Ai2 (Allen Institute for AI) Announces Partnership with Google Cloud to Accelerate Open AI Innovation

Ai2 makes portfolio of open source AI models available in Vertex AI Model Garden to re-shape how regulated industries, the public sector, and leading enterprises deploy AI at global scale

Seattle and Sunnyvale, Calif. — April 8, 2025 – Today, Ai2 announced a partnership with Google Cloud to make its portfolio of open AI models available in Vertex AI Model Garden. The collaboration will help set a new standard for openness that leverages Google Cloud's infrastructure resources and AI development platform with Ai2's open models that will advance AI research and offer enterprise-quality deployment for the public sector.

While AI reshapes industries and enterprises at an unprecedented pace, there remains a persistent gap in its adoption and impact due to the lack of model transparency and data privacy concerns. Open source AI models provide another option to unlock this untapped potential and enable both the public sector and highly regulated industries to meaningfully benefit from AI. Together, Ai2 and Google Cloud will give customers the ability to more easily access, openly customize, and deploy their AI solutions to innovate more quickly.

"This partnership is a reflection of our shared vision with Google Cloud to address the current limitations in AI deployment by building the next generation of truly open and trustworthy AI," said Ali Farhadi, CEO, the Allen Institute for AI. "While openness is key to building trust and advancing the development of AI, accessibility and data privacy are similarly fundamental to the acceleration of innovation across industries. By providing access to its AI-enabled infrastructure to support advanced and efficient model training, Google Cloud is critical to ensuring we can deliver state-of-the-art, transparent AI solutions while adhering to the strictest privacy standards, paving the way for a more secure and impactful future with AI. This will ultimately empower entire industries and even regulated sectors like government, health, sciences and more to do something they simply can't today: transform their domains with truly open AI."

This partnership will also help deliver transformative change to these agencies and enterprises by providing efficiencies and powerful computing tools in an open and transparent environment.

"In the public sector, trust is paramount. Transparency and accessibility are not just ideals; they are the foundation for deploying AI that improves citizen services," said Karen Dahut, CEO, Google Public Sector. "True openness, particularly in the creation and validation of AI models, is crucial for building that trust. Ai2's commitment to fully open training code, evaluation methods, and data aligns perfectly with our mission to drive responsible and effective AI innovation within government. We are proud to partner with Ai2 to deliver AI solutions directly to researchers, government agencies, and other public sector leaders to ensure widespread benefits to citizens."

Through this partnership, Ai2 will make its full portfolio ofstate-of-the-art open AI models available in Google Cloud's Model Garden on Vertex AI. This includes the recently released OLMo 32B, Ai2's new, fully-open language model trained onGoogle Vertex AI and designed to provide more audiences with access to a powerful, cost-effective model.

Vertex AI & Model Garden
Ai2's truly open models, including OLMo, Tulu, and Molmo model families, will be available on Google Cloud's Vertex AI Model Garden. OLMo32B will be available via API as part of Vertex AI's model as a service (MaaS). In addition to the models, AI researchers and developers will be able to access Ai2's open models, datasets, and recipes via Vertex, including all Tulu recipes to enable Google Cloud customers to fine-tune state-of-the-art models for their specific use cases.

Introducing OLMo 32B: accessible open-source AI
Ai2's OLMo 32B is a new open-source model designed to bring cutting-edge AI within reach for more organizations. The model was trained on a 160 node GPU cluster provided by Google Cloud. Each node had 8 Nvidia H100 GPUs connected with the GPUDirect-TCPXO interconnect running at 1800 tokens per second. OLMo32B, optimized to run efficiently on a single GPU box, 32B strikes the ideal balance between performance and accessibility: delivering the necessary accuracy for a wide variety of use-cases without the excessive costs of oversized deployments. Ai2's latest release reflects its commitment to open, practical, cost-effective AI—giving businesses, researchers, and developers the power to innovate without overspending.

About Google Cloud
Google Cloud is the new way to the cloud, providing AI, infrastructure, developer, data, security, and collaboration tools built for today and tomorrow. Google Cloud offers a powerful, fully integrated and optimized AI stack with its own planet-scale infrastructure, custom-built chips, generative AI models and development platform, as well as AI-powered applications, to help

organizations transform. Customers in more than 200 countries and territories turn to Google Cloud as their trusted technology partner.

About Ai2
Ai2 is a Seattle-based non-profit AI research institute with the mission of building breakthrough AI to solve the world's biggest problems. Founded in 2014 by the late Paul G. Allen, Ai2 develops foundational AI research and innovative new applications that deliver real-world impact through large-scale open models, open data, robotics, conservation platforms, and more. Ai2 champions true openness through initiatives like OLMo, the world's first truly open language model framework,Molmo, a family of open state-of-the-art multimodal AI models, and Tulu, the first application of fully open post-training recipes to the largest open-weight models. These solutions empower researchers, engineers, and tech leaders to participate in the creation of state-of-the-art AI and to directly benefit from the many ways it can advance critical fields like medicine, scientific research, climate science, and conservation efforts. For more information, visit allenai.org**.**

Contacts
Ai2
Sophie Lebrecht
sophiel@allenai.org

---

https://www.googlecloudpresscorner.com/2025-03-08-Ai2-Allen-Institute-for-AI-Announces-Partnership-with-Google-Cloud-to-Accelerate-Open-AI-Innovation