AI in the Humanities & Social Sciences

Ethics, Practice, and Pedagogies

University of Tennessee, Knoxville
April 25-26, 2025

Session Information and Abstracts

MARCO INSTITUTE

AI in the Humanities and Social Sciences
Ethics, Practice, and Pedagogies

Session #1: Ethical Issues and Social Implications of Creative AI
Discussant: Julia Abramson (University of Oklahoma)
Friday, April 25 • 2:00-3:30pm • Haslam 103

Kevin Taylor (University of Memphis): "Does AI Have Buddha-Nature? Hakuin Zen and Generative AI"

This paper argues that there are parallels that can be drawn between the teachings of Zen master Hakuin Ekaku (1685-1768) and pedagogical/moral questions surrounding the use of generative AI such as ChatGPT.

Generative AI through what are termed "hallucinations" certainly fail to satisfy Buddhist right speech insofar as ChatGPT is a people pleasing technology that will give false or misleading information. As a moral component of the noble eightfold path, generative AI is therefore possibly unethical and certainly pedagogically dangerous if accepted as infallible.

Hakuin, like other Zen masters, urges us not to confuse the finger pointing to the moon for the moon itself. This metaphor for language and the experience of enlightenment is analogous to students accepting ChatGPT responses as true without critical inquiry. Furthermore, Hakuin taught his students a combination of Great Doubt and post-*satori* practice to overcome the misguided confidence brought about by small insights, mistaking small enlightenments for great enlightenment. This type of fallibilism further can be seen in Hakuin's use of *nembutsu* practice. Hakuin was opposed to "other-power" *tariki* practices as opposed to Zen "self-power" *jiriki* practices and yet Hakuin occasionally employed *nembutsu* practices as skillful means when teaching students, urging them that his own *koan*, the sound of one hand, was vastly superior.

Hakuin Zen thus can be said to teach students to doubt the insights offered by generative AI as "confusing the moon with the finger" and to continually investigate beyond surface level answers as infinitely superior, while simultaneously seeing the possibility of generative AI as a tool when used in the right hands to help students go beyond small achievements

Alex Bentley (University of Tennessee, Knoxville): "What Is AI Model Collapse and How Do We Avoid It?"

In the film, "Heretic," Hugh Grant describes centuries of monotheistic religions as "Iterations. Over time. Diluting the message. Obscuring the original." In my own research with colleagues in Barcelona, we call this the "Dilution of Expertise." It could accelerate dramatically with AI, into a world of gibberish called "AI Model Collapse," which occurs when a socially-learning AI agents, like ChatGPT, train on their own content. Over successive 'generations', the agents 'forget' aspect of the original training data, eventually spewing nonsense. This is similar to evolutionary drift, which shows us what strategy to take. One lesson is to keep human experts in the loop. Event if we don't know exactly where the experts are, their expertise can diffuse through a population of social learners, like chatBots. This "pied piper" effect can greatly improve AI training data and the performance of generative AI as a result.

Veljko Dubljević (NC State University): "Responsible Use of Generative AI Technologies"

Generative AI has the potential to transform scholarly publishing. Authors, peer reviewers, and editors might use AI in a variety of ways, and those uses might augment their existing work or might instead be intended to replace it. Editors of bioethics and humanities journals, who have been contemplating the implications of this ongoing transformation, recently articulated an opinion that generative AI may pose a threat to the goals that animate scholarly work but could also be valuable for achieving those goals. In the interests of fostering a wider conversation about how generative AI may be used, a preliminary set of recommendations for its use in scholarly publishing was developed, in the hope that the recommendations and rationales will help the scholarly community navigate toward a deeper understanding of the strengths, limits, andchallenges of AI for responsible scholarly work.

AI in the Humanities and Social Sciences
Ethics, Practice, and Pedagogies

Session #2: Languages, Literatures, and Technology
Discussant: Hilary Havens (University of Tennessee, Knoxville)
Saturday, April 26 • 9:00-10:30am • Haslam 103

Timothy Beal (Case Western Reserve University): "Face of the Deep: Perverse Applications of AI and Machine Learning"

How might we in the humanities use AI and other emerging technologies perversely, to facilitate difficulty, frustrate felicitousness, and resist automation?

As an example, I will share work on "Face of the Deep," a project that seeks new models and methods for hosting a user experience with the processes of translating ancient texts that attends to the rich ambiguities and polyvocalities of the text in translation. How might emerging technologies enable a kind of "brooding over the face of the deep" (Genesis 1) that opens up creative, meaning-making possibilities without the need to decide or close anything off?

