IBM Unveils watsonx Generative AI Capabilities to Accelerate Mainframe Application Modernization

- IBM is designing this product to refactor, transform, and validate COBOL code to help speed time-to-value and augment skills for critical application modernization

- The product will be enabled by a 20 billion parameter large language model (LLM) for code

ARMONK, N.Y., Aug. 22, 2023 /PRNewswire/ -- IBM (NYSE: IBM) today announced watsonx Code Assistant for Z, a new generative AI-assisted product that will help enable faster translation of COBOL to Java and enhances developer productivity on the platform. This product will be generally available in Q4 2023, and is being designed to help accelerate COBOL application modernization. Watsonx Code Assistant for Z will preview during TechXchange, IBM's premier technical learning event in Las Vegas, Sept 11-13.

Watsonx Code Assistant for Z is a new addition to the watsonx Code Assistant product family, along with IBM watsonx Code Assistant for Red Hat Ansible Lightspeed, scheduled for release later this year. These solutions will be powered by IBM's watsonx.ai code model, which will have knowledge of 115 coding languages[1] having learned from 1.5 trillion tokens.[2] At 20 billion parameters, it is on target to become one of the largest generative AI foundation models for code automation.[3] The watsonx Code Assistant product portfolio will extend over time to address other programming languages, to improve time to value for modernization and address growing skills challenges for developers.

Watsonx Code Assistant for Z is being designed to assist businesses in leveraging generative AI and automated tooling to accelerate their mainframe application modernization – all with the goal of preserving the performance, security and resiliency capabilities of IBM Z.

The COBOL data processing language supports many vital business and operational processes at organizations globally. At scale, using watsonx Code Assistant for Z in comparison to other approaches could make it easier for developers to selectively and incrementally transform COBOL

business services into well architected high-quality Java code – with estimated billions of lines of COBOL code as potential candidates for targeted modernization over time. Generative AI can help developers to more quickly assess, update, validate and test the right code, allowing them to more efficiently modernize large applications and focus on higher impact tasks.

IBM is designing these capabilities to provide tooling for each step of the modernization journey. The solution is expected to include IBM's Application Discovery and Delivery Intelligence (ADDI) inventory and analysis tool. Following ADDI, key steps on the journey include refactoring business services in COBOL, transforming COBOL code to Java code with an optimized design, and validating the resulting outcome, including using automated testing capabilities. Potential benefits for clients include:

- Accelerating code development and increasing developer productivity throughout the application modernization lifecycle
- Managing total cost, complexity, and risk of application modernization initiatives, including translation and optimization of code in-place
- Expanding access to a broader pool of IT skills and accelerating developer onboarding
- Achieving high quality, easy to maintain code through model customization and the application of best practices

"Our collaboration with IBM is an important element in our drive to leverage generative AI interfaces to challenge legacy approaches with material productivity gains, and reinvent our Capital Markets solutions," said Roger Burkhardt, CTO, Capital Markets and AI, Broadridge Financial. "We have had excellent client response to our generative AI investments and we are intrigued by the opportunity to further our efforts by leveraging IBM watsonx Code Assistant for Z to address a broader range of platforms."

AI-assisted mainframe application modernization is an imperative

According to new research from the IBM Institute for Business Value, organizations are 12x more likely to leverage existing mainframe assets rather than rebuild their application estates from scratch in the next two years. At the same time, however, the study shows that the number one challenge for those same organizations is a lack of resources and skills.

"By bringing generative AI capabilities through watsonx to new use cases, we plan to drive real progress for our clients," said Kareem Yusuf, PhD, Senior Vice President, Product Management and Growth, IBM Software. "IBM is engineering watsonx Code Assistant for Z to take a targeted and optimized approach. It's built to rapidly and accurately convert code optimized for IBM Z, accelerate time to market and broaden the skills pool. This can help enhance applications and add new capabilities while preserving the performance, resiliency, and security inherent in IBM Z."

There are many application modernization approaches available today. Some options include rewriting all application code in Java, or migrating everything to public cloud, which may sacrifice capabilities that are core to the IBM Z value proposition while failing to deliver on expected cost reduction. Tools that convert COBOL applications to Java syntax can produce code that is hard to maintain and can be unrecognizable to a Java developer. Generative AI is promising, but current AI-assisted partial re-write technology lacks COBOL support and doesn't optimize the resulting Java code for the given task.

