

JUN 17, 2024

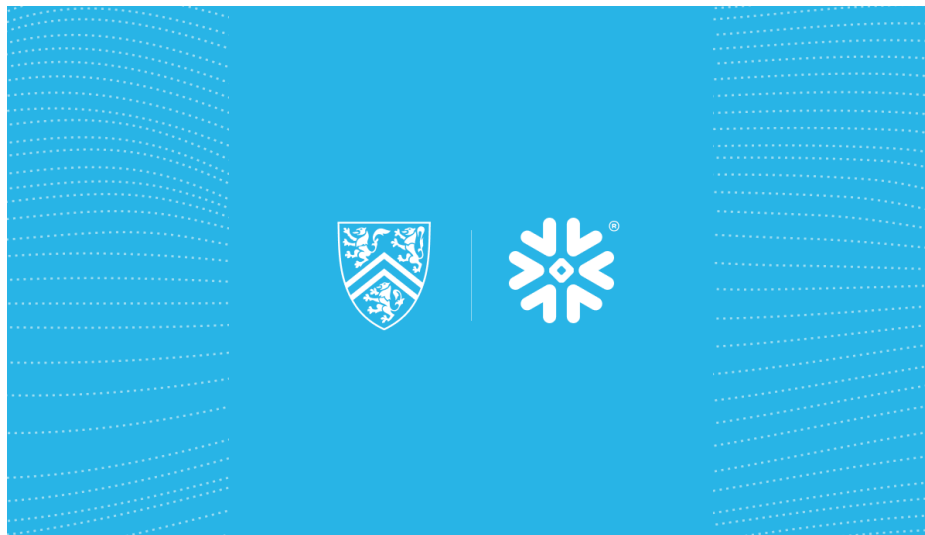
AUTHOR

Snowflake AI Research

SHARE

Moving Beyond MTEB and BEIR: Snowflake AI Research Joins Forces with the University of Waterloo to Evolve RAG and Retrieval Benchmarks

Gen AI



To accurately answer business questions using LLMs, companies must augment models with their data. Retrieval Augmented Generation (RAG) is a popular solution to this problem, as it integrates the organization's factual, real-time data into the prompt for the LLM. While the adoption of RAG has increased, an open question remains: How do enterprises know how effective their system is?

Due to the growth of interest in improving retrieval quality, open and collaboratively developed benchmarks, such as [BEIR](#), [MTEB](#), and [MSMARCO](#), have made it easier to compare and evaluate the surge of new retrieval systems. These benchmarks evolved from




collections of independent data sets in well-studied workloads that we, along with many other retrieval experts, used to quantify the performance of Arctic-embed model. As we continue to develop more advanced and efficient retrieval that enables enterprises to talk to their data, it's crucial to ensure that the benchmarking data sets represent these use cases directly.

AUTHOR

Snowflake AI Research

Building on Snowflake's broadly used data cloud, we aim to openly and collaboratively support the evolution of retrieval benchmarks to propel the industry forward.

SHARE

To help the broader ecosystem continue to improve performance, we're thrilled to announce a unique collaboration between    a team of retrieval experts from the University of Waterloo, which is renowned for its research prowess under [Professor Jimmy Lin](#). Together, we're embarking on a mission to build the next generation of retrieval evaluation benchmarks to understand better and evaluate how RAG agents perform.

"As a researcher, I'm thrilled to collaborate with Snowflake on this joint mission to build an improved representation of real-world retrieval applications," said Prof. Lin. "The expertise in practical enterprise AI from Snowflake, combined with our academic insights, promises to unlock new frontiers in AI innovation."

At Snowflake, we aim to empower our customers to get the most out of their enterprise data. From efficient and scalable elastic computing, to the best tools and frameworks to talk to your data, we strive to deliver insights quickly, accurately, and efficiently. With the growth of RAG-like systems and workflows, it quickly became apparent that we must qualify and quantify how well these systems perform.

Like all prior benchmarks, metrics, and tasks become saturated, and the gap between improvements on a leaderboard and reality begins to widen. In our work on our open source embedding model family, [Snowflake Arctic embed](#), we found MTEB crucial to quick iteration and qualification but saw a growing gap in improvements on existing benchmarks compared to our internal benchmarks.

Our collaboration is not about creating novel retrieval models. It's about creating novel open-source data sets and tasks to




revolutionize the field. We’re fostering a community-driven approach to research and development, a strategy that promises to bring about exciting and groundbreaking changes.

AUTHOR

Snowflake AI Research

TREC RAG: Using the past experiences of Professor Lin and Snowflake’s own Dr. Daniel Campos in creating world-class benchmarks and data sets, this RAG track focuses on understanding and evaluating the quality of cited and grounded generation and how it is influenced by the quality of retrieval, generation mode and use case.

SHARE

   Benchmarking Evaluation of Information Retrieval): Building on **Nandan Thakur’s** experience in building the first BEIR benchmark and expertise with commercial search systems, we seek to create a new and improved retrieval benchmark that is more representative of the workloads people use embedding models for.

We’re not just excited about this journey; we’re thrilled. Thrilled to shape the future of information retrieval and AI with the University of Waterloo, Professor Jimmy Lin and his brilliant researchers. Stay tuned for updates on our progress and the breakthroughs that will emerge from this collaboration. We’re confident that they will be nothing short of remarkable!

Join us at the Snowflake Data Cloud Summit in San Francisco this June to learn more about our AI research.

SHARE



Start your 30-day free trial

START NOW!



	PLATFORM	SOLUTIONS	RESOURCES	EXPLORE	ABOUT
AUTHOR Snowflake AI Research	Cloud Data Platform	Snowflake for Financial Services	Resource Library	News	About Snowflake
	Pricing		Webinars	Blog	Investor Relations
	Marketplace	Snowflake for Advertising, Media, & Entertainment	Documentation	Trending	Leadership & Board
	Security & Trust		Community	Guides	
SHARE		Snowflake for Retail & CPG	Procurement	Developers	Snowflake Ventures
			Legal		Careers
		Healthcare & Life Sciences Data Cloud			Contact
		Snowflake for Marketing Analytics			



Sign up for Snowflake Communications

diana.shaw@sn

United States

By submitting this form, I understand Snowflake will process my personal information in accordance with their **Privacy Notice**. Additionally, I consent to my information being shared with Event Partners in accordance with Snowflake's **Event Privacy Notice**. I understand I may withdraw my consent or update my preferences [here](#) at any time.

SUBSCRIBE NOW

Privacy Notice | Site Terms | Cookie Settings | Do Not Share My Personal Information

© 2024 Snowflake Inc. All Rights Reserved | If you'd rather not receive future emails from Snowflake, [unsubscribe here](#) or [customize your communication preferences](#)

https://www.snowflake.com/engineering-blog/snowflake-ai-research-joins-forces-with-the-university-of-waterloo/?mkt_tok=MjUyLVJGTy0yMjcAAAGU1rxV1lqsrI...

4/4