

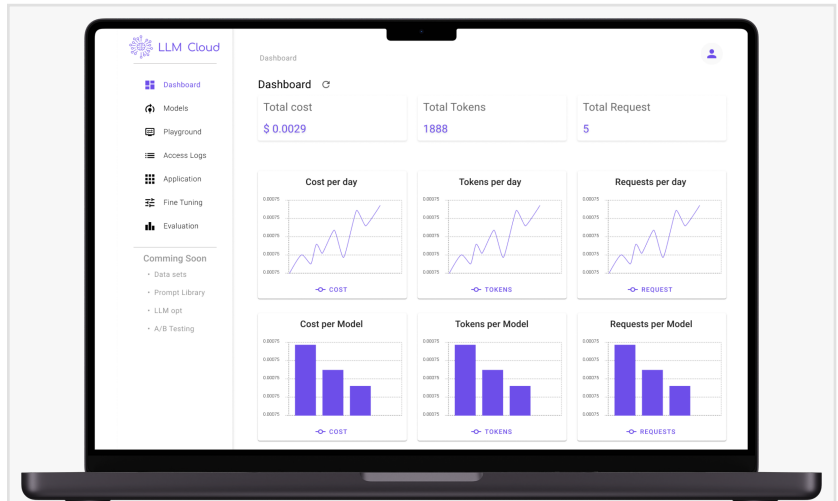
Predera launches LLMcloud.ai, A Fast Generative AI Cloud with Inference and Fine Tuning of Open Source LLMs

Predera introduces LLMcloud.ai, for streamlined access to Generative AI, with an expanding collection of open Large Language Models (LLM)

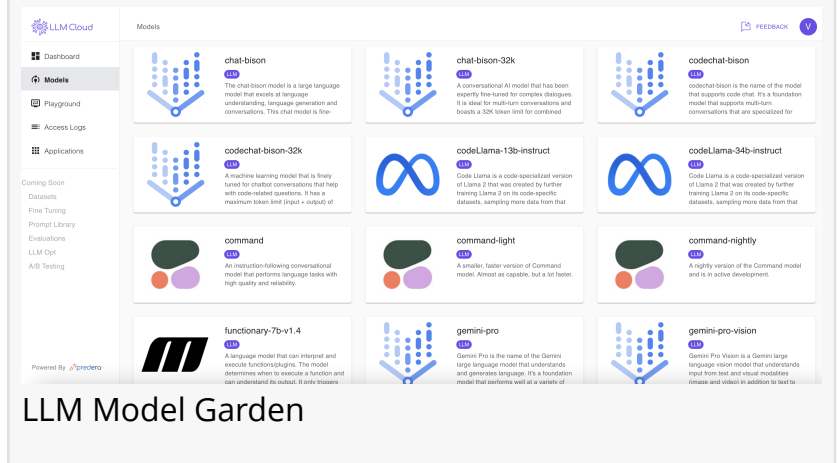
PLEASANTON, CALIFORNIA, USA, January 29, 2024 /EINPresswire.com/ -- [Predera](https://www.predera.com) Inc., an emerging leader in AI solutions, proudly announces the launch of [LLMcloud.ai](https://www.llmcloud.ai), an innovative Generative AI [cloud](https://www.llmcloud.ai) that streamlines deployment and fine-tuning of large language models (LLM), both from the open domain (Llama, Mistral, Llava etc) and from commercial providers (such as OpenAI, Cohere, Google).

LLMcloud.ai represents a significant leap forward in making advanced LLMs and Generative AI capabilities accessible to companies, enabling them to drive innovative products and operational effectiveness, securely. LLMcloud.ai introduces a full suite of advanced features to empower developers and business users to embrace Generative AI.

- **Unified API**:** A streamlined gateway offering a unified interface simplifies the access and integration of diverse LLMs, saving developers time and complexity in managing multiple models.
- **Open-Source LLM Garden**:** An expansive range of open-source LLMs, available via a one-click deployment feature that encourages rapid experimentation and prototyping in an enterprise's dedicated infrastructure.



Unified LLM Dashboard



LLM Model Garden

3. **Fine Tuning API**: Businesses can personalize LLM output with their unique data using a fast fine-tuning API, leading to sharper, more accurate AI insights. This module also supports continuous evaluations and benchmarking in response to base models and user feedback.

