COHESITY

Exploring the benefits of secure, AI-ready data

It’s time to bring your data into the future.

INTRODUCTION

In the past few years, we've heard all about AI and its promised benefits, largely around increasing productivity and replacing menial tasks. But the power of AI is growing rapidly and organizations are already using it in a wide range of industries, including healthcare, finance, transportation, and technology, among others. In 2022, the global AI market was valued at $428 billion. It's projected to surge to $2 trillion by 2030. If this seems outrageous, consider the growth of Open AI and ChatGPT in under a year. According to OpenAI, ChatGPT gained over one million users in just five days after its November 2022 launch. AI is here to stay and it's no longer a future technology, as tech companies are announcing new AI products and integrations every day.

What are the benefits of AI to large global enterprises?

At Cohesity, we're looking at AI through a data security and data management lens. And that's really where we see AI as a game changer. For example, we can secure your data with AI-powered early threat detection so you can get more value from your backed up data by exposing zero cost clones to your AI and ML workflows—whether your data sits in a data center, or a public or private cloud.

AI can help you differentiate your business, driving growth and improving customer outcomes.

In this eBook, we dive into some emerging AI use cases and explore how we can make your data AI ready while keeping it secure.

DATA AND AI

Data is an essential aspect of AI. It's the fuel that provides context to AI algorithms and vectors. But AI models are only as good as the data they're trained on, and the quality of this training data can significantly impact the validity, effectiveness, and power of the model. If the data available to AI models is biased or inaccurate, the model may produce biased or inaccurate results.

AI models also require ongoing access to new data to continue learning and improving over time. This is why data is so important to AI's development and deployment. Organizations that invest in collecting, storing, and analyzing high-quality data are better positioned to use AI's power to gain a competitive advantage.

In today's modern, distributed architecture, though, collecting, collating, and leveraging data from workflows across an organization's data estate can be complex. Organizations are generally running infrastructure in a variety of locations, spanning private data centers, single or multiple clouds, SaaS applications hosted by other organizations, and edge locations like stores, IoT devices, and more.

They're routinely storing petabytes (or more) of data without classifying, indexing, or tracking it. This unstructured data is often referred to as "dark data," which Gartner defines as "the information assets organizations collect, process and store during regular business activities, but generally fail to use for other purposes (for example, analytics, business relationships and direct monetizing)."

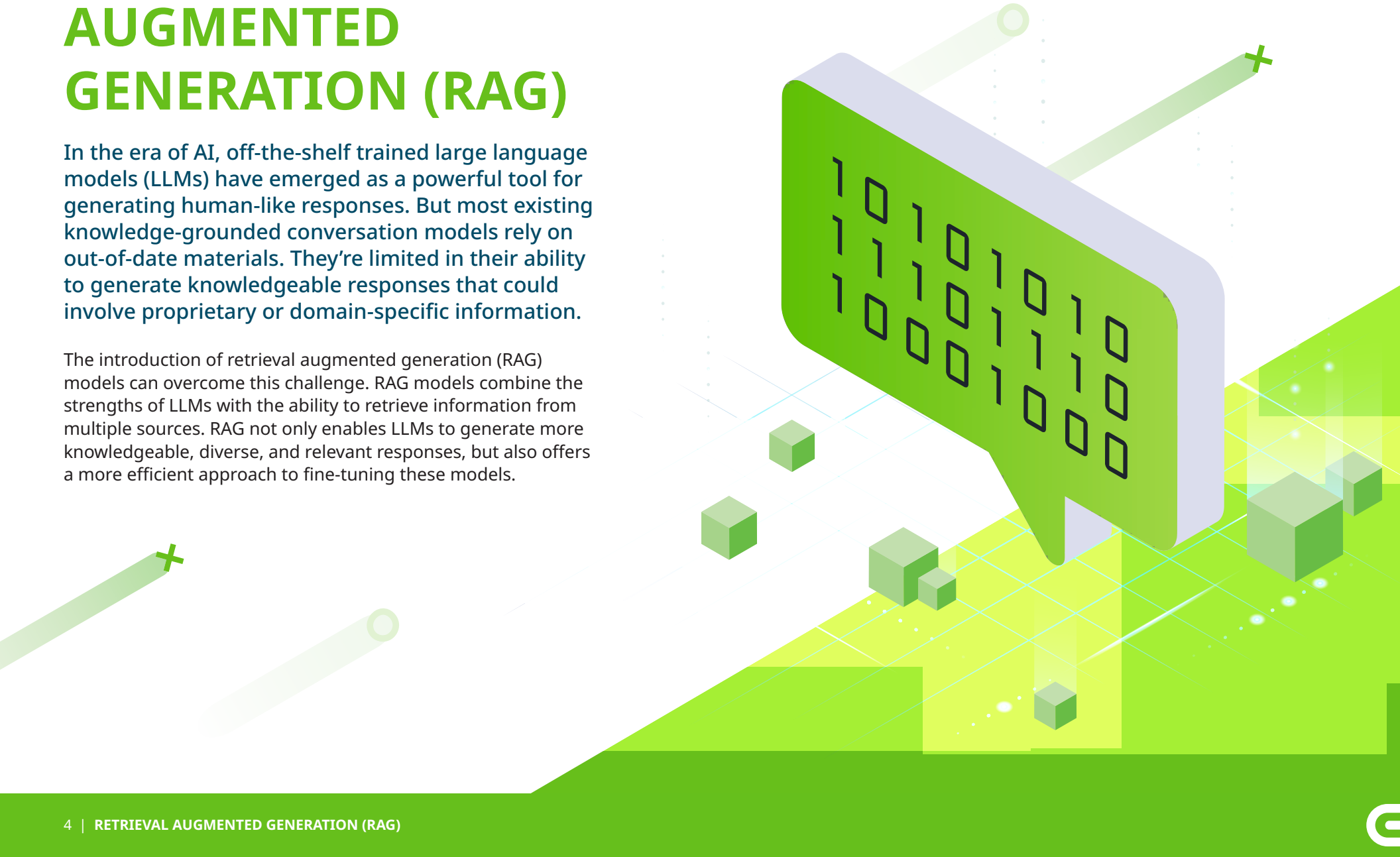
By most analyst estimates, unstructured data makes up **80-90%** of total data.

