

## EAAI-22 Blue Sky Ideas in Artificial Intelligence Education from the AAAI/ACM SIGAI New and Future AI Educator Program

**Michael Guerzhoy** (University of Toronto; [guerzhoy@cs.toronto.edu](mailto:guerzhoy@cs.toronto.edu))

**Marion Neumann** (Washington University in St. Louis; [m.neumann@wustl.edu](mailto:m.neumann@wustl.edu))

**Emmanuel Johnson\*** (University of Southern California; [ejohnson@isi.edu](mailto:ejohnson@isi.edu))

**David Johnson\*** (Uppsala University; [david.johnson@im.uu.se](mailto:david.johnson@im.uu.se))

**Henry Chai\*** (Carnegie Mellon University; [hchai2@andrew.cmu.edu](mailto:hchai2@andrew.cmu.edu))

**Daniel Garijo\*** (Universidad Politécnica de Madrid; [daniel.garijo@upm.es](mailto:daniel.garijo@upm.es))

**Zhuoyue Lyu\*** (University of Toronto; [zhuoyue@cs.toronto.edu](mailto:zhuoyue@cs.toronto.edu))

**Christopher J. MacLellan\*** (Drexel University; [christopher.maclellan@drexel.edu](mailto:christopher.maclellan@drexel.edu))

DOI: [XX.XXXX/XXXXXXXX.XXXXXX](https://doi.org/10.1111/XXXXX.XXXXXX)

\*These authors contributed equally to this article.

### Introduction

The 12th Symposium on Educational Advances in Artificial Intelligence (EAAI–22, co-chaired by Michael Guerzhoy and Marion Neumann) continued the AAAI/ACM SIGAI New and Future AI Educator Program to support the training of early-career university faculty, secondary school faculty, and future educators (PhD candidates or postdocs who intend a career in academia). As part of the program, awardees were asked to address one of the following “blue sky” questions:

- How could/should AI courses incorporate AI Ethics into the curriculum?
- How could we teach AI topics at an early undergraduate or a secondary school level?
- AI has the potential for broad impact to numerous disciplines. How could we make AI education more interdisciplinary, specifically to benefit non-engineering fields?
- How should standard AI courses evolve?
- How could we leverage AI education to promote diversity in the field?

This paper is a collection of their responses, intended to help motivate discussion around these issues in AI education.

### AI education can be leveraged to promote diversity in the field through various approaches

**Emmanuel Johnson (Information Sciences Institute, University of Southern California)**

In this essay, I will focus on two of the most promising approaches. First, we must make AI education more accessible to diverse populations. For instance, there are 101 Historically Black Colleges and Universities [1]; however, less than 20 offer AI courses [2]. We must find ways to bring AI education to these HBCU students. One method to do this is through a Distributive Teaching Collaborative (DTC). Through a DTC, professors from institutions with a strong AI program partner with those at institutions without a strong AI program to leverage online learning platforms to co-teach courses in Artificial Intelligence. These courses are taught simultaneously where a professor at one institution may present the lecture while students at another institution can ask questions and engage with the professor lecturing. Through this method, students can also be paired in groups across institutions, and this type of learning provides benefits for all students. Students are exposed to classmates from different backgrounds who may approach problems in different ways. This could also help to reduce bias in AI systems that may arise from a lack of exposure to different experiences or test cases. In addition to making AI education more accessible, we must design AI curricula that help

students from diverse backgrounds develop the skills needed to solve their community's problems. The groups underrepresented in AI are often overrepresented in other fields. For example, African Americans are underrepresented in STEM majors but overrepresented in social work [3]. One possible explanation is that fields like social work are critical for community and political movements in low-income and minority communities [4]. Thus, students pursue these degrees in hopes of having an impact in their community. It is worth considering how artificial intelligence curriculum can address the needs of minority communities thus providing an avenue for attracting more students who have an interest in building their communities. Initiatives like AI for Social Good provide a model that we can follow. By building a more inclusive AI curriculum that empowers students from all backgrounds and providing access to such a curriculum, we can help to increase the diversity in artificial intelligence.

**References** [1] National Center for Education Statistics). (2020, January 10). Historically Black Colleges and universities and higher education desegregation. Home. Retrieved 01/11/2022 from <https://nces.ed.gov/fastfacts/display.asp?id=667>

[2] Earl, C. (n.d.). Artificial Intelligence at historically Black Colleges. ai-hbcu. Retrieved 01/11/2022 from <https://charlescearl.github.io/ai-hbcu/>

[3] Carnevale, A. P., Fasules, M. L., Porter, A., Landis-Santos, J. (2016). African Americans: College majors and earnings.

