

How Advances in AI Changed the Game for Search and Knowledge Discovery



Hayley Sutherland
Research Manager,
Conversational AI and Knowledge Discovery, IDC

Table of Contents



CLICK BELOW TO NAVIGATE TO EACH SECTION IN THIS DOCUMENT.

In This InfoBrief	3	The Next Phase: Harnessing the Potential of GenAI	10
The Rise of GenAI: Impacts on Search and Knowledge Discovery	4	The New AI Imperative: Search’s Time to Shine	11
Legacy Search Keeps Organizations Mired in Challenges	6	Mind the Gap: GenAI Urgency Versus Data Readiness	12
Sea Changes in Search: 2019–2024	7	From Search to Search-Powered AI	13
The Importance of GenAI for Synthesized Insights	8	Essential Guidance	15
The Benefits of Modern Search	9	Appendix: Supplemental Data	16
		About the IDC Analyst	17
		Message from the Sponsor	18

In This InfoBrief

Generative AI (GenAI) is changing how forward-looking organizations approach search, knowledge management, and other forms of knowledge discovery.

Search is foundational for helping organizations discover, analyze, and utilize key data. However, an increasing deluge of data means legacy search systems can struggle to help business users quickly find what they need and avoid wasting time.

Modern search systems have made great leaps in the accuracy, relevancy, and usefulness of results thanks to AI-based capabilities such as natural language processing, machine learning-based relevancy, and vector (aka semantic) search. The combination of GenAI and search technologies has the potential to take knowledge discovery to the next level, allowing organizations to finally unlock the value of “hidden” or unanalyzed data.

This InfoBrief uses IDC's *Search/Knowledge Discovery Survey* (2019 and 2023), historical *Value of Hidden Information Survey* (2013 and 2015), and *Future Enterprise Resiliency and Spending Survey* (2023 and 2024) data to demonstrate how the combined power of AI and search technology can help organizations maximize productivity and gain competitive advantage in today's highly competitive markets.



The Rise of GenAI: Impacts on Search and Knowledge Discovery

GenAI triggered the opening of a new chapter in the digital business era, with organizations racing to make business use of promising advances in AI.

GenAI can rapidly glean insights and generate summaries from vast volumes of data. It also has the potential to drastically **reduce the time and long-term costs** associated with developing solutions across a wide range of use cases, increasing productivity and efficiency within organizations.

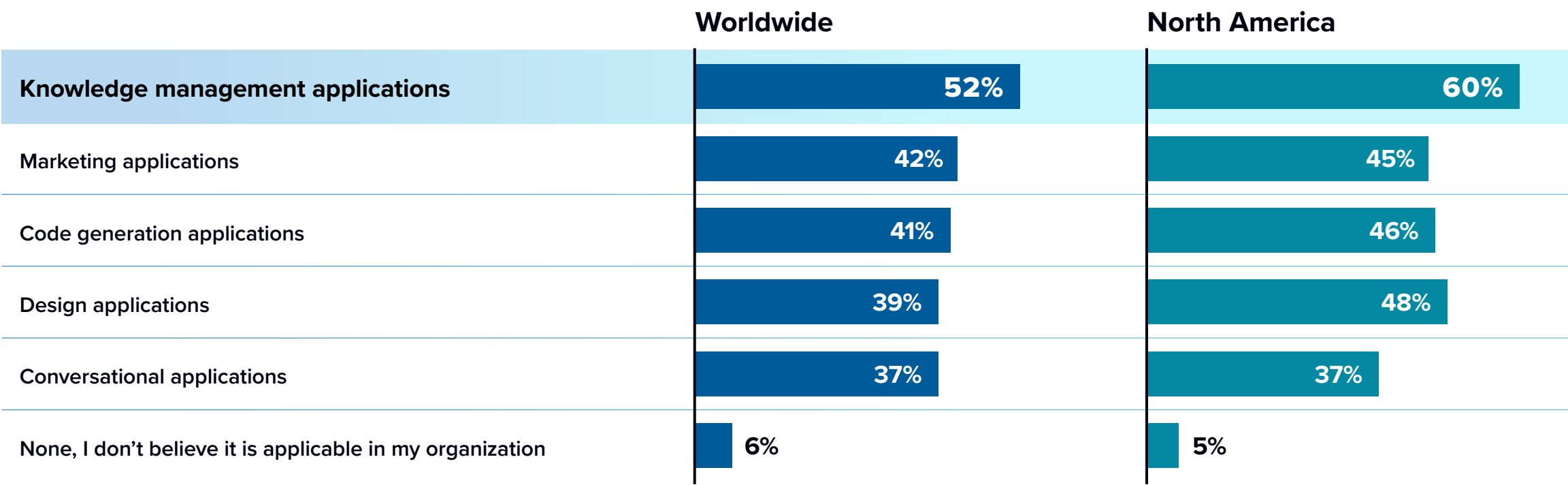
Some of the most promising initial use cases for this technology relate to **knowledge management and knowledge discovery.** This includes using GenAI to analyze, classify, and categorize unstructured and structured business knowledge, answer questions, and provide insights and summaries for users. At the same time, search technologies have become critical for connecting large language models (LLMs) to valuable business data, including real-time data.

