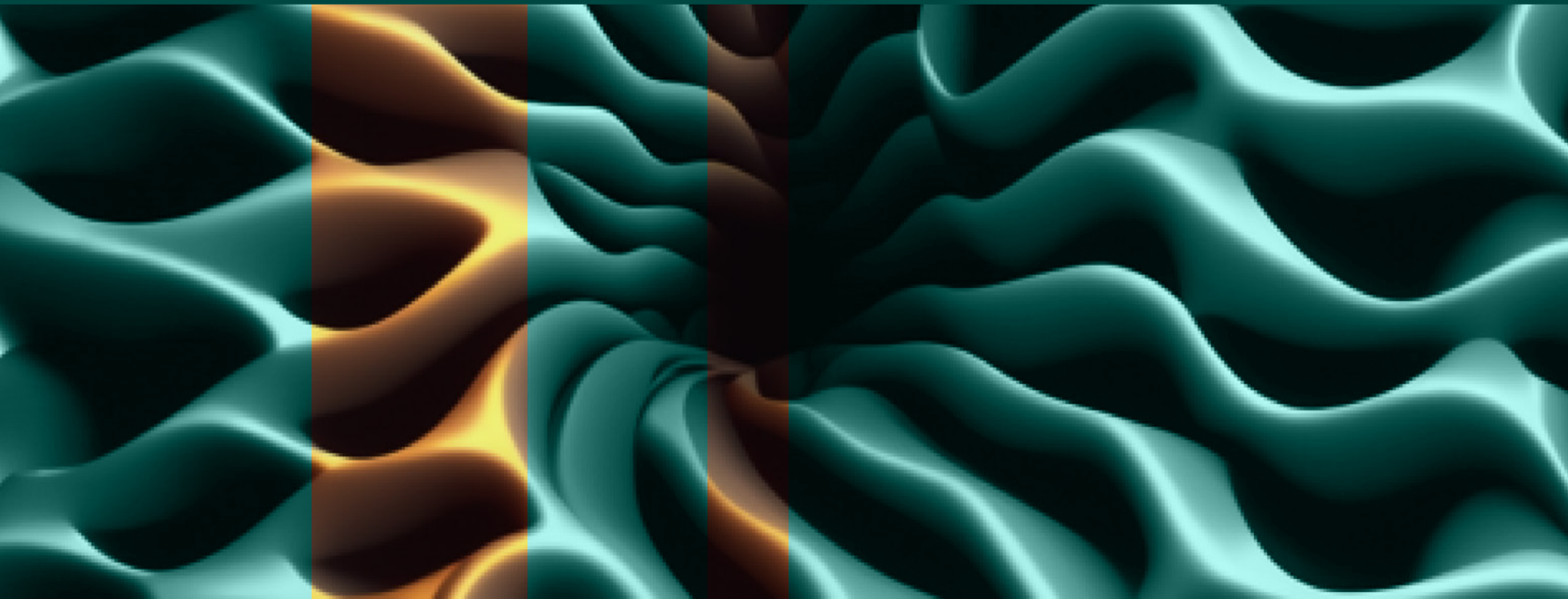


Chat with your docs

How generative AI is fast-tracking
the way employees interact with
corporate data

Globant▶
Enterprise AI



The era of sifting through folders or relying solely on metadata for unstructured data retrieval is gradually coming to an end. The advancements in Generative AI, particularly Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG), are setting the stage for a more intuitive, conversational approach to accessing your documents.

While Large Language Models, like GPT-4, bring a wealth of knowledge to the table, they aren't exempt from limitations—particularly when it comes to integrating new or organization-specific data. Traditional methods like fine-tuning can modify how these models communicate but are less effective for incorporating fresh data, which is often a critical business need. Additionally, fine-tuning demands significant resources when it comes to data quality, computational power, and time—luxuries not everyone can afford.

Retrieval-Augmented Generation (RAG) offers a dynamic and cost-effective approach to model adaptation. RAG is powerful and also simple enough for us to delve into its mechanics in this article, which will explore the principles behind RAG, its real-world applications, and how it addresses the limitations of traditional model fine-tuning. Whether you face constraints related to data, budgetary considerations, compliance requirements, or time limitations, there is a way forward for you to harness the power of AI in a manner that is effective and prioritizes security and efficiency. Using our cutting-edge accelerator platform, you can expedite productivity enhancements even further.

