IBM Plans to Make Llama 2 Available within its Watsonx AI and Data Platform

ARMONK, N.Y., Aug. 9, 2023 /PRNewswire/ -- As part of the continued roll-out of our enterprise-ready AI and data platform, watsonx, IBM (NYSE: IBM) plans to host Meta's Llama 2-chat 70 billion parameter model in the watsonx.ai studio, with early access now available to select clients and partners. This will build on IBM's collaboration with Meta on open innovation for AI, including work with open source projects developed by Meta – such as the PyTorch machine learning framework and the Presto query engine used in watsonx.data.

This will also support IBM's strategy of offering both third-party and its own AI models. Currently in watsonx.ai, AI builders can leverage models from IBM and the Hugging Face community, which are pre-trained to support a range of Natural Language Processing (NLP) tasks including question answering, content generation and summarization, text classification and extraction.

The future availability of Llama 2 in watsonx.ai will be another milestone on IBM's generative AI roadmap and is expected to be followed by the coming releases of its AI Tuning Studio, additional AI models in watsonx.ai, and FactSheets in watsonx.governance.

IBM is committed to keeping trust and security principles at the forefront as it continues to roll out its generative AI capabilities. For instance, when users run the Llama 2 model through the prompt lab in watsonx.ai, they can toggle on the AI guardrails function to help automatically remove harmful language from the input prompt text as well as the output generated by the model. Meta also provides an account of their fine-tuning methodology used in their large language models.

Furthermore, IBM Consulting has the expertise of 21,000 data, AI and automation consultants in addition to its Center of Excellence for Generative AI comprised of more than 1,000 consultants with specialized generative AI expertise. These experts can work with clients to help tune and operationalize models for targeted use cases aligned to their specific business requirements.

IBM, like Meta, is a supporter of open innovation. There is value in engaging a robust and diverse community of AI builders and researchers to test, share feedback and collaborate on these technologies to drive further innovation. We are excited to see what these innovators will build with Llama 2 and other models on the

watsonx platform.

Statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

AI builders can explore watsonx.ai today via the free trial: https://www.ibm.com/products/watsonx-ai

About IBM

IBM is a leading global hybrid cloud and AI, and business services provider, helping clients in more than 175 countries capitalize on insights from their data, streamline business processes, reduce costs and gain the competitive edge in their industries. Nearly 3,800 government and corporate entities in critical infrastructure areas such as financial services, telecommunications and healthcare rely on IBM's hybrid cloud platform and Red Hat OpenShift to affect their digital transformations quickly, efficiently and securely. IBM's breakthrough innovations in AI, quantum computing, industry-specific cloud solutions and business services deliver open and flexible options to our clients. All of this is backed by IBM's legendary commitment to trust, transparency, responsibility, inclusivity and service. For more information, visit www.ibm.com.

Contact:
Amy Angelini
alangeli@us.ibm.com

Additional assets available online: Photos (

https://newsroom.ibm.com/2023-08-09-IBM-Plans-to-Make-Llama-2-Available-within-its-Watsonx-AI-and-Data-Platform