AI will be everywhere in knowledge management and knowledge discovery applications moving forward, with GenAI becoming a critical component of advanced, modern knowledge solutions that enable organizations to derive competitive advantage from data.



The Rise of GenAI: Impacts on Search and Knowledge Discovery (continued)

What GenAI use cases do you anticipate having the most promise for your organization?



n = 952 (Worldwide), n = 370 (North America); Source: IDC's *Future Enterprise Resiliency & Spending Survey, Wave 2*, March 2023 | For an accessible version of the data on this page, see [Supplemental Data](#) in the Appendix.

Legacy Search Keeps Organizations Mired in Challenges

Legacy search engines that have not significantly advanced in the last five years have struggled to keep up with the increasing volume and variety of organizational data.

These legacy engines typically use traditional keyword search and brittle, rules-based systems instead of adaptive, intelligent, semantic, and hybrid search. As a result, organizations using these tools must cope with poor relevancy ranking, outdated or broken query understanding, and basic findability challenges.

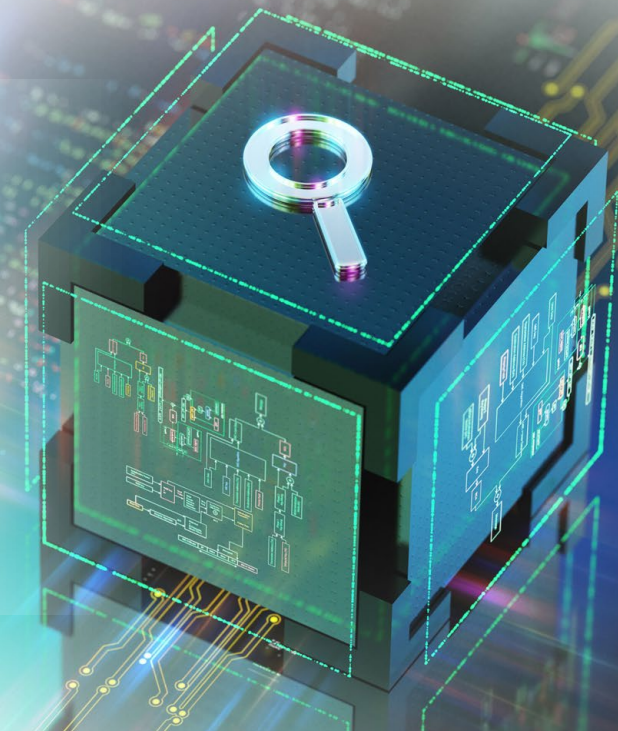
Examples of search activities where “time wasted” numbers got worse over time (as of 2019) include:

- ▶ **Recreating content that can’t be found:**
from ~3 hours in 2013 and 2015 to ~5 in 2019
- ▶ **Translating content:**
from ~1.7 hours in 2013 and 2015 to ~4 in 2019
- ▶ **Searching for info and never finding it:**
better in 2019 versus 2013, but worse than 2015

Average number of hours
wasted searching for,
but not finding, information:

2013	2015	2019
5.1	4.1	4.7
HOURS	HOURS	HOURS
per worker per week		

Sources: IDC's *Value of Hidden Information Study*, 2013 and 2015; IDC's *AI-Enabled Search Survey*, 2019



Sea Changes in Search: 2019–2024

AI has brought markedly better capabilities for searching, translating, and combining information in the last few years. **Before ChatGPT grabbed headlines, AI was already bringing new capabilities to search applications.**

AI advances have brought increasingly sophisticated **natural language understanding** to search.

