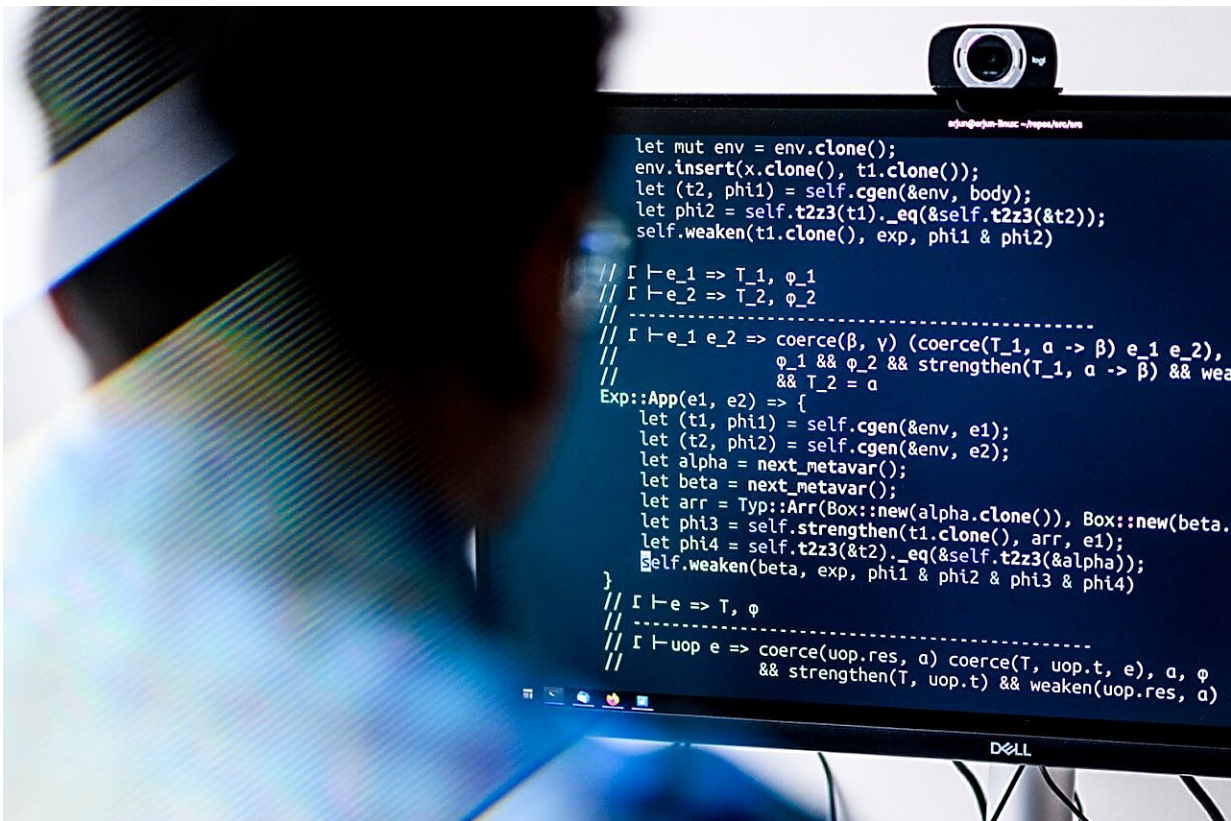


Responsible AI model for programmers being advanced by computer scientist

September 12 2023, by Alena Kuzub



Credit: Matthew Modoono/Northeastern University

Believing in open scientific collaboration on AI technology, a Northeastern professor joined others in creating a state-of-the-art open generative model for programmers that can be licensed and adapted for

different uses such as gaming and industrial automation.

Generative artificial intelligence and large language models have taken the world by storm in the last few years, says Arjun Guha, associate professor of computer science at Khoury College of Computer Sciences at Northeastern University. They are having a particularly significant impact on programming.