What is RAG?

The concept behind Retrieval-Augmented Generation (RAG) might seem complex, but it can be broken down into a sequence of straightforward steps:

1. User Query

The process kicks off with the user asking a question.

2. Document Retrieval

The system then scours an indexed document repository, often filled with proprietary data, to find documents that could answer the query.

3. Prompt Creation

Next, the system crafts a prompt that marries the user's question, the retrieved documents, and additional instructions. This prompt sets the stage for the Large Language Model (LLM) to generate an answer.

4. LLM Interaction

This tailored prompt is sent to an LLM for processing.

5. User Response

Finally, the LLM returns a contextually grounded answer, based on the information it was provided, back to the user in natural language.

While this explanation provides a high-level view of RAG, the devil is in the details, and implementing your own enterprise-grade solution might be completely out of the scope of your business. Is the investment worthwhile? What are the use cases that apply to your specific needs? Let's review them.



Conversing with your Documents is now possible

Imagine asking your documents questions in natural language and receiving immediate responses. With RAG, you can make this a reality. Here's how it can benefit your organization:

Instant Information Retrieval

You can query your documents as if you were chatting with a knowledgeable assistant. Need to find a specific clause in a contract, statistical data in a report, or historical information in your archives? Simply type in your question, and an AI assistant will fetch the relevant information in seconds.

Document Summarization

Another valuable capability of Generative AI is its ability to summarize lengthy documents quickly and accurately. Whether you need a concise overview of a long report or a summary of key contract clauses, Generative AI can generate summaries that save you time and effort. This feature is handy during meetings or when you need to grasp the essentials of a document without delving into its entire content.

Enhanced Decision-Making

Generative AI can help you make informed decisions by highlighting key insights and even providing recommendations based on the content. This capability is invaluable for legal departments reviewing contracts, financial analysts studying reports, or any department seeking data-driven insights.

Regulatory Compliance

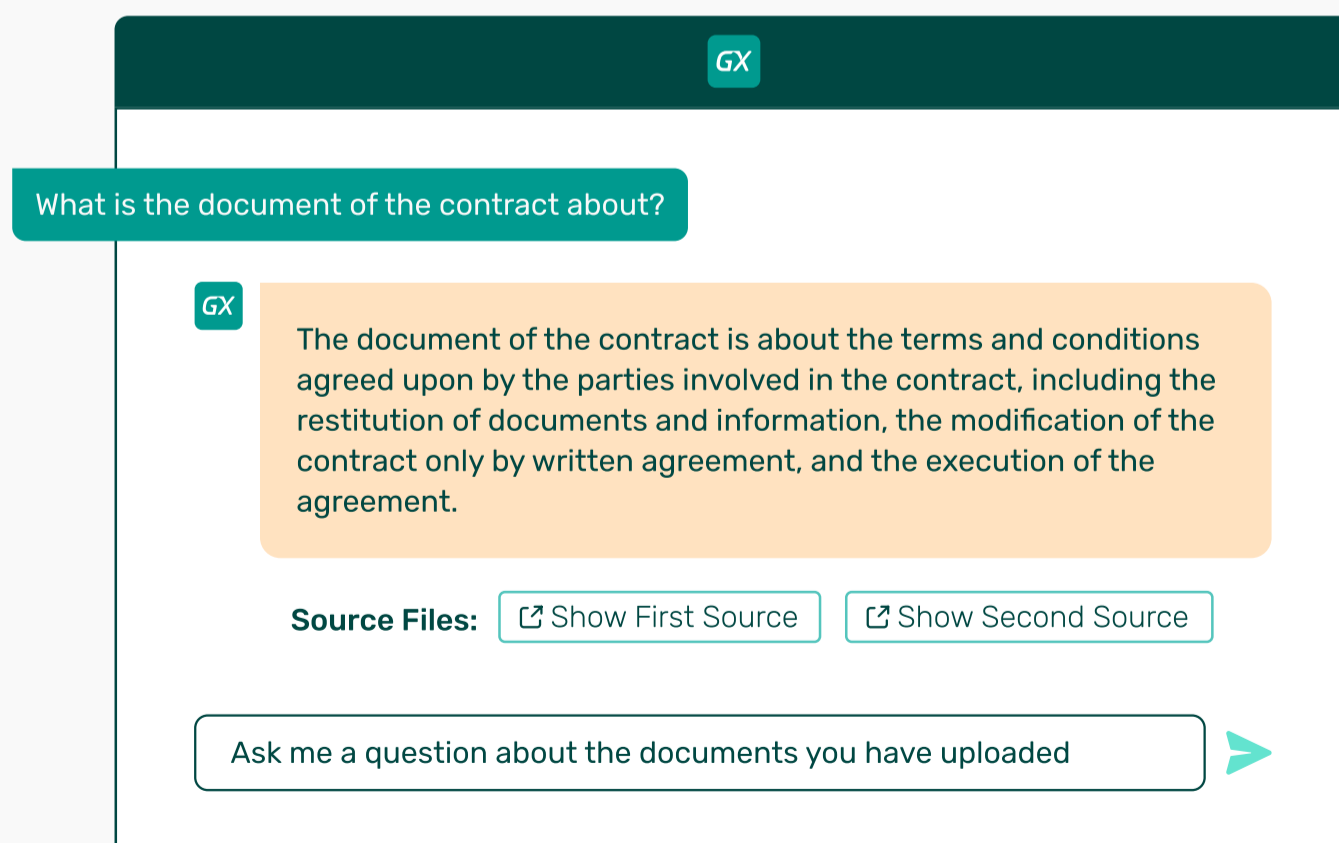
AI is a huge asset for industries with heavy regulations documented in PDFs or various documents, as they can ensure compliance by quickly extracting and interpreting regulatory information, reducing the risk of non-compliance. Having to set aside hours of labor to research and discern compliance information is a problem of the past.