Machine learning-based relevancy ranking became commonplace in the past few years.

Semantic search, aka vector search or similarity search, has been replacing or augmenting traditional keyword search. Semantic search leverages AI algorithms and vector databases to identify related terms and concepts for a given query, resulting in better recall overall and reducing “no results found” system responses.

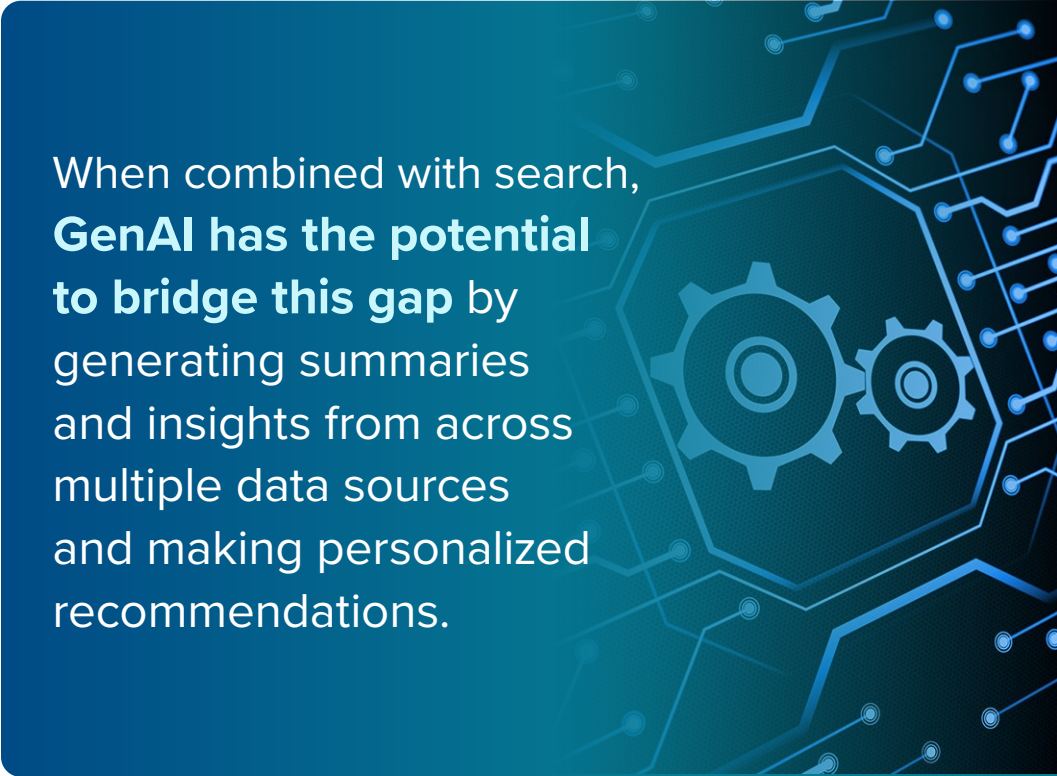
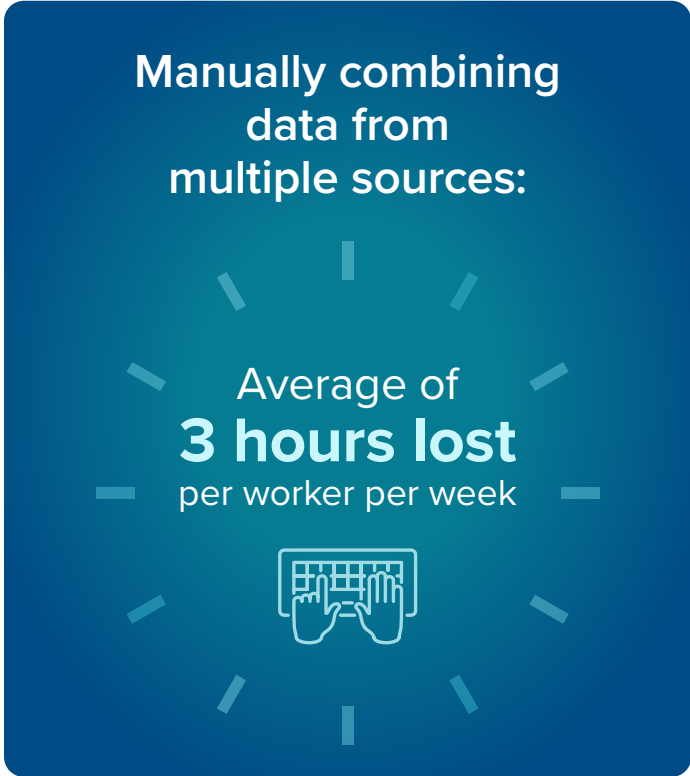