Computer scientists, programmers and smaller-market players, however, have very limited insight into the development process of these models, and that prevents them from developing a deeper understanding of the technology. It also excludes them from meaningful participation in its further expansion.

That is why Guha and his research group got heavily involved in the [BigCode project](#), launched by two private companies, Hugging Face and ServiceNow.

Hugging Face, a company that hosts a large open-source machine learning community, and ServiceNow, which helps businesses optimize technology solutions, teamed up to support individuals with professional AI research background in responsible development and use of open large language models for coding. They committed significant people and hardware resources to the project. As a result, StarCoder, a state-of-the-art, open generative [model](#) for programmers can be now licensed and adapted by others for different uses.

"You can spend an enormous amount of money building one of these things and not actually know if it's any good," Guha says.

The few multi-billion-dollar companies that have resources to build such learning models and "drop" them every now and then to stun the world, Guha says, are completely closed to the idea of sharing with the

community what this technology is capable of.

"If you ask the people who make them, 'What can I do with it?,' I think the answer they will always give you disingenuously is 'anything,' which is misleading," he says.

Guha believes that academic research has a role to play in shaping generative AI technology.

"An academic can come in and rigorously evaluate these things and say that here are its strengths and weaknesses. Yes, use it to do this, but please don't use it to do these other things without some serious guardrails," Guha says.

A much more pressing issue is people using this technology to make decisions that impact other people, for example, about a loan application or a job opening.

"We should talk about when it is not appropriate to use these models, when they are doing more harm than good," he says.

Guha dedicated a lot of energy to BigCode, which launched in September 2022, he says, leading a working group that focused on evaluating the open models, StarCoder and SantaCoder, created by the project.

Building an LLM first requires identifying the data that will be fed into the model to train it. When the model has been trained, Guha says, it should be evaluated on what it can and cannot actually do.

The models created by the BigCode project were trained at the Hugging Face cluster. Guha's group evaluated the majority of them at the Northeastern Discovery cluster at the Massachusetts Green High

Performance Computing Center, a high-powered parallel computing system that incorporates cutting-edge computing technologies and robust storage solutions.

They conducted an extensive evaluation in 19 different programming languages to understand the capabilities of the models.

"When this project launched, one of the goals was to have it work on lots and lots of languages to make several communities happy," Guha says.

The models were tested to implement such tasks as producing code from natural language descriptions, documenting code and predicting type annotations.

Other researchers carried out other analyses, such as a bias and toxicity analysis that showed that since the coding model was not trained on vast internet data, it consumed less toxic content and was not likely to produce toxic output.

Guha says the StarCoder model underwent the most extensive evaluation that ever occurred for a focused LLM, because of the massive collaborative nature of the BigCode project.

"It's been a great project that brought together a lot of researchers at various stages in their careers," he says.

The paper that came out of this part of the BigCode project in May had almost 70 co-authors. Several doctoral students and undergraduates, Guha says, were able to contribute to the model.

Anyone now can request to download and use StarCoder or SantaBase for free for research, commercial or non-commercial purposes as long as they sign the BigCode Open Responsible AI Licenses agreement and

follow restrictions that apply, including to the modified material.

For example, Guha is collaborating with MathWorks, a corporation that specializes in mathematical computing software for engineers and scientists, and Roblox, an online global game platform, on exploring how they could use StarCoder, bring it in-house and customize to their needs.

A number of researchers are using the model as well, Guha says.

The BigCode project is very transparent and explicit, Guha says, about what data its models are using. People can file a request if they want the project to stop using their data. So far, only a couple dozen people have done so.

BigCode is ramping up for the next round of the [project](#) and expects to make announcements on further developments soon.

Provided by Northeastern University

Citation: Responsible AI model for programmers being advanced by computer scientist (2023, September 12) retrieved 7 May 2024 from <https://techxplore.com/news/2023-09-responsible-ai-programmers-advanced-scientist.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
