

Ginkgo Bioworks Launches New Protein LLM and Model API Built on Google Cloud Technology

- Protein large language model (LLM) designed to help enterprises accelerate drug development coming to Google Cloud's Vertex AI Model Garden soon; one of the first-of-its-kind in the industry
- Model API programmable interface now available for individual scientists and researchers to quickly test and advance work

BOSTON, Sept. 17, 2024 [PRNewswire/](#) -- Ginkgo Bioworks Holdings, Inc. (NYSE: DNA), which is building the leading platform for cell programming and biosecurity, today announced two new offerings that will make it easier for pharmaceutical and biotech companies to develop new medicines, building on the [partnership announced](#) with Google Cloud last year. The first launch is one of the first-of-its-kind in the industry – a protein large language model (LLM) built in collaboration with Google Cloud Consulting that will give individual researchers and enterprise companies the ability to develop medicines with insights from Ginkgo's private data. Secondly, Ginkgo is launching its [model API](#), a powerful tool designed to bring biological AI models directly to machine learning scientists. The API is publicly available on [Ginkgo's website](#) today, and enterprise companies will be able to access the protein-based LLM Google Cloud's Vertex AI [Model Garden](#) soon.

Jason Kelly, CEO of Ginkgo Bioworks: "I'm excited to see how the community builds on top of these models and our API. AA-0 is the first model we have released trained on Ginkgo's proprietary data – we're opening it to data scientists and bioinformaticians so they can build new models and applications on top. We believe that the low cost for tokens and our other customer-friendly terms, like the lack of royalties, and our commitment to not re-use customer data, will allow users to build tools like iterative protein design programs that call our protein generation API or to use our embedding API to compute features for a clustering algorithm."

These new offerings demonstrate how Ginkgo is enabling the life sciences industry in new ways, helping them improve and accelerate the drug development process.

- **Protein LLM for individual researchers and enterprise companies:** Built on Vertex AI in collaboration with Google Cloud Consulting and trained on Ginkgo's extensive proprietary dataset, this and future LLMs empower companies to generate novel insights and accelerate the discovery of new therapeutics. By harnessing the power of AI to analyze and understand complex protein structures and interactions, researchers and enterprises can streamline their research pipelines, optimize lead identification, and ultimately bring life-saving medicines to market faster and more efficiently. Building on models that learn from Ginkgo's private data, unavailable to the public, can enable companies to unlock hidden patterns and potential therapeutic targets that would otherwise remain elusive.
- **Open API for scientists and researchers:** With this programmer-friendly ultra-low cost API, Ginkgo is making its internally-developed AI tools available to anyone. The interface provides an easy and scalable way to access sophisticated models trained on protein and DNA data, starting with its first release: a machine learning model trained on a proprietary Ginkgo dataset. (Read more about Ginkgo's first model — ginkgo-AA-0-650m, a large-scale model trained on 2+ billion proprietary Ginkgo protein sequences — [here](#).)

Chris Sakalosky, vice president of Strategic Industries, Google Cloud: "Ginkgo's new protein LLM and open API mark a major step forward in making advanced AI tools accessible for drug discovery and biological research. By leveraging Google Cloud's infrastructure and AI capabilities, Ginkgo is empowering both enterprises and individual scientists to accelerate their work and drive innovation in the life sciences. Ginkgo is leading the way in democratizing access to cutting-edge AI models, to increase value to pharma companies using Ginkgo's platform and to ultimately help people live healthier lives."

Ginkgo has a multitude of models under development, spanning machine learning methods like language modeling and diffusion for conditional design. Ginkgo's first protein language model release will support two use-cases:

- Generation via Masked Language Modeling: given a sequence of amino acids with one or more <mask> tokens, the model will complete the sequence.
- Embedding calculation: Calculate the final hidden layer of the trained model to extract valuable representations for downstream tasks. To begin, Ginkgo's model returns the mean-pooled representation across the length axis.