This dark data represents a missed opportunity. Organizations can't gain insights and make informed decisions, dramatically reduce their data costs, or secure and protect this data—because they don't even know it exists.

RETRIEVAL AUGMENTED GENERATION (RAG)

In the era of AI, off-the-shelf trained large language models (LLMs) have emerged as a powerful tool for generating human-like responses. But most existing knowledge-grounded conversation models rely on out-of-date materials. They're limited in their ability to generate knowledgeable responses that could involve proprietary or domain-specific information.

The introduction of retrieval augmented generation (RAG) models can overcome this challenge. RAG models combine the strengths of LLMs with the ability to retrieve information from multiple sources. RAG not only enables LLMs to generate more knowledgeable, diverse, and relevant responses, but also offers a more efficient approach to fine-tuning these models.



RETRIEVAL AUGMENTED GENERATION (CONTINUED)

At Cohesity, we're working on providing robust and domain-specific context to RAG-driven AI systems through our patented SnapTree and SpanFS architectures. By leveraging this robust file system, we're making our platform 'AI ready' for RAG-assisted LLMs through an on-demand index of embeddings that will be provided just-in-time to the AI application requesting the data. We also secure the data through our role-based access control (RBAC) models.

Our patent-pending retrieval-augmented response generation technology currently under development accepts a user- or machine-driven input—such as a question, or a query. That input is then used to filter the petabytes of an enterprise's backup data to filter down to a smaller subset of data. It then selects representations from within those documents or objects that are most relevant to the user or machine query. That result is packaged, along with the original query, to the LLM to provide a context-aware answer. This innovative approach ensures that the generated responses are relevant to the enterprise's domain-specific content.

Innovative RAG-driven AI systems like ours will present a unique opportunity for technology and business executives to leverage the power of data-driven insights and enhance the quality of conversations across various platforms. Harnessing the power of our data protection and data management capabilities, enhanced by AI, helps organizations unlock new levels of efficiency, innovation, and growth.

Of course, before you do anything with your data, you need to secure it. Let's look at the first of two use cases: data security.

USE CASE 1: USING AI TO BOOST DATA SECURITY

Anomaly detection with AI

We help customers back up their entire data estate and unlock new ways to protect and manage data, in addition to improving cyber resilience with data isolation, threat detection, and data classification. While we already provide deep insights and analytics that improve security posture—thanks to our unique distributed file system—soon we'll fully leverage the same data we're already securing and managing using AI models. Our new collaboration with Azure OpenAI makes this possible.

We recently previewed how Azure OpenAI may be integrated into the Cohesity platform to streamline anomaly detection, provide human-readable analyses of threats, and help organizations reduce recovery time. We demonstrated how we could take our foundation of log and system data, combined with insights built into our Cohesity DataHawk threat intelligence solution, to use AI to query all data and generate interactive reports for CISOs and practitioners alike.

Using Azure OpenAI, we generated an Insights Summary that found objects on virtual machines that could potentially be affected by ransomware. We used this information to explore how a CISO might assess business impact, and how practitioners could streamline responses.

