

On the paper “Exploring the MIT Mathematics and EECS Curriculum Using Large Language Models”

On June 15th, Iddo Drori posted on arXiv a working paper associated with a dataset of exams and assignments from dozens of MIT courses. He did so without the consent of many of his co-authors and despite having been told of problems that should be corrected before publication. One of us only learned of the posting on Sunday June 18 after traveling over the weekend.

In the process of addressing this matter, we discovered that contrary to what Iddo Drori had conveyed to us and to the students collecting data for the project, Iddo did not have permission from all the instructors to collect the assignment and exam questions that made up the dataset that was the subject of the paper. Instructors for some of those courses only learned of the existence of this dataset and the inclusion of their course material in it when the paper appeared on social media and when Iddo posted samples of the data online without permission from anyone.

These are serious matters that are being addressed through institutional channels, so we did not take lightly making such a public statement about them, but we feel it is important to explain why this paper should never have been published and must be withdrawn. We have asked Iddo to withdraw the paper from arXiv and have also contacted arXiv directly explaining the situation.

We want to emphasize that all the student authors in this paper worked really hard on what could have been a very interesting and valuable paper had the data been collected with consent. The many problems with the published work were not the fault of the students.

And no, GPT-4 cannot get an MIT degree.

Armando Solar-Lezama, Professor in EECS and COO and Associate Director of CSAIL, MIT

Tonio Buonassisi, Professor of Mechanical Engineering, MIT

Yoon Kim, Assistant Professor in EECS and CSAIL, MIT