The result: productivity improvements of **over 12 hours per week** just from 2019 to 2023

n = 522; Source: IDC's *North America Knowledge Discovery Survey*, February 2023

The Importance of GenAI for Synthesized Insights

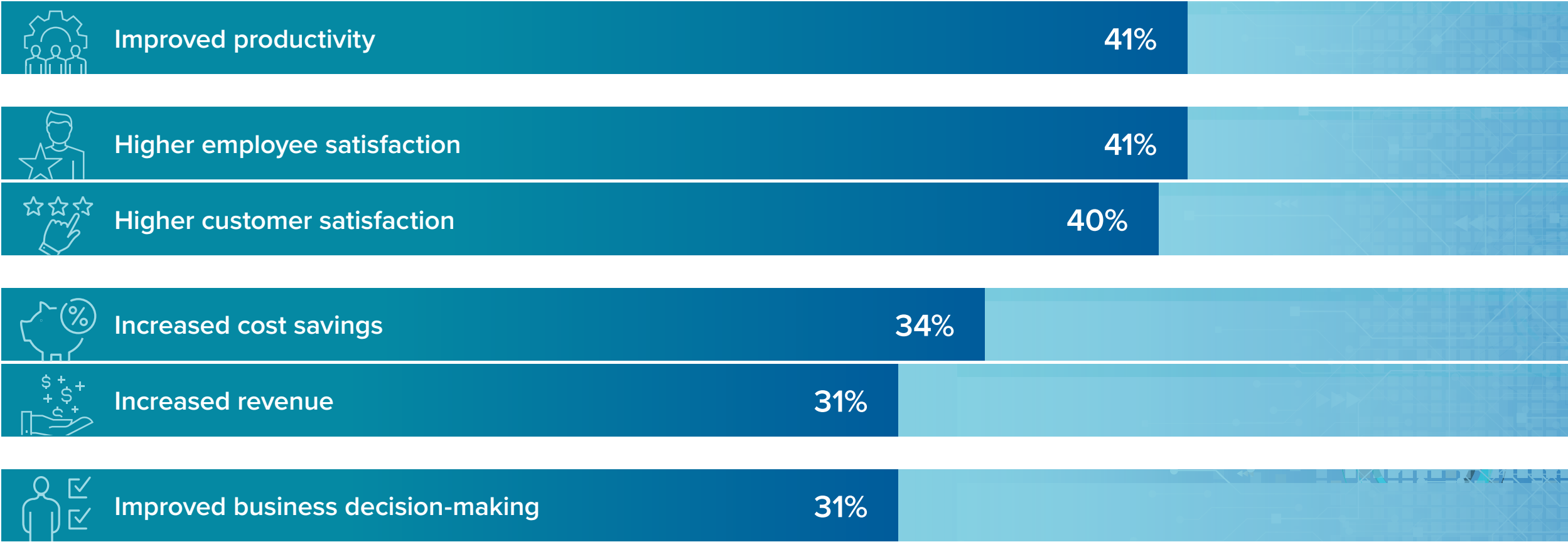
Despite improvements over the years, the process of searching across and combining data from multiple sources still represents the biggest waste of time for knowledge workers.



n = 522; Source: IDC's North America Knowledge Discovery Survey, February 2023

The Benefits of Modern Search

Companies that implement modern search/knowledge discovery systems see a wide range of operational and tactical benefits.



n = 522; Source: IDC's North America Knowledge Discovery Survey, February 2023

The Next Phase: Harnessing the Potential of GenAI

GenAI provides greater value when connected to business data.



