IBM Advances watsonx AI and Data Platform with Tech Preview for watsonx.governance and Planned Release of New Models and Generative AI in watsonx.data

- Watsonx AI and data platform to include new generative AI models and enhancements designed to help enterprises scale AI

- Capabilities on display for thousands of developers at IBM TechXchange Conference Sept. 11-14 in Las Vegas

ARMONK, N.Y., Sept. 7, 2023 /PRNewswire/ -- IBM (NYSE: IBM) today announced plans for new generative AI foundation models and enhancements coming to watsonx – its AI and data platform with a set of AI capabilities designed to help enterprises scale and accelerate the impact of AI. These enhancements include a technical preview for watsonx.governance, new generative AI data services coming to watsonx.data, and the planned integration of watsonx.ai foundation models across select software and infrastructure products.

Developers will be able to get their hands on many of these new capabilities and models September 11-14 at TechXchange, IBM's premier technical learning event in Las Vegas.

The new IBM and third-party generative AI models coming to watsonx.ai include:

- Granite series models: IBM plans to introduce its Granite series models later this month. The Granite models use the "Decoder" architecture, which underpins the ability of today's large language models (LLMs) to predict the next word in a sequence, and can support enterprise NLP tasks, such as summarization, content generation and insight extraction. IBM plans to provide a list of the sources of data as well as a description of the data processing and filtering steps that were performed to produce the training data for the Granite series of models. (Planned availability Q3 2023)
- **Third-party models:** IBM is now offering Meta's Llama 2-chat 70 billion parameter model and the StarCoder LLM for code generation in watsonx.ai on IBM Cloud. (Available now)

IBM has established a training process for its foundation models – centered on principles of trust and transparency – that starts with rigorous data collection and ends in control points for enabling responsible deployments of models and applications for governance, risk assessment, privacy concerns, bias mitigation, and compliance.

IBM is also announcing plans to launch new capabilities across the watsonx platform.

Watsonx.ai:

- **Tuning studio**: IBM plans to release the first iteration of its Tuning Studio, which will include prompt tuning – an efficient, low-cost way for clients to adapt foundation models to their unique downstream tasks with their own enterprise data. (Planned availability 3Q23)

- **Synthetic data generator**: IBM today launched synthetic data generator to assist users in creating artificial tabular data sets from custom data schemas or internal data sets. This will allow users to extract insights for AI model training with reduced risk, thereby augmenting decision making and accelerating time to market. (Available now)

Watsonx.data:

- **Generative AI:** IBM plans to infuse watsonx.ai generative AI capabilities in watsonx.data to help users discover, augment, visualize, and refine data for AI through a self-service experience powered by a conversational, natural language interface. (Planned tech preview 4Q 2023)

- **Vector database capability:** IBM plans to integrate a vector database capability into watsonx.data to support watsonx.ai retrieval augmented generation use cases. (Planned tech preview 4Q 2023)

Watsonx.governance:

- **Model risk governance for generative AI:** IBM is launching a tech preview for watsonx.governance. Clients in the tech preview will be able to explore capabilities for automated collection and documentation of foundation model details and model risk governance capabilities that allow stakeholders to view relevant metrics in dashboards of their enterprise-wide AI workflows with approvals' so humans are engaged at the right times.

"As demonstrated by the ongoing rollout of the watsonx platform within just a few months since launch, we are here to support clients through the entire AI lifecycle" said Dinesh Nirmal, Senior Vice President, Products, IBM Software. "As a transformation partner, IBM is collaborating with clients to help them scale AI in a trustworthy way – from helping to institute foundational elements of their data strategies to tuning models for their specific business uses cases to helping them govern models beyond that."

The IBM watsonx AI and data platform will be complemented by a set of AI assistants designed to help clients scale and accelerate the impact of AI with their trusted data across key enterprise use cases, such as:

- **Application modernization:** IBM watsonx Code Assistant products, coming later this year, will use tailored foundation models to transform code and generate code recommendations for developers. Recently, IBM announced two upcoming AI-assisted code products– watsonx Code Assistant for Z to enhance developer productivity and accelerate COBOL application modernization and watsonx Code Assistant for Red Hat Ansible Lightspeed to help developers of all levels write Ansible playbooks.

- **Customer care:** IBM watsonx Assistant will help deliver consistent and intelligent customer service solutions with conversational AI. For instance, IBM Support Insights Pro, which is expected to be available later this month, will use watsonx Assistant to help clients find insights in their multivendor IT infrastructures to proactively assess support patterns and remediate risks, resulting in higher availability and security.

- **HR & Talent:** IBM watsonx Orchestrate will help HR professionals automate repetitive, high-friction tasks and back-office

processes such as interview scheduling or posting open jobs, through a conversational interface.

IBM is also planning to embed watsonx.ai innovations across its hybrid cloud software and infrastructure products, including:

- **Intelligent IT automation:** Entering tech preview next week, IT Automation products Instana and AIOps Insights will include Intelligent Remediation, which embeds watsonx.ai generative AI foundation models to assist IT Ops practitioners with summarization of incident details, as well as provide prescriptive workflow suggestions to help engineers quickly implement solutions.
- **Developer services for watsonx**: To help simplify and accelerate developers' ability to bring watsonx capabilities closer to their companies' data on IBM Power for SAP workloads, SAP ABAP SDK for watsonx is expected to add to the ways clients can use AI to inference near their data on Power systems and deploy AI algorithms on their most sensitive data and transactions. (Planned availability 1Q24)

For more information and demonstrations of these products, be sure to tune into live sessions at TechXchange next week.

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

About IBM

IBM is a leading provider of global hybrid cloud and AI, and consulting expertise. We help 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. More than 4,000 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 effect their digital transformations quickly, efficiently, and securely. IBM's breakthrough innovations in AI, quantum computing, industry-specific cloud solutions and consulting deliver open and flexible options to our clients. All of this is backed by IBM's long-standing commitment to trust, transparency, responsibility, inclusivity, and service.

Visit www.ibm.com for more information.

Media Contact:

Amy Angelini
alangeli@us.ibm.com

SOURCE IBM