Artificial intelligence (AI) for science is a growing area of interdisciplinary computer science research focused on solving some of the most pressing global issues. While many cite AI’s technical advances as the innovative force of the endeavor, I argue that interdisciplinarity, democratization, and cogent justification toward global citizens are driving forces to be fostered in the program’s development.

AI for science (AI4Science) reimagines the scientific discovery process to incorporate advances in artificial intelligence (AI), machine learning, and high-performance computing (HPC). To meet the challenges of the current generation, discovery needs to be significantly accelerated to create new materials for climate goals or rapid drug development for emergent pandemics and diseases. But should sheeer technological development through AI and HPC be imagined as the driving force toward this future?

To meet AI4Science’s goal, I argue three main hurdles need to be overcome and are necessary investment points: interdisciplinarity, democratization, and an orientation toward cogent justification. Any technical discoveries led by AI4Science must be supported by including disciplinary perspectives, the global scientific community, and a communicative stance to citizens. Without these hurdles met, discoveries will fail to reach the scale desired, the creativity needed, and the uptake sufficient for their success. The long path to democratizing science is based not only on principled FAIR data sharing or reproducibility currently underway but also requires investment in institutions and practices for interdisciplinary engagement where AI is an aide, not the focal point, for resource sharing and collective, open science. We should broadly interpret the AI as algorithmic thinking aimed at producing structures for collective intelligence—which science has long been known as from scientist Ludwik Fleck—rather than opaque algorithms that threaten interdisciplinarity, resource sharing, and most importantly, a communicative stance toward citizens and theory.

**Interdisciplinarity for science**

AI4Science broadly represents “the next generation of methods and scientific opportunities in computing, including the development and application of AI methods … to build models from data and to use these models … to advance scientific research.”1 The titular technology invokes the recent advances of AI. But scientific informatics laid the groundwork over 30 years ago through translational research.

Informatic disciplines such as cheminformatics or bioinformatics translate fuzzy and competing scientific theories between computerized schematic models. The informatician must satisfy the scientific theorists’ due to axiomatic assumptions and generalizations, challenge HPC workflows, and extract knowledge from the scientific community. For example, cheminformatics began by interpreting small molecules as graphs with nodes and edges as atoms and bonds. It is a reductive translation that loses out on many nuances of chemical theory. Interdisciplinary translation requires erasing, adding, modifying, hiding, mutating, and co-designing. Translation requires constant communication back and forth.

This classical assumption about molecules as graphs has produced many results and is promising for accelerating the future of molecular dynamics simulations. At the same time, complacency with the model has led many computer scientists toward only performing in silico validation. State of the art benchmark performance, reproducibility studies, novel architectures, and so forth are essential pieces of computer science research but often fail to meet and engage other disciplines. This is not to put down such research (which I continue to work in) but rather to reorient focus toward interdisciplinary engagement, which benchmark tasks can lose sight of. Metrics in learning problems often conceal the underlying workflow and its values. For example, in drug discovery, the real measure of a good model is that (1) identifies a molecule that is synthesizable, safe, and effective, (2) reduces experimental cost (false positives), and (3) comes packaged with a justification that is convincing to medicinal chemists, pharmaceutical companies, and the community at large to invest large sums of money into. The justification to the disciplines that control the gate for downstream drug design is often lost in pure benchmark machine learning tasks.

One exemplary project that has taken interdisciplinarity seriously, at least internal to the project’s development, is AlphaFold2.2 AlphaFold2’s protein folding capabilities have astounded the structural biology community. AlphaFold2 is deeply interdisciplinary, especially compared to other approaches. Rather than following a traditional machine learning paradigm based on constructing a deep neural network that predicts some measure from a set of features, AlphaFold2 builds on top of multiple sequence alignment (MSA) in an iterative algorithm. MSA is a technique for aligning protein sequences based on similarity to determine common structural motifs or understand phylogeny or homology and has a strong basis from the structural biology community.3

The engagement with a rich tradition of structural biology and the integration with
artificial intelligence has transformed structural biology and brought two communities exceptionally close together. By following the deep engagement with disciplinary knowledge and practice, AI4Science projects can open new discourse and terrains for discovery.