[4] Downs, K. (2016, February 10). African-Americans over-represented among low-paying college majors. PBS. Retrieved 01/11/2022 from <https://www.pbs.org/newshour/education/african-americans-over-represented-among-low-paying-college-majors>

## Embodied AI Ethics

### David Johnson (Uppsala University)

Democratizing AI can only be achieved by mainstreaming AI education and a key challenge is in lowering the barriers to understanding AI concepts. Something that we explore in our teaching of AI is not presuming of our students a deep understanding of computation or mathematics, and by conveying key AI concepts to students outside

of the computer – an “unplugged” approach. We submit that teaching AI ethics can be “unplugged” by putting students *in situ* and embodying the components of an AI/ML workflow. The constituent approaches are not entirely novel; an established approach to computer science education is in CS Unplugged [1] and there are several recent efforts that are self-labelled “AI Unplugged”, while embodied design has long been integrated into mathematics pedagogies [2].

The idea of embodying AI ethics is captured in a workshop that we run with our students about algorithmic accountability that we describe in brief here: Students work in teams of 4, where each team is tasked with developing a movie ranking service and are given a small example database of movies to analyse. Team members are set specific roles of a manager, data pre-processor, predictor, and a visualizer. The team, altogether, first determines what features to select from the example data and develops rules to apply to generate their ranking. Next, some new unseen data is distributed to each team, and each team manually processes it. The manager is responsible for collecting the data and signalling when the ranking is completed. The data pre-processor receives the raw data from the manager and manually copies it to a new table according to any agree feature selection and pre-processing. The predictor then receives the pre-processed table and applies the prediction rules to add a column with predictions. Finally, the visualizer takes this table from the predictor and copies it manually to a rank-ordered list of movie names that is passed back to the manager.

Each student in the team takes a physical role with which to fulfil the movie rank prediction workflow. When performing the workflow, each not only does some processing but also receives and sends data. The team embodies the workflow, and the individuals gain a unique experience (within their team) in their embodied role. Fairness and accountability are then discussed as a class relating to the data and algorithmic processing, where we would expect multiple viewpoints given the heterogeneous roles. This is just one example of learning AI ethics with such an approach. We plan to further develop a collection of embodied AI ethics activities with a view to evaluating them for inclusion in AI education curricula.

**References** [1] Bell, T., Alexander, J., Freeman, I., & Grimley, M. (2009). Computer science unplugged: School students doing real computing without computers. *The New Zealand Journal of Applied Computing and Information Technology*, 13(1), 20-29.

[2] Abrahamson, D., Nathan, M. J., Williams-Pierce, C., Walkington, C., Ottmar, E. R., Soto, H., & Alibali, M. W. (2020). *Frontiers in Education* 5, art. no. 147.

## Multidisciplinary Ethics in AI

**Henry Chai (Carnegie Mellon University)**

In my introduction to machine learning courses, I typically devote one lecture to fairness and bias in machine learning, usually near the end of the semester. However, recently I've been somewhat unsatisfied with these lectures: 80 minutes is just not enough time to cover such an important topic with any amount of meaningful depth. I find myself devoting most of that time to going over examples of bias in notable/interesting settings. These examples are really the highlight of that lecture: they resonate with students and make for easy engagement exercises, i.e., challenging students to identify which protected groups are at risk or how bias is introduced in each setting. Of course, I use real world examples in most if not all of my lectures to motivate the particular algorithm or method being taught: if examples are the most compelling way to discuss the ethical aspects of machine learning, then why not apply these bias and fairness exercises to all of those examples as well?

I'm excited to do exactly that this in future iterations of my introductory courses: instead of restricting the discussion of fairness and bias in machine learning to a single lecture, I intend to spread that content out over the semester and get students thinking about these issues regularly and in a variety of contexts. I plan to start early in the semester, in the very first lecture, where I introduce the basic notation and problem formulation of supervised machine learning using an example of a bank deciding who to extend credit and how much, an easy example with which to introduce ethics questions.