If the aim of machine translation is to erase ambiguity and make the processes of translation invisible and immediate, we aim to build models that use these technologies perversely, to slow down and make visible the complex processes of translation, in order to invite users to participate in those processes.

Harrison Meadows (University of Tennessee, Knoxville): "Digitizing the Early Modern Spanish-Language Archive: Challenges and Opportunities"

I am currently collaborating with University of Alabama colleagues (Drs. Sergei Gleyzer and Xabier Granja) and a global team of software engineers (Arsh Khan, Utsav Rai, Shashank Sheklhar Singh, and Yukinori Yamamoto) on a machine learning project called renAIssance. Under the auspices of the HumanAI foundation, the project was funded by the 2024 Google Summer of Code, and just accepted again for GSoC 2025. Our objective is to develop an end-to-end tool that can digitize 16th- and 17th-century Spanish print texts using machine learning techniques and optical character recognition (OCR). The tools that currently exist are either not sufficiently accurate to handle standard research tasks or require specialized training, so the tool stands to offer an efficient means to expand the digital archive and vastly increase the amount of searchable data for scholars of the early modern period. As a result, the broader vision for the tool is to participate in realizing a broadened, diversified, and more accessible digital archive for historical research in the Humanities. In this presentation, I will offer a Humanities perspective on the project so far, both in terms of the development status of the project to date, and also provide applicable takeaways for DH scholars based on my experience collaborating with software engineers on a Google-sponsored project.

Lauren Goodlad (Rutgers University): "We Have Always Been Action Theorists: What LLMs Mean for the 'L' in MLA"

This paper describes how large language models (LLMs), the underlying technology for generative AI tools such as ChatGPT, work — including their impacts, social implications, and the surrounding political economy. It concludes by arguing that scholars in literary studies should reconsider the poststructuralist tenets that have dominated since the 1980s and explore the field's less visible commitments to an action theory of languge that helps elucidate the language effects of "AI."

AI in the Humanities and Social Sciences
Ethics, Practice, and Pedagogies

Session #3: AI in Premodern Research and Pedagogies
Discussant: Angela Gibson (Modern Language Association)
Saturday, April 26 • 11:00am-12:30pm • Haslam 103

Carmella Matzza (Louisiana State University): "AI-Powered Pedagogy: Transforming the Study of Women in Medieval and Early Modern Spain"

This paper will discuss how AI tools can be used to enhance the study of women in medieval and early modern Spain, a field that often presents challenges in capturing the complexity and diversity of women's experiences. The presentation will highlight specific AI-powered classroom activities that help students engage with historical texts, images, and interactive maps, as well as strategies for assessing students' work through AI-enhanced rubrics. Dr. Mattza's focus will be on how these tools can help students develop skills such as empathy, critical thinking, and creativity, enriching their understanding of the past.

Alexa Alice Joubin (George Washington University): "'Such stuff as dreams are made on': Shakespeare and RenAIssance Technologies"

In this session, Dr. Joubin will explore how Shakespearean drama and AI technologies intersect, specifically examining how early modern theatrical culture and Shakespeare's soliloquies anticipate AI as a performative agent. Drawing connections between AI hallucinations and Shakespeare's fantastical depictions of dreams, this talk proposes new pedagogical frameworks that teach AI through the lens of performance studies. Dr. Joubin will also consider the ethical implications of AI as a tool for creativity, encouraging users to adopt a reflective, critical stance toward AI's outputs. By integrating AI with performance theory, she suggests that we foster a deeper understanding of AI's cultural and cognitive impacts in both the humanities and the classroom.

Paul Atkins (University of Washington): "Statistics, Machine Learning, and Classical Japanese Orthography"

Dr. Atkins will present his groundbreaking work in the field of classical Japanese orthography, demonstrating how machine learning and statistical analysis are being used to analyze hiragana—a system of writing that was once highly fluid and polygraphic. His presentation will focus on the application of AI to manuscript studies, showing how AI tools can identify scribes, attribute manuscripts, and analyze historical texts that have previously been difficult to study at scale. The session will include a discussion of the machine learning models used to extract and categorize orthographic data from handwritten documents, allowing for large-scale analysis and more accurate dating and attribution of medieval Japanese manuscripts.