

Artificial Humanities: AI, Language, and Fiction

Nina Begus

Human and Social Dynamics of Data and Technology Requirement

UC Berkeley, CDSS

Spring 2024

The course is expanded from a [study abroad](#) course, titled *AI and Humanities: Language, Literature, and Ethics*, designed in 2020 for the University of Washington.

Course number: Data 94

Unit value: [3 credits](#); 4 credits with an additional section

Instructional format and contact hours per week: two 90min lectures

Duration of course: Semester (15 weeks)

Prerequisites: none

Estimated total hrs of student work per week: 3 hours in-classroom, 4 hours reading literature and writing assignments

Restrictions: none

With the tumultuous arrival of AI systems and wider availability of generative AI, we have undoubtedly entered the AI age. The artificial humanities course offers a framework to explore the setbacks and opportunities of this era through the imaginative and conceptual power of both humanities scholarship and its objects of study. Artificial humanities explore and demonstrate how humanities—as broad as science and technology studies, history of technology, film and media studies, literary studies, philosophy, rhetoric, women and gender studies—can help address most pressing issues in AI development and use.

How has AI challenged our common conceptions of creativity, agency, intelligence, and language? We will address this question by looking at the history and contemporary use of chatbots, virtual assistants, virtual beings, and generative AI, occasionally paired with social robotics and communicative neurotechnology. We will address social implications of digital culture, addressing questions of regulation and copyright, privacy and bias, access and control, and the changing nature of work.

We will focus on fictional representations of AI as entangled with historical, conceptual, and technical development of AI technologies. Through considering AI narratives as epistemic sites of inquiry, students will develop a critical eye towards AI discourse and social ramifications. They will be able to employ their analytic, creative, or technical skills during the weekly assignments, in their presentations, and in the final project.

Goals:

- a) to learn about AI's broader cultural, philosophical, artistic, and social implications;
- b) to evaluate the influence of imaginary to actual representations of AI;

- c) to recognize philosophical, ethical, critical, and historical traditions embedded into AI technologies;
- d) to critically assess current and future AI systems through group discussions, presentations, and debate in a collegial and civic setting;
- e) to formulate a creative, analytical, or consulting project with a clear lesson on how an AI product, product category, or historical trajectory would benefit from humanistic knowledge.

Course Outline

Week 1: Introduction

Week 2: The Pygmalion Myth

- *Shaw's Pygmalion*
- *The Eliza Effect*

Week 3: Language Benchmarks

- *The Turing Test*
- *The History of the Concept of Language*

Week 4: Personification and Gender

- *Assistants*
- *Companions*

Week 5: Human Values and Social Bias

- *Value alignment*

Week 6: Writing Machines

- *Fictional Writing Machines*
- *Factual Writing Machines*

Week 7: Creativity

- *Creative Software*
- *Copyright Infringement + Debate*

Week 8: Language Development

- *Language Difference*
- *Voice and Voicelessness*

Week 9: Therapy

- *Myself*
- *Others*

Week 10: Medicine

- *Neurotechnology*
- *Caretaking*

Spring Break

Week 11: Artificial Humanities Projects

- *Student Proposals Workshop*
- *Proposal Presentations*

Week 12: Search: Social Critique

- *Cybernetics, Cyberpunk, Cyberspace*
- *Surveillance Capitalism*

Week 13: Risks and Benefits

- *Digital Labor*
- *Distributed AI systems*

Week 14: **Future: Speculations and Solutions**

- *To the Planetary Future*
- *Final Debate: AI: Good or Bad?*

Week 15: **RRR Week**

Requirements and grading

- **40%:** 10 weekly responses (each 4%). There are altogether 11 weekly assignments, and one with the lowest grade will be dropped in the final grade.
- **10%:** Class attendance.
- **15%:** Choose between 1) in-class presentation of readings, 2) midterm debate on AI creativity or 3) final debate on AI: good or bad?, or 4) final project proposal presentation.
- **5%:** Attend an AI event (artistic, industry, non-profit, academic) and briefly report back to class.
- **30%:** Final project (see more below under Assignments).

Assignments

Weekly Written Responses: submit on bCourses every week before class on Monday.

10 responses are mandatory out of 11 possible responses.

Assignments should be 300 words. We will distribute clear guidelines on how to prepare adequate written responses to our weekly questions.

Briefly, make sure to develop a clear argument and support it with textual evidence from the readings. It is important you incorporate all the daily readings into your response. You can choose between Monday or Wednesday readings.