The resulting Java code from watsonx Code Assistant for Z will be object-oriented. IBM is designing this solution to be optimized to interoperate with the rest of the COBOL application, with CICS, IMS, DB2, and other z/OS runtimes. Java on Z is designed to be performance-optimized versus a compared x86 platform.[4]

Building on a foundation of governance and innovation

According to a 2023 Gartner® report (For Gartner Subscribers only), "by 2028, the combination of humans and AI assistants working in tandem could reduce the time to complete coding tasks by 30%." The report further states that "the use of AI code generation tools is not replacing the quality assurance (QA) processes and security controls that are needed by developers for robust and secure product development, as well as for mitigation of inherited risks from using generative methods for code."[5]

Protecting sensitive data and customer intellectual property are critical when it comes to implementing generative AI. IBM for decades has followed core principles, grounded in commitments to Trust and Transparency. With this principle-based approach, the watsonx platform aims to enable enterprises to leverage their own trusted data and IP to build tailored AI solutions that are scalable across operations.

Additionally, IBM Consulting brings deep domain expertise in IBM Z application modernization with a focus on guiding clients that leverage the platform across key industries such as banking, insurance, healthcare and government. These dedicated consultants can help clients identify the right application areas to modernize in order to optimize the potential benefits of watsonx Code Assistant for Z.

For more information about AI-assisted mainframe application modernization, and to get started with IBM's optimized, targeted approach, please visit our website here and join us at TechXchange. Register today for our watsonx Code Assistant for Z webinar on Sept. 21 at 11 am ET here and

learn how IBM is bringing Gen AI to mainframe application modernization. You can also schedule a live demo with our team here.

IBM's plans, directions, and intentions may change or be withdrawn at any time at IBM's discretion without notice. Information about potential future products and improvements is provided to give a general idea of IBM's goals and objectives and should not be used in making a purchase decision. IBM is not obligated to provide any material, code, or functionality based on this information.

[5] *Gartner, Emerging Tech: Generative AI Code Assistants Are Becoming Essential to Developer Experience, By Radu Miclaus, Arun Chandrasekaran, Ray Valdes, Mark Driver, Eric Goodness, Published 11 May 2023*

GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.

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 affect 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 legendary commitment to trust, transparency, responsibility, inclusivity and service.

Media Contact:
Ashley Peterson
ashley.peterson@ibm.com

[1] List of coding languages used in the model can be found here: https://github.ibm.com/ai-models-architectures/Granite-Megatron-LM/blob/starcoder-experiments/sampling_proportions/starcoderdata-90/proportions.txt

[2] Number based on proprietary internal data.

[3] Previous largest was community model StarCoder at 15B https://www.marktechpost.com/2023/05/07/meet-starcoder-the-biggest-open-source-large-language-models-for-code/

[4] For example, as noted during the announcement of IBM z16, using IBM Semeru Runtime Certified Edition 11, run Business Rules Processing with IBM Operational Decision Manager 8.11.00 on Linux on IBM z16 for up to 70% higher throughput per core   versus running the same application on a compared x86 server. DISCLAIMER: Performance results are based on the average of measurements done using IBM Operational Decision Manager (ODM) 8.11.0 with IBM Java 8.0.7.10 and IBM Semeru Runtime Certified Edition 11.0.15.0 on IBM z16 and on a compared x86 server. Two different configurations were tested: executing 2005 rules (from a ruleset containing 14560 rules), and executing 80 rules (from a ruleset containing 300 rules). IBM z16 configuration: Linux on IBM Z LPAR with Red Hat Enterprise Linux 8.5 (Ootpa) and 4 IFLs (SMT). x86 server configuration: Red Hat Enterprise Linux release 8.6 (Ootpa) and 4 SMT-2 cores (Cascade Lake Intel(R) Xeon(R) Gold 6226R CPU @ 2.90GHz). Results may vary.

SOURCE IBM

Additional assets available online: ◼ Video (+)

https://newsroom.ibm.com/2023-08-22-IBM-Unveils-watsonx-Generative-AI-Capabilities-to-Accelerate-Mainframe-Application-Modernization?_gl=1%2A1jcyvyk%2A_ga%2ANDA3MTg1NDEyLjE2OTI3MjE0NTM.%2A_ga_FYECCCS21D%2AMTY5MjcyMTQ1Mi4xLjEuMTY5MjcyMTQ2OC4wLjAuMA..