Over 80% of organizations that IDC surveyed were either **exploring potential use cases or already investing** in GenAI technologies in 2023.



83% of IT leaders believe that **GenAI models that leverage their own business' data** will give them a significant advantage over competitors.

This will allow GenAI models to respond to specific customer queries, provide guided shopping experiences, make advisory business predictions and personalized recommendations, synthesize and surface new insights, and provide organizations with a clearer view of their knowledge assets.

Note: Data weighted by IT spend. n = 1,363; Source: IDC's *GenAI ARC Survey*, August 2023

The New AI Imperative: Search's Time to Shine

While AI has continued to improve search technology, on the flip side, search technology has also proven key for making the latest AI advances work for business use cases.



The rise of LLMs and GenAI is driving a surge of demand for **capabilities such as natural language/conversational search and question-answering**.

Business use of LLMs is also **fueling the need to combine search systems with conversational AI and GenAI systems** to provide accurate, relevant answers based on real-world business data.

Vector search and related vector databases have become increasingly popular for retrieving the most relevant information from proprietary business sources and **feeding this info into a LLM to help augment AI-generated responses**. This is a process called retrieval-augmented generation (RAG).

Compared with fine-tuning, which requires retraining an AI model, **RAG can be a more cost-effective and less time-consuming method of supplementing LLMs** with specific and/or proprietary data.

Mind the Gap: GenAI Urgency Versus Data Readiness

For enterprise GenAI use cases, LLMs must be connected to enterprise data — but is your data ready for GenAI?



In January 2024,
only 24%
of organizations believed
that their resources
were **extremely prepared
for GenAI.**

Only one-third or fewer of those organizations:

- ▶ Could ensure that **high-quality data** was being used with GenAI models
- ▶ Understood where **personal and corporate IP sensitivities** existed in data being shared with GenAI models
- ▶ Could **trace what the models were doing** with the data being sent through prompts

n = 881; Source: IDC's *Future Enterprise Resiliency & Spending Survey, Wave 1*, January 2024

From Search to Search-Powered AI

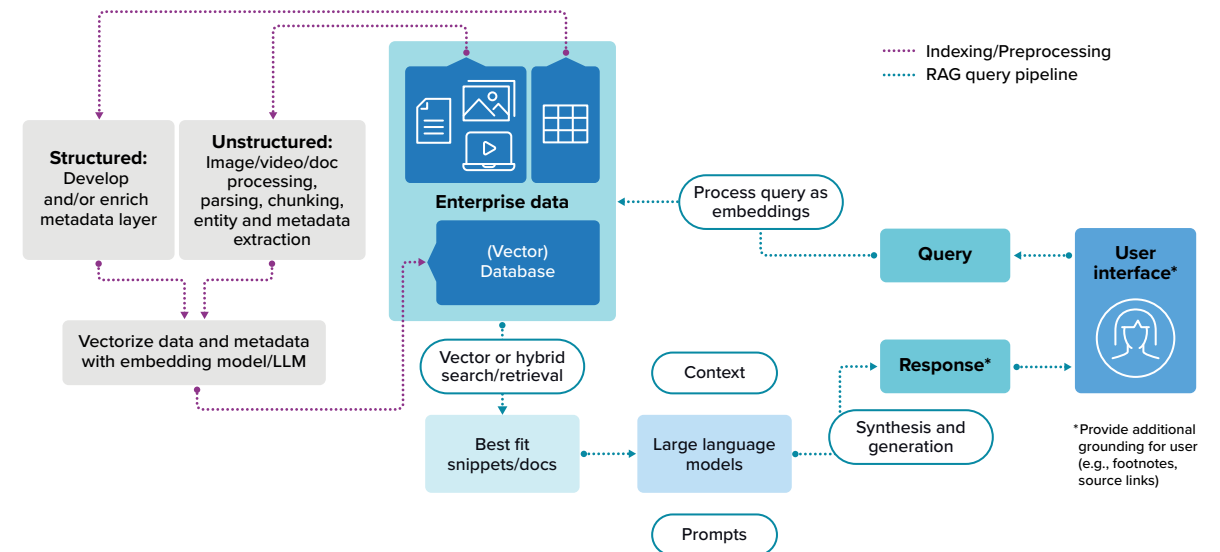
Search technologies bridge the gap between GenAI and enterprise data.

