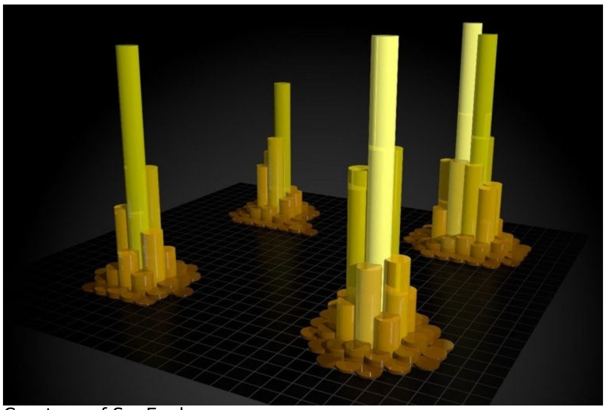
Boosting Scientific Research Using FAIR AI Models

Posted by Okachinepa 10/16/2024



Courtesy of SynEvol
Credit: Argonne Leadership Computing Facility Visualization and Data Analytics Group

The findable, accessible, interoperable, and reusable (FAIR) principles were first put forth by researchers to outline recommended practices for increasing the use of datasets by researchers and machines. These ideas have now been modified for usage with scientific datasets and research software with the goal of improving research transparency, reproducibility, and reusability while also encouraging software reuse as opposed to redevelopment.

These concepts are currently also followed by artificial intelligence (AI) models, which use different digital assets such as advanced computing, research software, and databases. A collection of useful, succinct, and quantifiable FAIR principles designed especially for AI models is presented in a recent work. It goes on to explain how scientific discovery can be greatly accelerated by combining datasets and FAIR AI models.

In this paper, the FAIR principles for AI models are precisely defined and their use in a particular kind of advanced microscopy is demonstrated. In particular, it shows how FAIR datasets and AI models may be combined to describe materials at the Advanced Photon Source at Argonne National Laboratory (ANL) twice as quickly as using conventional methods.

The paper also emphasizes how accelerating scientific discovery can be achieved by tying the Argonne Leadership Computing Facility and ANL's Advanced Photon Source together. This approach reduces hardware disparities, enables researchers to speak the same AI language, and accelerates AI-driven discoveries. The adoption of these FAIR standards for AI models is expected to stimulate new relationships between data, AI models, and high-performance computing as well as propel the development of next-generation AI technologies.

In this study, researchers generated an undeformed bi-crystal gold sample at Argonne National Laboratory's Advanced Photon Source, resulting in a FAIR experimental dataset of Bragg diffraction peaks. The Materials Data Facility published this FAIR and AI-ready dataset.

Then, using the ThetaGPU supercomputer and the open-source API PyTorch, the researchers used this dataset to train three different types of AI models at the Argonne Leadership Computing Facility (ALCF): a traditional model; an NVIDIA TensorRT version of the traditional PyTorch AI model; and a model trained on the SambaNova DataScaleTM system at the ALCF AI Testbed. These AI models include uncertainty quantification metrics that make it evident whether predictions made by the AI are reliable.

The researchers' suggested FAIR principles for AI models were then followed by the publication of these three distinct models in the Data and Learning Hub for Science. Then, using the ALCF's ThetaGPU supercomputer, they connected all of these resources, FAIR AI models, and datasets to do repeatable AI-driven inference.

Globus is used for process orchestration, whereas Globus Compute is used for workflow execution. The researchers asked their University of Illinois colleagues to independently confirm that the results could be replicated after using software to automate this task.