Cross-Functional Information Access

Gen-AI democratizes access to information within your organization. Whether you're a legal professional, analyst, or customer support agent, you can harness the power of Generative AI to access critical data and insights, regardless of your technical expertise, freeing up expert time to do higher-value tasks.

Cross-Lingual Information Retrieval

Generative AI enables users to search for information in one language and retrieve relevant results in another. This powerful function is particularly useful when dealing with multilingual data or when users are more comfortable searching in their native language. Imagine you have a partner in China and need to review their documentation. By leveraging the power of LLMs, you could chat with those documents in English to obtain the information you need.



Industry Specific Impact

AI assistants can handle repetitive or non-value-adding tasks, allowing your team to focus on more strategic jobs. Whether you're in the legal, insurance, or pharmaceutical industries, Generative AI can adapt to your specific needs, providing a tailored solution that enhances productivity. Let's take a look at some examples:

Claims Analysis in Insurance

An AI assistant can analyze relevant discrepancies or inconsistencies between the claimant's description and the documented evidence, which can help detect potential fraud or inaccurate claims.

Research and Development in Pharma

You can make searches and compare clinical trials and studies related to a specific disease to identify potential therapeutic approaches and contribute to developing new treatments.

Risk Analysis in Banks

The AI assistant can connect to customers' credit reports and solvency assessments to quickly analyze the information and assess creditworthiness. This helps banks make more informed decisions regarding loan approvals and credit lines.