Science for All
The second challenge for a new scientific paradigm is democratization. Democratization of science requires accepting the values of inclusivity, decision-making equality, and deliberation—the latter factor demanding more than just open science alone. Computer science has, in some respects, produced great strides toward open science through open-source software. Regardless of someone’s status or physical location, software contributors are welcome to fix bugs, add features, or improve efficiency. Furthermore, the data science ethos of FAIR features, or improve efficiency. Further-
not be thinking about in assessing the experimental results, the fields will talk past each other.

Furthermore, the ability for citizens to engage in the scientific project outside of the scientific community has immense potential for addressing issues of trust in science. From high school students working on cheminformatics problems to data scientists analyzing data in their spare time, citizens can develop a feeling of authorship toward science as a global and inclusive project. Political philosopher Thomas Cristiano writes alienation from decision making “is like playing a game whose rules do not make any sense to one.” For citizens to first make sense of and second identify with the rules and practices of science, science needs to orient itself toward mutual justification. Citizen identification with science is part of the long road of developing trust in science, which is ultimately paramount to the political adoption of any global-scale scientific recommendations that emerge from AI4Science’s campaigns. Therefore, the role of mutual justification needs to be taken seriously, not just as a solution to the current conflict but as a sustainable end to pursue new and novel technologies for the next generation of global crises.

**Conclusion**

AI4Science is a promising endeavor not only for its premise of accelerated discovery and a new scientific revolution. AI4Science has a greater potential at expanding the collective intelligence of scientists through tightly integrated interdisciplinary research, inclusion of the global community through a community-driven ethos, and reorientation of justification and cogency. AI4Science is really accelerating the very factor that philosophers of science attribute to the core success of science thought collectives. Drug discovery during the COVID-19 pandemic has offered a glimpse into the future of tightly integrating computation and science through various collective and consortium projects, which ought to serve as a model for future grand challenges. With new technological paradigms on the horizon, we should step back and look at the ways in which these technologies not only have potential to accelerate discoveries in a vacuum but ultimately create and foster the social conditions necessary for the success of science, attainment of global inclusion, and mutual justification to all. Without these conditions at the forefront, a myopic focus on results themselves will not achieve the science we desperately need.

**DECLARATION OF INTERESTS**

The author has no competing interests to declare.

**REFERENCES**

1. Stevens, R., Taylor, V., Nichols, J., Maccabe, A.B., Yelick, K., and Brown, D. (2020). AI for Science: Report on the Department of Energy (DOE) Town Halls on Artificial Intelligence (AI) for Science, Report from the U.S. Department of Energy, February 2, 2020 https://doi.org/10.2172/1604756.

2. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589.

3. Kurovski, M.A., and Bujnicki, J.M. (2003). GeneSilico protein structure prediction meta-server. Nucleic Acids Res. 31, 3305–3307.

4. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3, 160018.

5. Woolley, A.W., Chabris, C.F., Pentland, A., Hashmi, N., and Malone, T.W. (2010). Evidence for a collective intelligence factor in the performance of human groups. Science 330, 686–688.

6. Longino, H.E. (1990). Science as Social Knowledge: Values and Objectivity in Scientific Inquiry (Princeton University Press).

7. Achdout, H., Almon, A., Bar-David, E., Barr, H., Ben-Shmuel, A., Bennett, J., Bobby, M.L., Brun, J., Sarma, B.V.N.B.S., Calmiano, M., et al.; The COVID Moonshot Consortium (2020). COVID Moonshot: open science discovery of SARS-CoV-2 main protease inhibitors by combining crowdsourcing, high-throughput experiments, computational simulations, and machine learning. bioRxiv. https://doi.org/10.1101/2020.10.29.339317.

8. Cristiano, T. (2008). The Constitution of Equality: Democratic Authority and its Limits (Oxford University Press).

**Austin Clyde** is a PhD candidate at the University of Chicago and an assistant computational scientist at Argonne National Laboratory. His research focuses on scaling drug discovery techniques to exascale computing infrastructure using artificial intelligence. He is a visiting research fellow at the Harvard Kennedy School’s Program on Science, Technology, and Society focusing on the intersection of AI, science, and democracy.