In terms of a grander, more pie-in-the-sky answer, I suspect most in the field would agree that a course in ethics should be required for all students, graduate and undergraduate, studying data science, artificial intelligence, machine learning or any related field. I strongly believe that these courses should be heavily interdisciplinary: at a minimum, they ought to include material drawn from the fields of philosophy and law to address questions like "what does it mean for an algorithm to be just or fair?" and "why should one value those traits in an algorithm?" In my personal experience, the most successful machine learning projects have involved domain experts and I believe the same approach should extend to how we teach all aspects of machine learning, including fairness and bias.

To facilitate the development of such courses, I envision the creation of an entity or organization that functions much like an institution's center for teaching and learning (indeed, such a group could be initially housed in these centers) expect with an exclusive focus on unifying and supporting the instruction of ethics. Notably, this governing body would not be restricted to just AI ethics but ethics across all subject areas and thus, would be inherently multidisciplinary. The roles that such an organization could take on are many but some possibilities include reviewing existing courses and suggesting places to incorporate ethics modules, developing and tailoring the aforementioned ethics modules to different domains as well as establishing guidelines and curricula for standalone ethics courses.

## Bridging the gap between AI courses and Open Science

**Daniel Garijo (Universidad Politécnica de Madrid)**

Due to the rise in popularity of Machine Learning (ML) techniques, universities around the globe have started incorporating Data Science as part of their courses. Frameworks such as ScikitLearn [1], TensorFlow [2] and Pytorch [3] have eased the adoption of ML tools, and are used by students to address many types of classification and regression problems.

AI courses usually provide a theoretical explanation for popular AI techniques, but major challenges regarding data integration, representation, modeling, collection and cleaning remain, in many cases, unaddressed. Based in my experiences at both the University of Southern California and Universidad Politécnica de Madrid, I believe AI-focused courses should evolve towards including:

1. **Basic data modeling and representation:** Data quality is paramount to ML applications. The Web is full of rich data that can be used to train ML algorithms, but that can be difficult to find and integrate together. Students should learn how to understand and represent the data they feed into their algorithms. Having a good notion in data modeling and representation helps expanding training corpora, which may significantly improve the final results of the ML algorithm chosen at the end.
2. **Best practices for reproducible results:** A data science project never ends just by training and testing a model. In order for the results to be useful, students should learn how to document the features, biases and limitations of

their projects, to record the input and parameters used in training and to preserve the computational environment used to perform the analysis. AI courses should incorporate best practices for reproducibility, in order to make the results of a model (and the model itself) Findable, Accessible, Interoperable and Reusable (FAIR) [4] by others.

3. **Accessibility for non-experts:** AI courses use many research software components, some of which may not be easy to adopt by wider audiences. In order to welcome students from other areas besides Computer Science, AI courses should include introductory materials for using command line interfaces, teach how to read software documentation, and teach how to install virtual environments and package managers such as Pypi [5] and Conda [6]. These skills are crucial when testing the results from colleagues and exploring novel algorithms.

**References** [1] Pedregosa, F., Varoquaux, G., Gramfort, A. et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. vol 12, pp: 2825-2830. 2011. <https://doi.org/doi/10.5555/1953048.2078195>

[2] Abadi, M., Agarwal, A., Barham, P., et al. TensorFlow, Large-scale machine learning on heterogeneous systems. Zenodo. 2015. <https://doi.org/10.5281/zenodo.4724125>

[3] Paszke, A., Gross, S., Massa, F., et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems* 32, pp: 8024-8035. 2019. <https://doi.org/doi/10.5555/3454287.3455008>

[4] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

[5] <https://pypi.org/>

[6] <https://docs.conda.io/>

## AI has the potential for broad impact to numerous disciplines. How could we make AI education more interdisciplinary?

**Zhuoyue Lyu (University of Toronto)**

Our EAAI paper [1] explored the idea of teaching AI in an interdisciplinary setting integrating STEM with the humanities. We developed a high school

AI lesson focusing on teaching Variational Autoencoder (VAE) through philosophical metaphors, creative art, and music applications. The pilot studies with 22 students found that our approach was effective.

Students in our study were fascinated by the creativity of AI. To them, AI was not just something that predicted their next YouTube video or the Face ID system on their phones anymore. It's much bigger. It's lively, vibrant, and creative. They even understood the connection between AI and philosophy. Plato's cave is an allegory that is 2400 years old, and it's somewhat abstract and obscure. Nevertheless, students were able to utilize that to learn the concept of Latent Space in the VAE model without problems.

As a (non-professional) singer and podcaster myself, I enjoy finding the connection between STEM and humanities, which is also the vision of many researchers who built amazing AIs such as MusicVAE [2], SketchRNN [3], Pointflow [4], that received great feedback from the public. So let us keep exploring AI's connection with other disciplines and integrating more non-engineering examples into our curriculum.