During this demonstration, the integrated AI uncovered numerous affected, highly sensitive files showing behavior or changes associated with a high confidence rating for anomalies. If a CISO were to ask what files were impacted on the VMs, the executive-level summary output could provide a near real-time assessment of the risk profile of these anomalies. Immediately, the CISO could assess business risk while team members in the SOC could get details to stop, mitigate, and respond to the threat.

From there, you could ask for more details on the impacted systems, and the integrated AI could offer human-readable impact analyses. Using comprehensive metadata from the backup and recovery files, you could view contextual threat response information.

You can imagine how quickly cross-functional teams could use these conversational analyses to assess risk, align on impact, and kick off an action plan before attackers could further disrupt the business.

USE CASE 2: GETTING MORE FROM STORED DATA

So how do you gain even more insights from all this data you're securing? We spent years building AI into several of our products to help detect threats, classify large volumes of unstructured data, and protect critical data and workflows. The Cohesity Data Cloud, our modern data security and management platform, is unique in that it is "AI ready." It's architected to be easily searchable, with granular access controls. With global search, you can search across multiple workloads and histories of snapshots. Future integrations with AI and LLMs will be able to quickly answer critical business questions, while ensuring that only the right people see responses regarding the data they have access to.

In the same way backup data is stored and can be searched for threat analysis, it's also AI ready and contains metadata that can be used with LLMs. When a person asks questions about the data through the LLM, the model provides human-readable responses. Using authoritative data sources backed up on Cohesity can help ensure more accurate, actionable responses to user or machine queries.

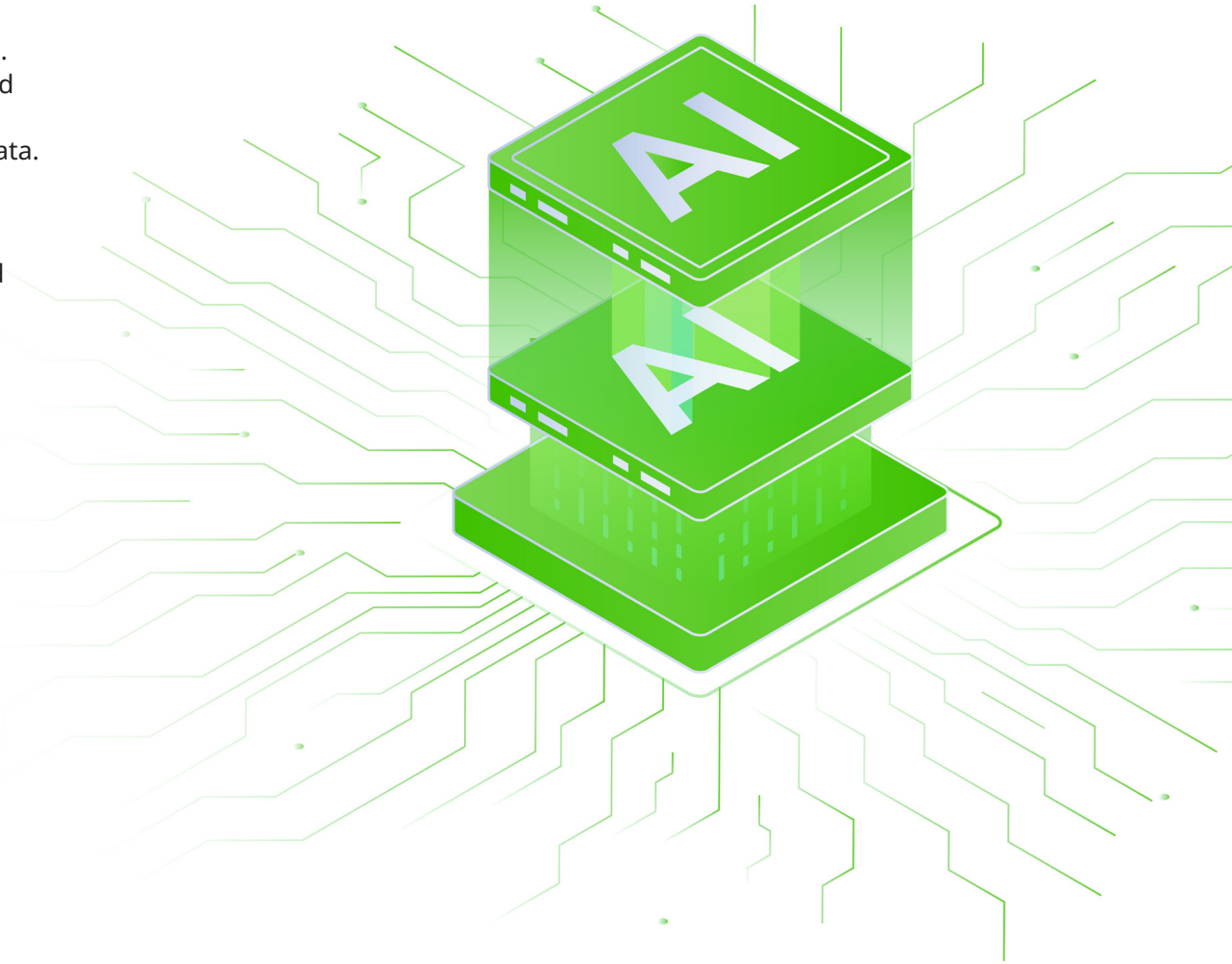
Since we index all backup data, APIs will return context-aware responses in a highly performant way that doesn't use up too much compute power.