RAG makes enterprise adoption of GenAI more feasible and practical by augmenting LLMs with enterprise data using search techniques.

Without RAG or other grounding techniques, LLMs lack real-time and domain-specific context and have been shown to frequently produce “hallucinations.”

The quality of the search/retrieval step is crucial to the quality of GenAI output in RAG. Vector or hybrid search (vector + keyword) is the predominant method for doing this with the greatest accuracy and relevancy.

Retrieval-Augmented Generation (RAG): Vector Search Augments LLM Responses with Specific Data

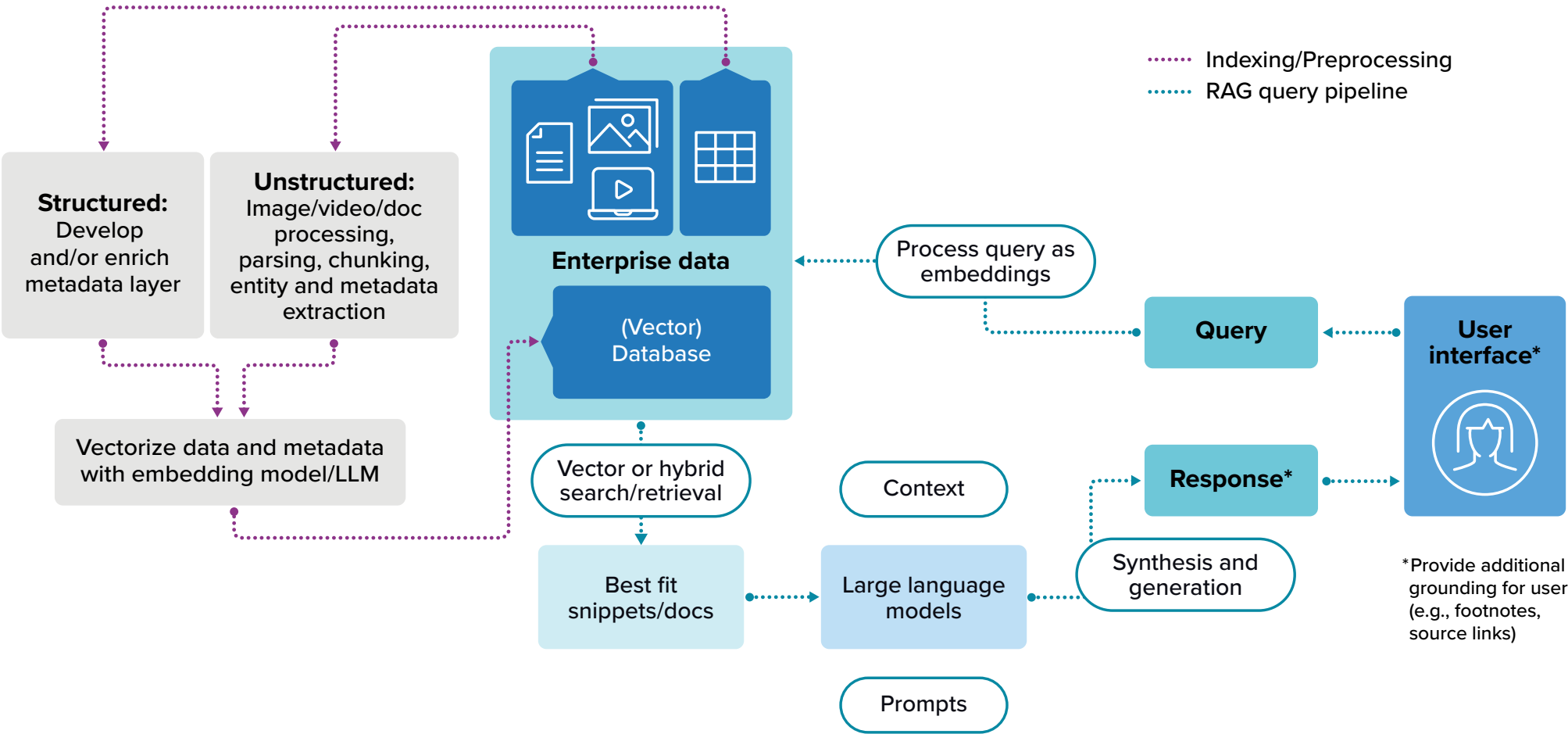


For a larger view of this diagram, see [next page](#).

