

Safe, Reliable Deployment of Mistral AI's Open-Source LLM Paired with Squirro's RAG Technology

Squirro delves into the immense potential of Mixtral-8x7B-Instruct—expounding on the impact of RAG as a key enabler to one of the best open-LLMs on the market.

ZURICH, SWITZERLAND, February 29, 2024 /EINPresswire.com/ -- Squirro, the generative AI-enabled semantic enterprise search provider, delves into the immense potential of Mixtral-8x7B-

Instruct—expounding on the impact of retrieval augmented generation (RAG) as a key enabler to one of the best-in-class open-LLMs on the market.

In a comprehensive [white paper](#) released this week, Squirro explains the tremendous potential of fusing [Mistral AI's](#) powerful open-source model (Mixtral-8x7B-Instruct) with an enterprise-ready information retrieval stack—uncovering the capabilities, drawbacks, and benefits of applying LLMs in highly regulated institutional contexts for generative AI adoption.

Mixtral-8x7B-Instruct: The Case for Open-Source LLMs

While excitement about generative AI remains high among business leaders, security and deployment challenges persist—posing questions about closed-source black-box models developed by big tech and the inherent risk of adoption lag. Hallucination, lack of transparency, and use with private data are some of the concerns faced by leaders and developers along their journeys in search of an enterprise-ready AI stack.

Combining AI with the power of open-source software is fast becoming one of the hottest topics for high regulation organizations looking to securely experiment with their data in 2024. To build trust and maximize data transparency, organizations that adhere to strict government and industry regulations can leverage open-source ecosystems to reduce costs, deploy faster with increased flexibility.

With the hype of generative AI consuming headlines at this year's World Economic Forum, French-based start-up Mistral AI took center stage as the preferred open-source LLM. With its



Squirro delves into the immense potential of Mixtral-8x7B-Instruct

open weight, high-quality Sparse Mixture of Experts (SMoE) model—allowing customers to easily adapt their models for on-premises self-deployment—Mixtral-8x7B-Instruct is strategically positioned as one of the best available products, outperforming LLaMA 2 70B and competing with GPT-3.5-Turbo on most benchmarks.

SquirroGPT: The First Enterprise-Ready RAG Stack

While open-LLMs offer profound value in agility and probability computation, they need a reliable companion to guide contextual information retrieval. This is where Squirro's LLM agnostic technology—rooted in RAG—steers accuracy, safety, and reliable deployment.

Whether deploying on-premises or in a private cloud setting, enterprise security remains the number one priority for central banks, insurance, and healthcare organizations. To help institutions address this fundamental need, Squirro's RAG-enabled technology has been delivering enterprise-ready solutions for the past 10 years. Now, by harnessing the power of open-source LLMs, regulatory companies can benefit from a fully integrated offering—enabling technical teams to deploy their technology stacks safely in a containable environment.

Reflecting on the power of open-LLMs and effectively integrating the Mixtral-8x7B-Instruct model, Saurabh Jain, CTO of Squirro, elaborates: "Squirro's RAG now supports the open-source LLM Mixtral-8x7B-Instruct, making it feasible for regulated industries to adopt GenAI solutions. We recognized early on that the transformative potential of GenAI for enterprises is often hindered by privacy concerns and regulatory compliance. It became apparent that to truly democratize the benefits of GenAI for enterprises where privacy is a key concern, we needed to develop a solution that is LLM-agnostic. This architectural decision early on allowed us to support open-source LLMs, starting with Mixtral. This approach ensures that every organization, regardless of the regulatory environment, can leverage this powerful technology with absolute control over their own data, without third-party data control. If data privacy is a concern for you, we invite you to explore how Squirro's RAG and Mixtral-8x7B-Instruct can enhance your approach to leveraging AI".

[Try Mixtral-8x7B-Instruct with SquirroGPT](#)

By combining the power of Mixtral-8x7B-Instruct and Squirro's RAG technology, your knowledge base will be deployed reliably without external party control. Squirro's enterprise-ready application stack provides a fully controlled solution at low risk that removes any security concerns, simplifying on-premises deployment.

About Mistral AI

Mistral AI is a French-based artificial intelligence company founded by researchers previously employed by Meta and Google DeepMind. The machine learning startup is backed by more than USD 500 million in funding and recently debuted its open weights, high quality Sparse Mixture of Experts (SMoE) model. The Mixtral-8x7B-Instruct model comprises eight different neural networks, featuring 46.7 billion parameters. To learn more, visit: <https://mistral.ai/>

About Squirro

Squirro is the leading enterprise ready generative AI solution for search, insights and automation. Squirro empowers organizations across the globe to transform enterprise data into knowledge, insights and recommendations. Squirro has a track record of more than ten years in marrying AI, machine learning, predictive analytics, generative AI, and symbolic AI-like knowledge graphs, together.

Founded in 2012, Squirro is a fast-growing company with dedicated teams in Switzerland, the United States, the UK, and Singapore. Our customers include the European Central Bank, the Bank of England, Standard Chartered Bank, Oversea-Chinese Banking Corporation, Henkel, Armacell, Library of Congress and Indicia Worldwide. To know more about us visit - <https://squirro.com/> or book a demo with us - <https://squirro.com/book-a-demo>.

Anupriya Chaturvedi

Squirro AG

[email us here](#)

Visit us on social media:

[Facebook](#)

[Twitter](#)

[LinkedIn](#)

[Instagram](#)

[YouTube](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/691934723>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2024 Newsmatics Inc. All Right Reserved.