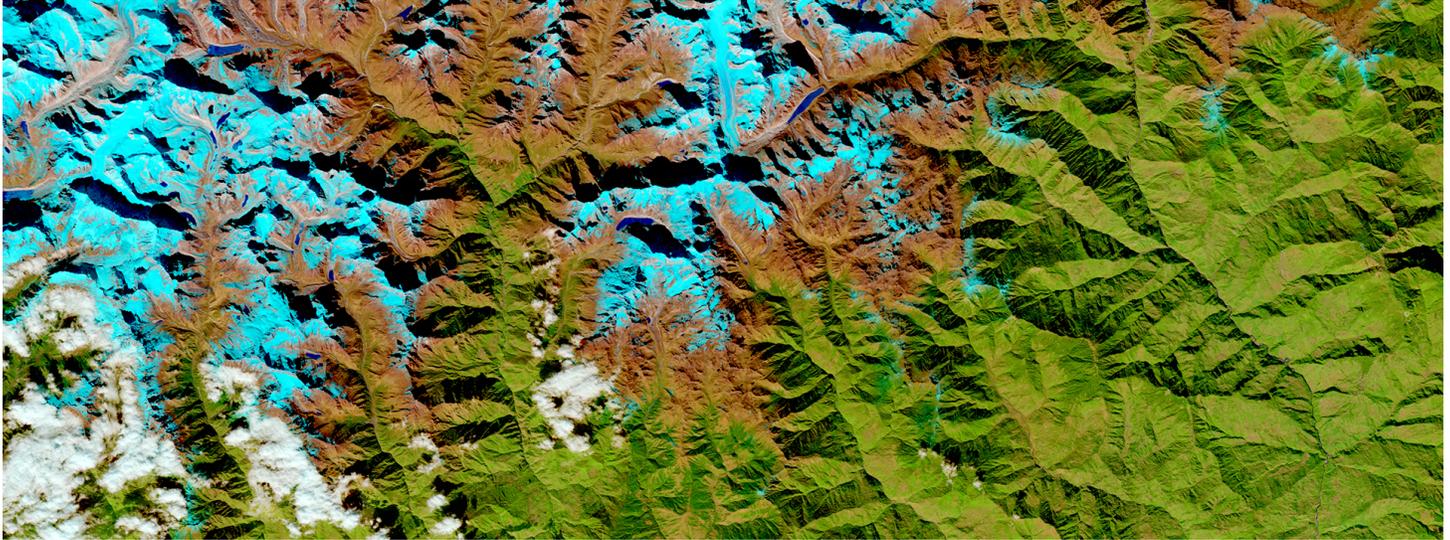# IBM and NASA Open Source Largest Geospatial AI Foundation Model on Hugging Face

Effort aims to widen access to NASA earth science data for geospatial intelligence and accelerate climate-related discoveries

YORKTOWN HEIGHTS, N.Y., Aug. 3, 2023 /PRNewswire/ -- IBM (NYSE: IBM) and open-source AI platform Hugging Face today announced that IBM's watsonx.ai geospatial foundation model – built from NASA's satellite data – will now be openly available on Hugging Face. It will be the largest geospatial foundation model on Hugging Face and the first-ever open-source AI foundation model built in collaboration with NASA.

Access to the latest data remains a significant challenge in climate science where environmental conditions change almost daily. And, despite growing amounts of data — estimates from NASA suggest that by 2024, scientists will have 250,000 terabytes of data from new missions — scientists and researchers still face obstacles in analyzing these large datasets. As part of a Space Act Agreement with NASA, IBM set out earlier this year to build an AI foundation model for geospatial data. And now, by making a geospatial foundation model available via Hugging Face — a recognized leader in open-source and a well-known repository for all transformer models — efforts can advance to democratize access and application of AI to generate new innovations in climate and Earth science.

"The essential role of open-source technologies to accelerate critical areas of discovery such as climate change has never been clearer," said Sriram Raghavan, Vice President, IBM Research AI. "By combining IBM's foundation model efforts aimed at creating flexible, reusable AI systems with NASA's repository of Earth-satellite data, and making it available on the leading open-source AI platform, Hugging Face, we can leverage the power of collaboration to implement faster and more impactful solutions that will improve our planet."

"AI remains a science-driven field, and science can only progress through information sharing and collaboration," said Jeff Boudier, head of product and growth at Hugging Face. "This is why open-source AI and the open release of models and datasets are so fundamental to the continued progress of AI, and making sure the technology will benefit as many people as possible."

"We believe that foundation models have the potential to change the way observational data is analyzed and help us to better

understand our planet," said Kevin Murphy, Chief Science Data Officer, NASA. "And by open sourcing such models and making them available to the world, we hope to multiply their impact."

The model – trained jointly by IBM and NASA on Harmonized Landsat Sentinel-2 satellite data (HLS) over one year across the continental United States and fine-tuned on labeled data for flood and burn scar mapping — has demonstrated to date a 15 percent improvement over state-of-the-art techniques using half as much labeled data. With additional fine tuning, the base model can be redeployed for tasks like tracking deforestation, predicting crop yields, or detecting and monitoring greenhouse gasses. IBM and NASA researchers are also working with Clark University to adapt the model for applications such as time-series segmentation and similarity research.

The news follows IBM's announcement earlier this year to collaborate with NASA to build an AI model that could speed up the analysis of satellite images and boost scientific discovery. It's also part of NASA's decade-long Open-Source Science Initiative to build a more accessible, inclusive, and collaborative scientific community. NASA, along with the White House and other federal agencies, has declared 2023 a Year of Open Science to celebrate the benefits and successes created through the open sharing of data, information, and knowledge.

The model leverages IBM foundation model technology and is part of IBM's larger effort to create and train AI models that can be used for different tasks and apply information from one situation to another. In July, IBM announced the availability of watsonx, an AI and data platform that allows enterprises to scale and accelerate impact of the most advanced AI with trusted data. A commercial version of the geospatial model, which is part of IBM watsonx, will be available through the IBM Environmental Intelligence Suite (EIS) later this year.

For more information about this collaboration, visit the IBM Research Blog

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 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.

Visit www.ibm.com for more information.

Contacts:

Sarah Benchaita
IBM Research
sarah.benchaita@ibm.com

Bethany Hill McCarthy

IBM Research

Bethany@ibm.com

SOURCE IBM