Why is this important? Because AI should be human-centered, and our next generations of scientists, engineers, and policymakers need to understand and agree with that idea. The humanities are a powerful way to achieve that as it's the utmost expression of human creativity and imagination that can resonate with people regardless of their nations and identities.

**References** [1] Lyu, Zhuoyue, Safinah Ali, and Cynthia Breazeal. "Introducing Variational Autoencoders to High School Students." *Proceedings of the AAAI Conference on Artificial Intelligence*. 2022. <https://raise.mit.edu/vae>

[2] Roberts, Adam, et al. "A hierarchical latent vector model for learning long-term structure in music." *International conference on machine learning* PMLR, 2018.

[3] Ha, David, and Douglas Eck. "A neural representation of sketch drawings." *arXiv preprint arXiv:1704.03477*. 2017.

[4] Yang, Guandao, et al. "Pointflow: 3d point cloud generation with continuous normalizing flows." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.

## How to make AI more interdisciplinary

**Christopher J. MacLellan (Drexel University)**

I argue that by developed specialized educational AI resources for Human-Computer Interaction (HCI) students and practitioners, we can make AI more interdisciplinary and empower this audience to create the human-centered AI systems of the future. As a researcher working at the intersection of AI and HCI, I have discovered that many HCI students and practitioners have a limited view of AI. They often view AI as a collection of isolated algorithms for specific purposes (e.g., classification or recommendation) and lack an understanding of how different AI components integrate to produce intelligent behavior. Further, their understanding also often omits key topics and theoretical ideas, such as expert systems, analogical reasoning, and creativity. HCI practitioners are eager to leverage AI advancements, but these fundamental deficiencies limit the kinds of human-centered technologies they might one day create.

To address these deficiencies, I have been developing a graduate-level course on Human-AI Interaction for Drexel University's HCI and User Experience (UX) master's program. The course aims to introduce students to AI concepts as well as the unique design challenges faced when integrating AI into user-facing technologies. Over the term, students read and discuss papers that intentionally include a broad range of perspective on AI topics (e.g., readings were selected to represent different sociocultural and AI perspectives). The central goal for the readings is to help students gain a broader understanding of AI approaches and how these approaches can be integrated to support human-centered design. In conjunction with the readings, students work on interdisciplinary teams to design a user-facing AI prototype.

Even though many of my students came to the course with limited technical experience, they were all able to successfully produce and present AI prototypes by the end of the term. Further, many of them developed compelling concepts at the intersection of AI and HCI. For example, one group of students explored the development of a novel interactive voice assistant that leveraged both a rule-based dialog system and speech-to-text functionality to support elderly people in tracking and managing chronic health conditions. Another group explored novel ways for users to provide temporally scoped preference information to recommender systems (e.g., I love this song, but I do not want to hear this song for the next few weeks) and reimaged what a system like Spotify would look like with these new capabilities.

While I originally designed the course to help HCI students advance their AI knowledge, I found that these students bring a fresh, interdisciplinary perspective to the AI discipline itself. HCI students almost always start by considering the human context for AI technologies and as a result they would often identify ways that widely use AI design patterns (e.g., recommender systems that only accept binary yes/no feedback) are insufficient. I argue that by creating more courses that bridge the HCI and AI disciplines, we can both increase the impact of AI technologies and enrich our discipline.

### Acknowledgments

The EAAI-22 AAI/ACM SIGAI New and Future AI Educator Program is partly supported by funding from ACM SIGAI and the Artificial Intelligence Journal.



**Michael Guerzhoy** is a Assistant Professor, Teachig Stream at the University of Toronto, and an Affiliate Scientist at the Li Ka Shing Knowledge Institute, St. Michael's Hospital. His professional interests are in computer science and data science education and in applications of machine learning

to healthcare.



**Marion Neumann** is a Teaching Professor and the Director of Data Science Programs at the Department of Computer Science and Engineering at Washington University in St. Louis. She teaches Machine Learning, Analysis of Networked Data, and Introduction to Data Science.

Her research interests include graph-based machine learning and capturing and analyzing student feedback and emotions in large computing courses using text mining and sentiment analysis.



**Emmanuel Johnson** is a postdoctoral research associate at the University of Southern California's Information Sciences Institute through a CI Fellowship. His research explores building AI systems for teaching interpersonal skills.

