The Need for New "AI Telescopes"

Pierre Baldi

University of California, Irvine

For centuries universities have been at the cutting edge of research in the natural sciences and the technologies derived from them. However, today universities are unable to train and study the most advanced forms of Artificial Intelligence (AI) at scale, including large language models with trillions of parameters, such as GPT-4. Universities' main problem is computational. No university has the in-house computational resources, primarily in the form of large clusters of Graphical Processing Units (GPUs) needed to train the most powerful AI models. Here we call for a very large-scale effort to address this dangerous situation.

A few people think this problem can be brushed away. There are still plenty of problems for academics to work on at the fringes—for instance by training much smaller models or fine-tuning and aligning existing models using specific data. Furthermore, this situation has happened before in the world of technology—after all, universities cannot build large ships, airplanes, or nuclear plants. Finally, the economic reality is that the cost and infrastructure required for building such models are prohibitive for academic institutions: GPT-4 is rumored to cost north of 100 million dollars to train over a period of many months using tens of thousands of GPUs.

Although there is wisdom in these arguments about resources and scale, they fail to recognize the ambiguous nature of AI as being not only a formidable technology but also a fundamental topic in the natural sciences and the study of our place in the universe. Telescopes and microscopes taught us that we are not central to the universe of stars or of living systems. Computers are teaching us that we are not central to the universe of intelligence. And in the same way that the science of fluid dynamics is fundamental for understanding and building safe airplanes, the science of AI is fundamental for understanding and building safe AI systems. The stakes related to AI are so fundamental that it seems dangerous to leave them exclusively to companies. No one knows for sure, but we may be reasonably close to being able to create intelligence that is greater than human intelligence. Do we want companies to be the first and only ones to peer into and exploit superintelligence?

How can we address this fundamental problem? To enable scientific AI research at scale, there have been proposals for distributed approaches with, for instance, small clusters of elite universities sharing resources around computer centers with budgets at most in the 10-100 million-dollar range. However, such proposals are not likely to work for at least two reasons. First, the scale of the computing resources that would be possible at such prices appears to be insufficient by at least one order of magnitude, and possibly more. Second, the proposed organizational structures are not likely to take advantage of the breadth of ideas and talent available throughout academia.

To address the problem, we need instead to ask whether there exist larger projects that have been undertaken by the scientific community. The answer is clearly yes, as evidenced by CERN's Large Hadron Collider (LHC) or NASA's James Webb Space Telescope (JWST). The lifetime cost of the JWST is projected to be around 10 billion dollars, and similar figures can be estimated for the LHC.

Thus, history shows that when political will and consensus are present, our society is capable of funding multi-billion-dollar scientific projects.

Thus, what is needed is at least one national or international multi-billion-dollar effort to build a new kind of AI "telescope", in the form of the largest data and computing center on Earth, that would enable scientists to build the most powerful AI systems and peer into the depths of the universe of intelligence while studying the related safety issues. This would enable both the science of AI and the application of AI to science at scale.

Ideally, like CERN, the effort should be international, both to share the cost burden among democratic nations but also to promote international cooperation and shared AI-safety standards. It is also possible that national security concerns could tilt the preferences toward a purely U.S. effort. In either case, the effort ought to be carried out by permanent staff in collaboration with thousands of participating academic laboratories organized in large collaborations, along the lines of those that are already routinely used in physics and astronomy and, in a somewhat different way, in biomedicine, to sequence the human genome or study complex diseases such as cancer.

No doubt such an approach would face formidable challenges. The first challenge is political, of course: the need to build enough consensus to start the project and allocate sufficient funding over the long haul against the current backdrop of the American elections, a divided Congress, and 30+ trillion dollars in national debt. International consensus would be even harder to secure, although the potential existential threats of AI ought to unite humanity beyond party lines or national boundaries.

The consensus challenge is heightened by the need for speedy execution at a time when AI capabilities are believed to be progressing at an exponential pace, notwithstanding the handicaps that universities have already accumulated. Speed of execution is likely to be important not only at the start of the project but throughout its duration, and thus careful consideration must be given to designing the decision processes within the new organization.

Unlike the LHC or the JWST, an AI telescope creates direct competition between academic scientists and industry. Thus, it may be wise to think about ways to also maximize synergies and buy-in from industry, in terms of cost and knowledge sharing, hardware, data sharing, personnel training, security issues, and so forth.

Finally, production poses challenges. Companies with large resources to train cutting-edge AI are also investing heavily in its production—that is, in serving it and monetizing it in ways that help sustain and accelerate their efforts. Careful thought will have to be allocated to studying how the cutting-edge AI produced by such a consortium could be used to serve humanity and whether it should be deployed and monetized in some form in part to help sustain the research effort. While these various challenges are formidable, none seem insurmountable. Furthermore, even in the worst-case scenarios, the majority of the budget would simply result in a positive injection into the economy through, for instance, hardware purchases and job training and creation.

Although creating a large-scale academic collaboration to investigate AI entails significant challenges, examples such as CERN and the JWST demonstrate that with political will, the national and international communities can come together to meet them. AI has now advanced to a

sufficient stage of sophistication that we broadly recognize the enormous opportunities and challenges it creates, including the potential to threaten human life across a range of fronts. The time may be ripe for us to coalesce around a large-scale scientific collaboration to build an AI telescope that would enable the development of the science of intelligence, and the study of the risks and potential of AI to harness it for our greater benefit.