Source: IDC TechBrief: Retrieval-Augmented Generation, April 2024, Doc #US51676224

From Search to Search-Powered AI (continued)

Retrieval-Augmented Generation (RAG): Vector Search Augments LLM Responses with Specific Data



Source: IDC TechBrief: Retrieval-Augmented Generation, April 2024, Doc #US51676224

Essential Guidance

For organizations that want to leverage their enterprise data to power GenAI use cases, IDC recommends the following:

Ensure that the organization is leveraging strong search technology, with high levels of accuracy and relevancy, for the retrieval step of RAG. Aspects to look for include hybrid search (keyword and semantic/vector), automated reranking, and low-code/no-code tools that make testing and tuning easy for a wide variety of users. This step is crucial for ensuring that LLMs provide the most relevant, useful, and actionable summaries and answers.

Assess the accuracy and freshness of data sources, and consider what tools will be used to connect, filter, or ingest data into the pipeline. Ensure that data governance and business rules, such as access permissions, are not lost in the process and that the system has strong security guardrails.

Determine what types of AI are best suited to different use cases; GenAI should be applied strategically to ensure that its usage is feasible, valuable, and responsible. If needed, select a provider with the experience to assist in prioritizing use cases and AI usage.

Look for a vendor that can assist with some or all of the steps required to connect enterprise data to LLMs, including parsing, chunking, embedding, storing, and using vector search to retrieve key information to feed to the LLM.

Appendix: Supplemental Data

The table in this appendix provides an accessible version of the data for the complex figure in this document. Click “Return to original figure” below this table to get back to the original data figure.

SUPPLEMENTAL DATA FROM PAGE 5

What GenAI use cases do you anticipate having the most promise for your organization?

	Worldwide	North America
Knowledge management applications	52%	60%
Marketing applications	42%	45%
Code generation applications	41%	46%
Design applications	39%	48%
Conversational applications	37%	37%
None, I don't believe it is applicable in my organization	6%	5%

n = 952 (Worldwide), n = 370 (North America); Source: IDC's *Future Enterprise Resiliency & Spending Survey*, Wave 2, March 2023

[Return to original figure](#)

About the IDC Analyst

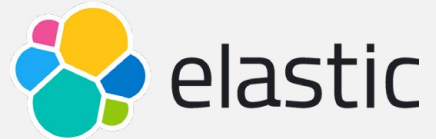


Hayley Sutherland
Research Manager,
Conversational AI and Knowledge Discovery, IDC

Hayley Sutherland is a research manager for Conversational AI and Knowledge Discovery within IDC’s software market research and advisory group. Her core research coverage includes conversational AI and search, with a focus on AI software development tools and techniques for chatbots and digital assistants, speech AI and text AI, machine translation, embedded knowledge graph creation, intelligent knowledge discovery, and affective computing (also known as emotion AI).

[More about Hayley Sutherland](#)

Message from the Sponsor



Elastic's complete, cloud-based solutions for search, observability, and security are built on the Elastic Search Platform, used by more than 50% of the Fortune 500. By taking advantage of all structured and unstructured data and securing and protecting private information, Elastic helps organizations use the precision of search to deliver on the promise of AI.

[Explore Elastic AI solutions](#)

IDC Custom Solutions

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. This IDC material is licensed for external use and in no way does the use or publication of IDC research indicate IDC's endorsement of the sponsor's or licensee's products or strategies.



IDC Research, Inc.
140 Kendrick Street, Building B, Needham, MA 02494, USA
T +1 508 872 8200

[idc.com](https://www.idc.com)

[in @idc](https://www.linkedin.com/company/idc)

[X @idc](https://twitter.com/idc)

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives.

©2024 IDC. Reproduction is forbidden unless authorized. All rights reserved. [CCPA](#)