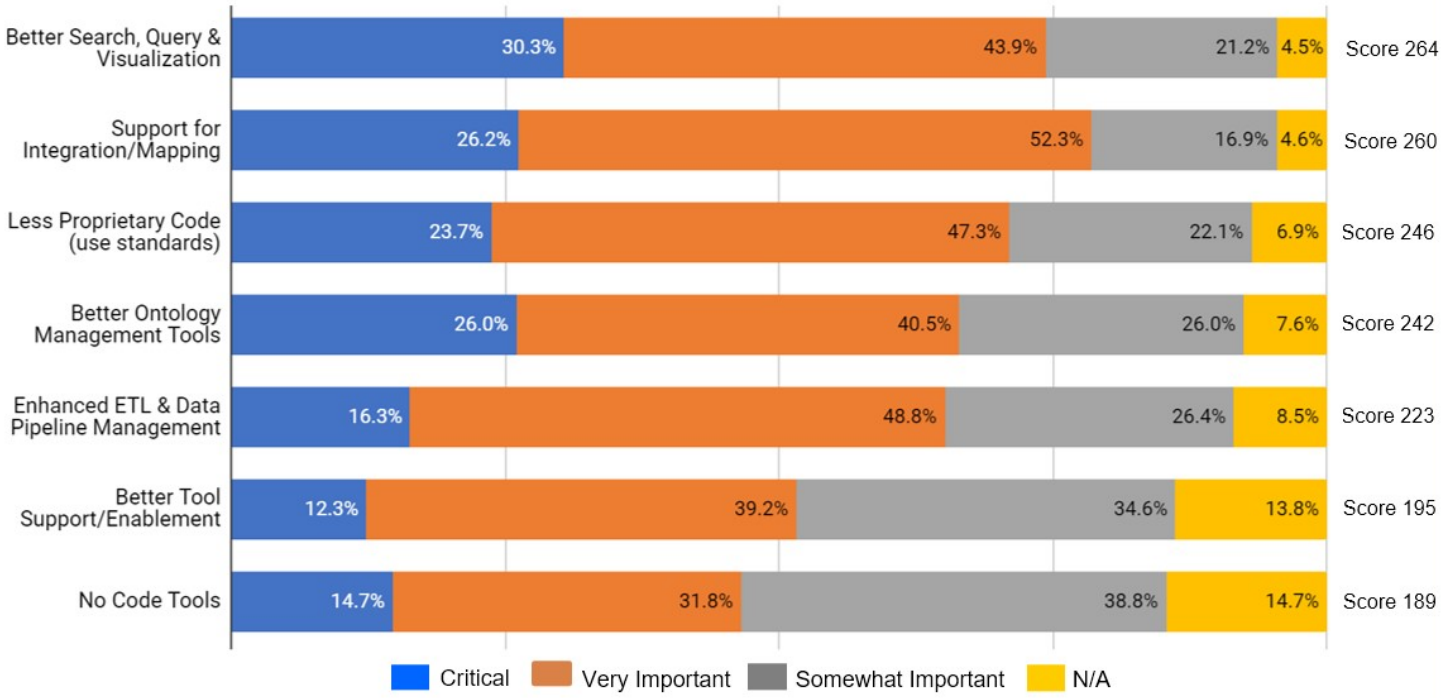


- **Standards** [score 262] **–** Consistency and interoperability of standards
- **Methodology** [score 256] **–** Methodology, best practices and implementation guidance
- **Business Case** [score 246] **–** Business case, ROI research and benchmarking
- **Datasets** [score 245] **–** Curated datasets and shared ontologies
- **Expertise** [score 239] **–** Communities of practice and access to expertise
- **Use Cases** [score 231] **–** Use cases and case studies
- **Directory** [score 225] **–** Directory of tools (open source and vendor specific)
- **Education** [score 217] **–** Education, training and certification programs

Analysis: This question focused on areas of priority for the knowledge graph community. The top priority (when you combine “business case” and “use cases”) is about ensuring appropriate positioning of knowledge graph and semantic standards to business and executive stakeholders. This is about messaging and the importance of crafting a narrative that resonates. There are two other priorities of note. The first relates to interoperability of standards and our ability to integrate knowledge graph capability across platforms. The second relates to the establishment of best practices for implementation and governance-related issues.

Viewpoint: Pull it together – knowledge graph justification using the language of business, standards for interoperability and implementation best practices. A wise agenda for the industry.