Compliance and Regulation in Automotive

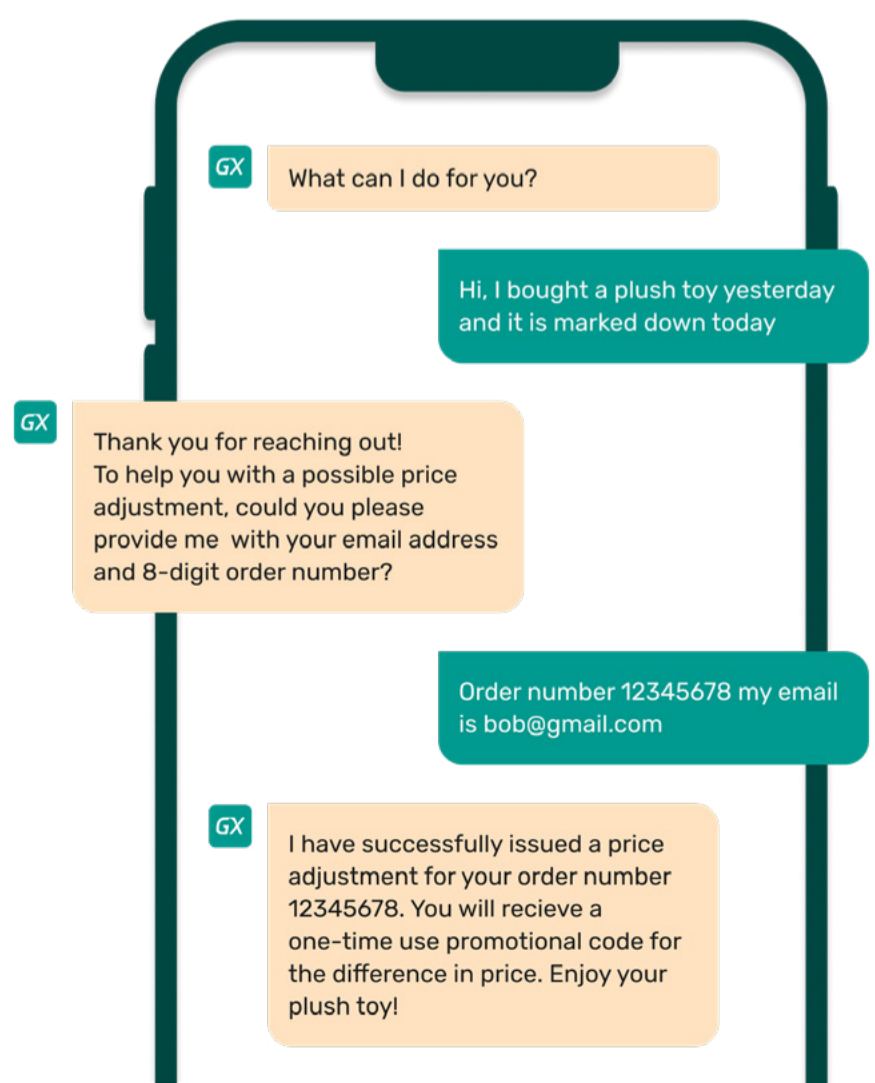
You can analyze industry-specific regulations and standards to ensure that manufactured vehicles comply with safety standards and legal requirements.

SLA Compliance in Telcos

The AI assistant can analyze the service level agreements (SLAs) established with corporate clients and compare them with actual performance and service quality data. This allows for assessing SLA compliance and taking corrective measures in case of any deviations.

Incident Report Analysis in Airlines

An AI assistant can examine incident reports, such as flight delays, cancellations, baggage issues, or onboard service problems, to identify patterns and trends. This application helps the airline process vast amounts of unstructured data to understand problem areas and then take corrective actions to prevent recurring issues.



How can you implement these solutions in a cost-effective and fast way?

This is the point at which Globant Enterprise AI comes into play. Globant Enterprise AI is the backbone that connects companies to LLMs in a monitored and cost-effective way, providing agility to embrace the power of Generative AI. It enables the creation of AI-powered assistants that integrate and interact with organizational operations, processes, systems, and doc-

uments. These AI assistants can be tailored to specific corporate requirements, like the ability to chat with documents for valuable outcomes. Moreover, it allows you to securely connect your organization with LLMs, avoiding any leak of private information and providing monitored access. But there's more to it. Globant Enterprise AI can help you:

Unlock the adoption of Gen-AI through:

Enhanced Security

Monitor LLM access using your authentication and authorization definitions.

Solution Management

Supervise costs and interactions of each Gen-AI solution for further analysis and control.

Scalability

Scale your solutions to multiple channels, such as Web, Slack, Teams, or WhatsApp.

Cost Control

Manage quotas per solution and control your spending.

Unleash the creation of Gen-AI solutions through:

Automated execution

Automatically create APIs and version your assistants and prompts.

Optimized Exploration

Easily switch between different LLMs like OpenAi, Google, AWS, and Azure to discover the best fit for your needs without altering what you've already built.

Reduced Coupling

Reduce the level of dependency between the AI applications you create and the underlying LLMs.

A step forward for AI-inspired organizations

Generative AI and Retrieval-Augmented Generation (RAG) are changing the game when it comes to interacting with documents in your organization. Whether you're looking to improve unstructured data retrieval, summarization, contract analysis, or any other document-related task, Generative AI offers a powerful and efficient solution. And alongside Globant Enterprise AI, you can quickly unlock their potential to boost productivity and unleash innovation within your organization. What are you waiting for? Fuel your AI motion today.

Get in touch