GETTING MORE FROM STORED DATA
(CONTINUED)

Indexing and instant search with Cohesity

We'll also take a differentiated approach to storing and indexing massive amounts of data, enabling organizations like yours to instantly search across your entire data estate. With Cohesity instant search, you can quickly search for and locate specific data and find the information you need. You don't have to manually search through large amounts of data.

Think of it this way: Say you're a lawyer and you want to research information from a case your firm litigated years ago. You could pull out boxes and dig through them to find the brief in question, like a needle in a haystack. Soon, you could simply use an AI chatbot to query the data that already lives on Cohesity.



With our holistic view of data over time, you can search decades-old data, instantly, and see different variations of it during different periods. Other companies store data without indexing file and object metadata and have limited history available (only days) for searching or changes.

Thanks to the Cohesity Data Cloud, your data is AI ready.

GETTING MORE FROM STORED DATA
(CONTINUED)

The Cohesity Data Cloud offers:

- **Data aggregation and unification:** We aggregate data from different sources and data types, including on-premises, cloud, and edge locations. This makes it easier to analyze and grant secure access to data for AI applications and identify patterns, trends, and anomalies that may not be visible in siloed data, while also dramatically reducing or eliminating dark data. Using backup data that's indexed and aggregated allows you to leverage AI in a highly performant way.

- **Data optimization:** We efficiently deduplicate and store data in compact structures which can be fortified with appropriate metadata that makes search more robust.

- **Data protection:** We provide enterprise-grade backup, recovery, and disaster recovery capabilities—including a way to isolate data in a virtual air-gapped environment so a known clean copy can be retrieved and restored in the event of a cyberattack. Organizations can get back up and running quickly using instant mass restore (IMR) to instantly recover thousands of VMs, databases, files and other data. This data is available, resilient, and recoverable when needed—which is critical for AI applications that rely on large volumes of data.

- **Data security:** Our comprehensive data security proactively detects threats within data to identify anomalies, like possible malware, using security threat feeds. Data classification can identify where sensitive and critical data sits, and cyber vaulting ensures if an attack happens the blast radius is mitigated. AI applications continue running to help ensure that business operations don't stop, even in the event of an attack or disaster.

- **Data access:** Our granular role-based access controls (RBAC) for backup data helps prevent users from accessing data they don't have permissions for, like sensitive data (patient data/PII, trade secrets, financials, and more). This approach applies to AI, where the AI model only queries data and provides responses that align to users' permissions.

Using the power of AI, we offer countless opportunities for organizations like yours to unlock the power of your **entire data picture** over time. You'll have AI-ready data that's **resilient and available** when needed.

CONCLUSION

AI is poised to be the next big disruptor in today's macro business environment—from streamlining operational processes to transforming customer experience via new data insights-driven services. And that's just a start. At Cohesity, we're focusing on the power of AI to improve data protection and security via advanced threat and anomaly detection. We create AI-powered backups that are fully hydrated and ready for nearly instant zero cost cloning, and safe from harm, too—turning data backups into treasure troves of data insights.

Glossary:

Retrieval augmented generation (RAG) is a NLP technique that combines the benefits of retrieval-based and generative-based approaches to improve the quality of text generation tasks, such as question-answering, summarization, and conversational AI.

Generative AI uses algorithms to generate new content (written content, images, video, audio, and computer code, etc...) based on user input. Unlike earlier versions of AI, generative AI can create new content, like news articles, poetry, or cyber threat analyses presented in a conversational UI.

Conversational AI simulates human conversations through natural language processing (NLP) and related techniques such as transformers. It can interpret user input and generate appropriate responses based on an understanding of the user's intent.

Cognitive AI aims to replicate human cognitive abilities such as perception, reasoning, problem-solving, and decision-making. It uses ML, NLP, and related techniques to create intelligent systems that can learn and improve based on interactions with users and the environment.

Retrieval AI involves searching and retrieving information from large datasets. It uses techniques such as NLP and ML algorithms to understand the intent of a user's query and retrieve the most relevant results from a database.

6100021-001 EN 5-2023