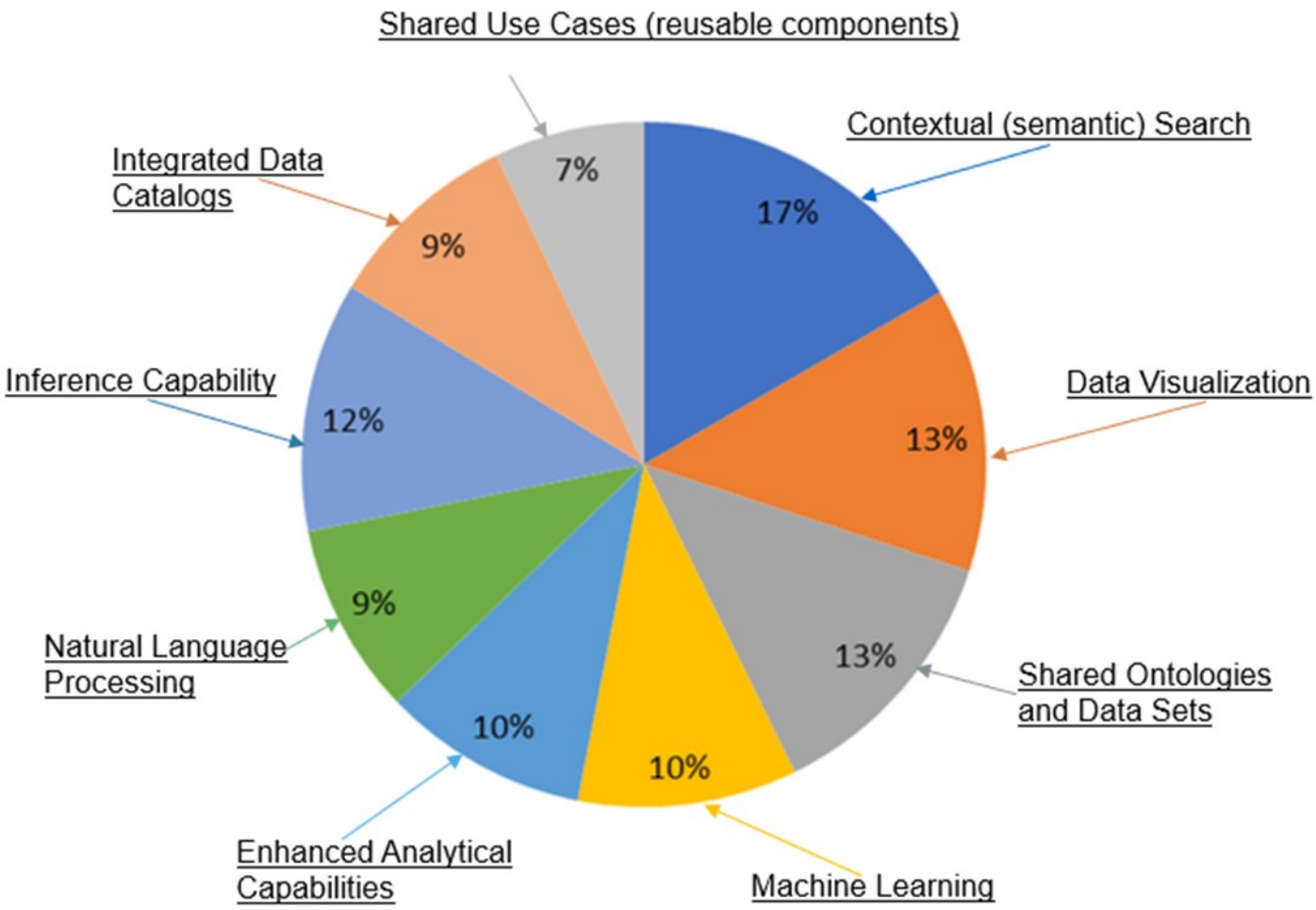
Knowledge Graph Vendors



Analysis: The offerings and capabilities of the knowledge graph vendors are maturing. Concerns about vendors are not viewed as inhibitors to adoption. At the top of the wish list for vendors is contextual search and visualization capability, as well as support for integration and mapping. It is not surprising that a maturing industry seeks support for navigation, enhanced search and information discovery. There is good news on the data integration front, as many are relying on machine learning as one of the most promising pathways to both address the content mapping challenge and reduce ambiguity.

Viewpoint: The bottom line for knowledge graph vendors is about global interoperability across platforms. This means the adoption of standards to reduce friction across the data production process from planning through delivery. Pay attention to the DataOps process to improve collaboration among stakeholders. The focus is on development of logical/physical models, metadata management, ontology development, mapping and production testing. At the end of the day, the goal is to improve time-to-market deployment, reduce failure rates and enhance time of recovery for anything that impacts the quality of a production system.

Future of Knowledge Graph Technology



Analysis: Responses to the question about the future of knowledge technology mirror the enhancements the practitioners would like to see from vendors and suppliers. The top two are about semantic search and data visualization – all designed to help end users get value out of the graph.

There is a significant amount of interest in tools for better pipeline management including machine learning, natural language processing and inference capabilities. Rounding out the list are the notions of shared (industry-level) ontologies and curated datasets. These help reinforce the message about interoperability and support for a common language across all platforms, publishers and consumers.

Four Key Takeaways

Early Stage: Knowledge technologies are still emerging as the pathway out of the (very real) challenges of incongruence, silo operations and structural rigidity.

Investment is Required: A foundational capability is necessary (*i.e., discovery, traceability and integration*) before the delivery of business value. Don't ignore the importance of "data as infrastructure" – it is the driver!

Not Technology: Overcoming organizational inertia is a serious obstacle. Perhaps we should stop talking about how it works. It is not as important as making it work.

Interoperability: The knowledge industry (providers) have an opportunity to expedite adoption by focusing on global interoperability across platforms. Keep up the good work.

Contact Information

Michael Atkin, Managing Director +1.240.602.8390
atkin@content-strategies.com *(Principal Analyst)*

François Scharffe, Co-Founder, Knowledge Graph Conference +1.917.756.9535
francois@knowledgegraph.tech

Thomas Deely, Co-Founder, Knowledge Graph Conference +1.917.385.4927
thomas@knowledgegraph.tech

Acknowledgements

Dennis Wisnosky, Chairman, Enterprise Knowledge Graph Foundation

Claude Robotham, Communications and Research Associate, Knowledge Graph Conference, clauderobo@outlook.com

Paco Nathan, Managing Partner, Derwen, Inc.