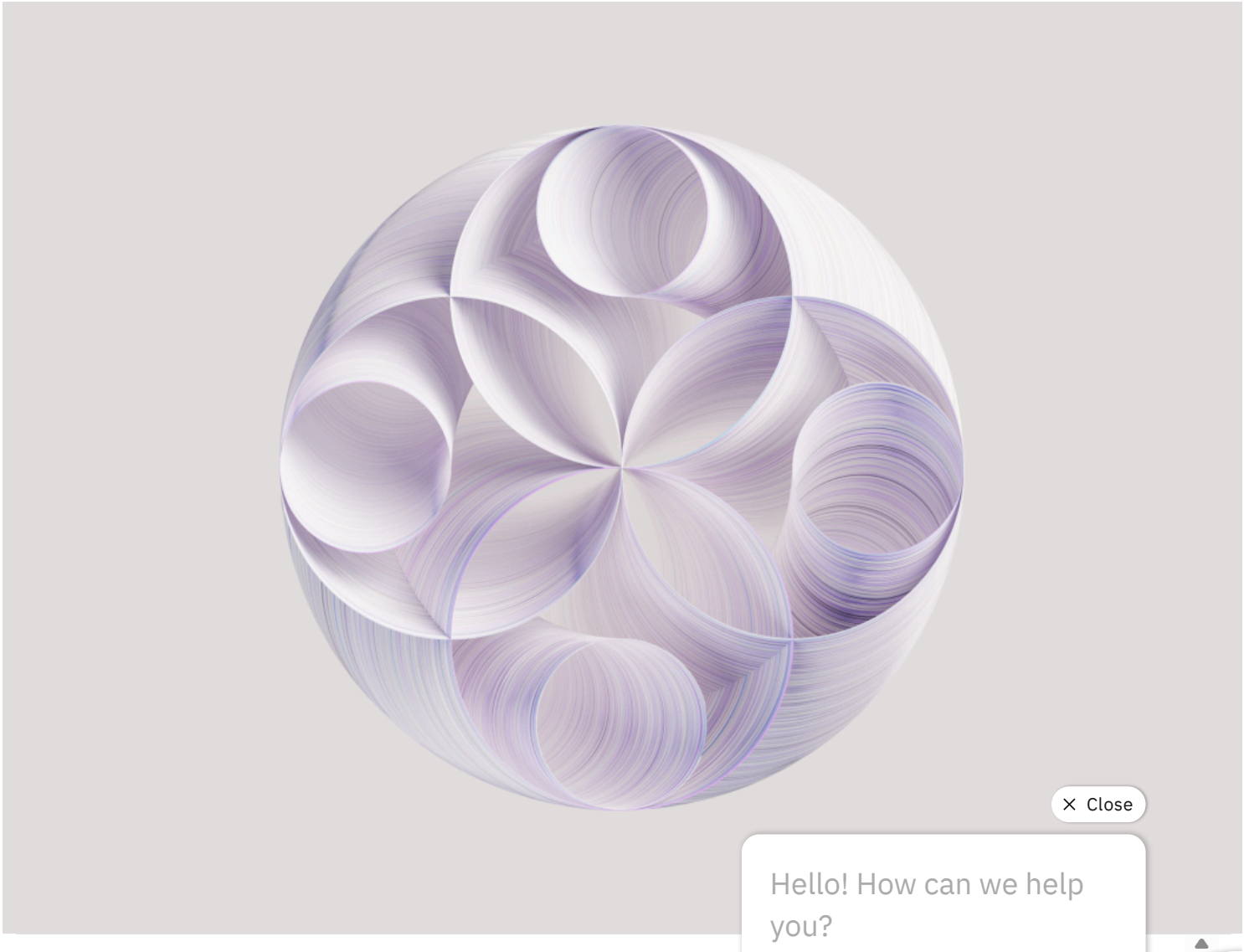
Harnessing the power of generative AI for structured data

GigaSpaces + IBM



About GigaSpaces Watsonx.ai is helping companies custom-h

GigaSpaces, a global pioneer of in-memory computing and an IBM Business Partner, is building one of the market's first Digital Integration Hubs (DIHs), an out-of-the-box solution that simplifies organizations' digital transformation and enables them to develop and launch digital services at a rapid pace. GigaSpaces' solutions are utilized by tier-1 organizations all over the world, enabling mission-critical services that rely on real-time data.

Building upon its existing track record in delivering real-time data to digital services, GigaSpaces is developing a new solution, enterprise retrieval augmented generation (eRAG) (link resides outside of ibm.com), that will enable its customers to use generative AI (gen AI) large language models (LLMs) to query structured data that is typically organized and cataloged in a complex and unintuitive manner.

near 95%

accuracy rate with eRAG approach is the mission

days

instead of months to prototype an assistant based on eRAG

The GigaSpaces team is building the new solution using IBM® watsonx.ai™ and IBM watsonx Assistant with the support of the IBM Build Lab. It will contain a virtual assistant whose front-end is powered by watsonx Assistant, and whose query and summarization capabilities are powered by wasonx.ai. The vision is for GigaSpaces' customers to be able to query enterprise data in natural language. The primary goal of the eRAG platform is to empower organizations to access enterprise data in natural language, enabling them to make reliable business decisions and actionable insights through AI. By using the advanced LLMs on IBM watsonx.ai and applying techniques like semantic enrichment and vector caching, GigaSpaces is able to enhance the metadata reasoning and retrieval modules of eRAG.

Once the new eRAG platform is fully operational, GigaSpaces' customers will be able to use natural language to query structured data such as a company's internal sales records, inventory management, order payments, or marketing analytics. They can expect eRAG to deliver high levels of accuracy in the query results it returns. Currently, engineers who use AI to work with structured information such as databases see an up to 55% accuracy rate. GigaSpaces' mission is to help its customers achieve a near 95% accuracy rate based on initial results when using this new eRAG approach.

By integrating IBM watsonx Assistant, GigaSpaces significantly reduced its time to prototype an assistant based on eRAG to within several days, instead of months, allowing the R&D team to concentrate on the core challenge of enhancing eRAG technology and streamlining the development of human language for structured information at a very high level. Going forward, GigaSpaces plans to explore IBM watsonx.governance™, a solution that uses software automation to mitigate risks, manage regulatory requirements and address ethical concerns for both gen AI and machine learning (ML) models. GigaSpaces aims to harness the platform to provide confidence and security within the eRAG solution, ensuring evaluation, monitoring and control over the models integrated into the solution. The introduction of IBM watsonx.governance will mark a pivotal advancement toward achieving this objective.

About GigaSpaces

GigaSpaces (link resides outside of ibm.com) is a software company and an IBM Business Partner, that offers in-memory computing and operational data technologies with one of the market's first Digital Integration Hubs (DIH).

Solution components

- IBM® watsonx™ Assistant →
- IBM watsonx.ai™ →
- IBM watsonx.governance →

Watsonx.ai is helping companies custom-build AI solutions to suit their specific needs.

Explore watsonx.ai →

Start your free trial →

Hello! How can we help you?

Legal

Produced in the United States of America, March 2024.