4. **Smart Model Router**: An intelligent API to dynamically direct user prompts to the ideal LLM based on specific requirements for cost, accuracy, or speed, ensuring an optimal balance of AI resource utilization.

5. **Multimodal Model Support**: LLMCloud.ai extends beyond text, accommodating AI that comprehends both textual and visual data. The addition of multimodal models opens up a broader spectrum of applications in media-rich industries.

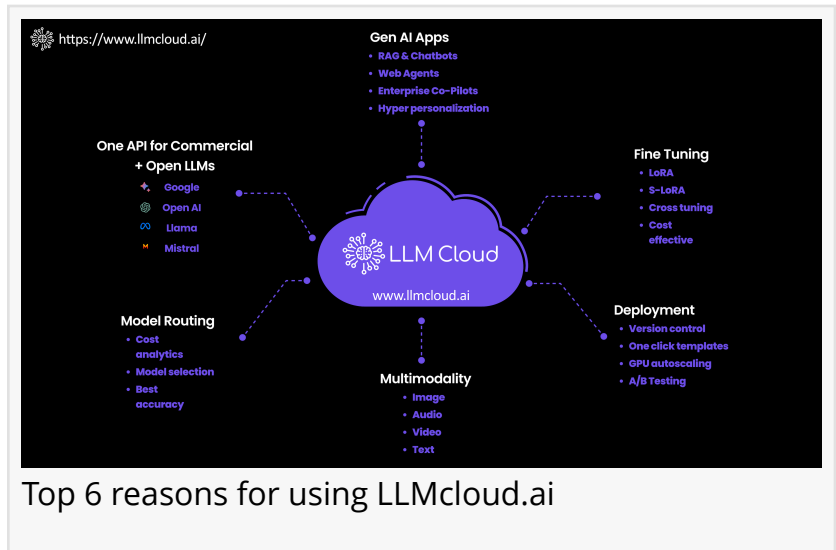
6. **Monitoring & Observability**: Real-time metrics on model performance, quality, and expenditure allow for strategic adjustments and resource allocation, helping you streamline AI operations.

7. **Gen AI Apps**: An expanding collection of app templates including RAG, Co-Pilot, and Chatbot frameworks, so businesses can rapidly develop and deploy AI-driven applications, significantly cutting down the route from development to launch.

8. **AI Agents**: Agents serve as the cognitive bridge between Generative AI systems and users, effectively 'bringing to life' enterprise LLM models. LLMCloud provides popular agents with customizable integration points allowing for real-time data and enhanced learning to deliver on enterprise automation.

Vamshi Ambati, the CEO of Predera, remarks, "With the introduction of LLMcloud.ai, we are building on our established expertise in MLOps and secure scaling of AI systems. This platform epitomizes our commitment to solving the 'last mile' challenge in Generative AI by removing complexities associated with exploring LLMs and AI app development in a private enterprise setup."

"LLMcloud.ai is not just a technological stride but an embodiment of responsible AI," adds Predera's Chief Technology Officer, Nazeer Shaik. "We've embedded the principles of Responsible AI into the platform's core to ensure our clients can rely on us for AI solutions that are secure, transparent, and fair."



Companies across various sectors can now leverage LLMcloud.ai to establish a competitive edge in the intelligence-powered market space. For more information on LLMcloud.ai's capabilities or to schedule a demonstration, please visit <https://www.llmcloud.ai/>

About Predera

Predera is a leading provider of AI-driven solutions, with a mission to help businesses leverage the power of AI/ML and Generative AI to drive growth, optimize operations, and reduce costs. Predera's flagship platform is a cloud-native, secure, MLOps & LLMOps platform that simplifies AI/ML/LLM life-cycle management for modern enterprises. With an innovative operations platform, industry-specific ML models, and high-quality solutions expertise, Predera has served customers in the Retail, Finance, Healthcare & Pharma industries. With the launch of LLMcloud.ai and its enhanced LLM features, Predera continues to be at the forefront of enabling business transformation in an AI-first way.

For more information on Predera's products and solutions, or to make inquiries, please visit - <https://www.predera.com/> or go to <https://www.llmcloud.ai/>

Press Communications

Predera

+1 669-289-9768

info@predera.com

This press release can be viewed online at: <https://www.einpresswire.com/article/684008942>

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.