Cerebras Delivers End-to-End Training and Inference for Jais 2, the World's Leading Open Arabic LLM

December 9, 2025

Trained on the industry's largest, highest-quality Arabic-first dataset, Jais 2 sets new standards for accuracy, fluency, and cultural intelligence

SUNNYVALE, Calif.--Cerebras Systems, in partnership with G42's Inception and MBZUAI's IFM, today announced the release of Jais 2, the leading open-source Arabic LLM – the first frontier language model both **trained and deployed for inference on Cerebras Systems**. The organizations combined their expertise with leading machine learning techniques, uniquely enabled on Cerebras wafer-scale clusters, to achieve state-of-the-art quality on Jais 2, using only a fraction of compute used to train similar-sized models in the past.

> "This marks the first time a frontier-grade LLM has been trained end-to-end and deployed in production for inference on Cerebras hardware, demonstrating a new, efficient blueprint for sovereign AI development," Natalia Vassilieva, Field CTO, Cerebras

"This marks the first time a frontier-grade LLM has been trained end-to-end and deployed in production for inference on Cerebras hardware, demonstrating a new, efficient blueprint for sovereign AI development," said Natalia Vassilieva, VP and Field CTO, Cerebras. "Training and serving a large-scale Arabic model at scale on a unified compute architecture greatly simplified operations, reduced costs, and accelerated time to capability."

Jais 2 models — including the new 8B and 70B generations — were trained entirely on Cerebras wafer-scale clusters, using ML methodologies uniquely enabled by the system's massive compute, memory, and communication bandwidth. In production, the Jais 2 chat application now runs inference on Cerebras hardware as well, delivering up to 2,000 tokens per second, making it one of the fastest LLMs in the world.

Built for the Arab World, Beyond the Limits of Western-Optimized Models

Arabic is spoken by more than 400 million people, yet most frontier models are optimized primarily for English. Arabic with its rich morphology and cultural diversity continues to be underrepresented. Jais 2 closes this gap through a model family purpose-built for the Arab world — combining frontier-level intelligence with deep Arabic linguistic, cultural, and contextual fidelity.

> "Arabic has long been underserved in AI development due to limited high-quality data for training large language models," said Professor Preslav Nakov, Department Chair and Professor of Natural Language Processing at MBZUAI. "Today, with Jais 2, we share a model that is built not only with scale, but with cultural and linguistic fidelity at its core. By dramatically expanding the quality and diversity of Arabic data, we created a foundation that reflects the richness of the Arabic language."

Large Western general-purpose models, while powerful, are not aligned with Arabic culture, dialects, or norms. This often leads to poor model performance in key domains such as:

- regional dialect variation
- politeness, honorifics, and deference norms
- religious, ethical, and legal reasoning
- culturally specific idioms, humor, and proverbs

Existing Arabic-specific models, meanwhile, have typically been too small to preserve the reasoning depth and multilingual fluency needed for modern applications.

Jais 2 bridges this divide: a model with frontier-level capability and native Arabic cultural alignment. Jais 2 excels in domains deeply rooted in Arab life, such as poetry, religion, cuisine, and dream interpretation.

A Step Change from the Original Jais Family

Building on the success of the original Jais Family of bilingual models, Jais 2 introduces:

- New 8B and 70B models

- A redesigned architecture optimized for efficient training
- A dramatically expanded and cleaner Arabic-first dataset
- A more rigorous fine-tuning and alignment pipeline
- Best-in-class performance on a key Arabic leaderboard: AraGen

Jais 2 70B establishes new state-of-the-art accuracy for Arabic LLMs while maintaining competitive multilingual reasoning and comprehension.

Jais 2 is available on Hugging Face, and the updated Jais Chat app is live on web, iOS and Android.

For more information, visit JaisChat.ai or read our technical blog.

About Cerebras Systems

Cerebras Systems builds the fastest AI infrastructure in the world. We are a team of pioneering computer architects, computer scientists, AI researchers, and engineers. We make AI blisteringly fast through innovation and invention because we believe that when AI is fast it will change the world. Our flagship technology, the Wafer Scale Engine (WSE) is the world's largest and fastest AI processor. At 56 times larger than the largest GPU, the WSE uses a fraction of the power per unit compute while delivering inference and training more than 20 times faster than the competition. Leading corporations, research institutes and governments on five continents chose Cerebras to run their AI workloads. Cerebras solutions are available on premise and in the cloud. For further information, visit cerebras.ai or follow us on LinkedIn, X and/or Threads.

Contacts

ZM Communications
pr@zmcommunications.com