Over the next year, Ginkgo will roll out more models and expand the API's capabilities, building a robust suite of tools that will enable you to solve complex problems in drug discovery, synthetic biology, genomics, and more using the latest machine learning methods. [Access the portal today and be among the first to explore our new API](#)

Ankit Gupta, General Manager of Ginkgo AI: "Flexibility is everything. Alongside our first proprietary model, which leverages unique datasets from Ginkgo, you'll also have access to publicly available models like [ESM2](#). This means you can explore and experiment with different approaches, all through a single streamlined platform. We're also deeply committed to making

advanced machine learning tools accessible, which is why our API comes with competitive pricing and a free tier. We've structured our costs to make it easy for you to jump in, experiment, and get predictions without worrying about high fees. Our initial models will have a free tier and our introductory pricing is approximately \$0.18 per million tokens. This means for a protein with around 500 amino acids, users should be able to get predictions on 2000 sequences for roughly 10 cents. In the age of generative biology, with engineers designing thousands to millions of sequences at a time, we hope to enable them with enormous amounts of computational scale."

About Ginkgo Bioworks

Ginkgo Bioworks is the leading horizontal platform for cell programming, providing flexible, end-to-end services that solve challenges for organizations across diverse markets, from food and agriculture to pharmaceuticals to industrial and specialty chemicals. Ginkgo Biosecurity is building and deploying the next-generation infrastructure and technologies that global leaders need to predict, detect, and respond to a wide variety of biological threats. For more information, visit ginkgobioworks.com and ginkgobiosecurity.com, read our [blog](#), or follow us on social media channels such as X (@Ginkgo and @Ginkgo_Biosec), Instagram (@GinkgoBioworks), Threads (@GinkgoBioworks), or [LinkedIn](#).

GINKGO BIOWORKS INVESTOR CONTACT:

investors@ginkgobioworks.com

GINKGO BIOWORKS MEDIA CONTACT:

press@ginkgobioworks.com

Forward-Looking Statements of Ginkgo Bioworks

This press release contains certain forward-looking statements within the meaning of the federal securities laws, including statements regarding the capabilities and potential success of the partnership, Ginkgo's model API and Ginkgo's cell programming platform. These forward-looking statements generally are identified by the words "believe," "can," "project," "potential," "expect," "anticipate," "estimate," "intend," "strategy," "future," "opportunity," "plan," "may," "should," "will," "would," "will be," "will continue," "will likely result," and similar expressions. Forward-looking statements are predictions, projections and other statements about future events that are based on current expectations and assumptions and, as a result, are subject to risks and uncertainties. Many factors could cause actual future events to differ materially from the forward-looking statements in this press release, including but not limited to: (i) volatility in the price of Ginkgo's securities due to a variety of factors, including changes in the competitive and highly regulated industries in which Ginkgo operates and plans to operate, variations in performance across competitors, and changes in laws and regulations affecting Ginkgo's business, (ii) the ability to implement business plans, forecasts, and other expectations, and to identify and realize additional business opportunities, (iii) the risk of downturns in demand for products using synthetic biology, (iv) the uncertainty regarding the demand for passive monitoring programs and biosecurity services, (v) changes to the biosecurity industry, including due to advancements in technology, emerging competition and evolution in industry demands, standards and regulations, (vi) the outcome of any pending or potential legal proceedings against Ginkgo, (vii) our ability to realize the expected benefits from and the success of our Foundry platform programs, (viii) our ability to successfully develop engineered cells, bioprocesses, data packages or other deliverables, and (ix) the product development or commercialization success of our customers. The foregoing list of factors is not exhaustive. You should carefully consider the foregoing factors and the other risks and uncertainties described in the "Risk Factors" section of Ginkgo's annual report on Form 10-K filed with the U.S. Securities and Exchange Commission (the "SEC") on February 29, 2024, Ginkgo's most recent quarterly report on Form 10-Q, and other documents filed by Ginkgo from time to time with the SEC. These filings identify and address other important risks and uncertainties that could cause actual events and results to differ materially from those contained in the forward-looking statements. Forward-looking statements speak only as of the date they are made. Readers are cautioned not to put undue reliance on forward-looking statements, and Ginkgo assumes no obligation and does not intend to update or revise these forward-looking statements, whether as a result of new information, future events, or otherwise. Ginkgo does not give any assurance that it will achieve its expectations.

Additional assets available online:



<https://www.googlecloudpresscorner.com/2024-09-17-Ginkgo-Bioworks-Launches-New-Protein-LLM-and-Model-API-Built-on-Google-Cloud-Technology>