In-class presentations and debates: SIGN UP FOR ONE OPTION AMONG

a) **Presentations of readings:** 10-minute verbal presentation of the day's readings during lecture. Feel free to use visuals and slides.

b) **Final project proposal presentations:** for students who wish to introduce their final project proposals to the class and get constructive feedback from their peers.

c) **Mini debate** on AI creativity. Prepare a stance for or against during the weekly response and present your argument eloquently to the class.

d) **Final debate** on AI repercussions: The final debate will take the whole lecture and will be conducted in a parliamentary form: the debaters' goal is to convince the voting body in the audience to vote for their argumentation.

AI event attendance: Being at UC Berkeley and close to Stanford University, UCSF, and Silicon Valley, we are right at the heart of AI technological efforts. Attend an event of your choice that focuses on AI (not necessarily entirely, but at least in some significant aspect) and briefly report

to class your observations: what you learned or your critique, concerns, enthusiasm, and extrapolation.

Final Project: Please discuss the project proposal with us. We will help you to develop your final project during week 11 and are always accessible on office hours and via email. It is important to incorporate topics and readings from the class into the final project. The project can be an academic paper, a technological artifact, or a creative piece. Choose one track for your project:

- **Creative/artistic:** Create an artistic or creative piece on your own and accompany it with an artist's statement.
- **Critical/analytical:** Write a critical analysis of AI practice or philosophy.
- **Computational/technical:** Present a technological artifact and accompany it with written critical reflection.

Syllabus

Week 1: What is Artificial Humanities?

The first class will introduce the course content and requirements. It will demonstrate with a few examples how students will use the knowledge of artificial humanities to analyze fictional works thematizing AI, historical testimonies of its technological development, and actual products.

No readings for this week.

Week 2: The Pygmalion Myth

3 Shaw's *Pygmalion*

The framework for our approach to computers and language will be the Pygmalion myth, a powerful trope that brings personification of the nonhuman (nowadays, machines) to the extreme. We will look at fictional examples of the myth throughout the course, focusing on language acquisition, difference, and training in humans and machines.

Most classes will consist of a fictional, analytical, and technological example. We start with **fiction** readings:

- Pygmalion's story from Ovid's *Metamorphoses* (8 AD) ([verse](#) and [prose](#) translation)

Choose one reading or viewing:

- EITHER George Bernard Shaw, play *Pygmalion* (1911)
- OR film [Pygmalion](#) (1938)
- OR film [My Fair Lady](#) (1964)

First assignment due. All are due every Monday before class.

4 *The Eliza Effect*

Personification of machines has been a known tendency since the 1960s. We will examine how this tendency was addressed in chatbots then and now.

- **Analysis:** Excerpt from Joseph Weizenbaum, *Computer power and human reason: From judgment to calculation*, 1976.
- **Fiction:** Excerpt from Richard Powers, novel [Galatea 2.2](#), 1995
- **Fiction:** Excerpt from Blaise Agüera y Arcas, novella [Ubi sunt](#), 2022
- **Analysis:** Lemoine Blake, [The New York Times](#), 2022

Week 3: Language Benchmarks

5 *The Turing Test*

This week we will look at technologists, particularly Alan Turing from the early days of AI, who prominently addressed conversational language in computers. We will parallel their views with the ontological substratum they are operating within, provided by Descartes's philosophy that marks the modern era.

- **Tech:** Alan Turing, '[Computing Machinery and Intelligence](#),' *Mind* 1950, P. 433–460.
- **Analysis:** René Descartes, [Discourse on the Method](#), Part V, and [Second Meditation](#).
- **Fiction:** Blade Runner, [the Voight-Kampff test scene](#), 1982.

Second assignment due.

6 *The History of the Concept of Language*

- **Analysis:** Tobias Rees, '[Non-Human Words: On GPT-3 as a Philosophical Laboratory](#),' *Daedalus* (2022) 151 (2): 168–182.
- **Analysis:** Emily Bender, Timnit Gebru, et al. '[On the Dangers of Stochastic Parrots](#),' *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (March 2021). Pages 610–623.
- **Analysis:** Noam Chomsky, '[The False Promise of ChatGPT](#),' *New York Times Opinion*, March 2023.
- **Fiction:** Ted Chiang, very short story [The Great Silence](#), 2015

Week 4: Personification and Gender

7 Assistants

What kinds of philosophical, cultural, and imaginary limitations has digital media of the 2010s and the 2020s inherited? We will first look at the idea of the technological aid as a servant—personified service.

- **Analysis:** Simone Natale, Excerpt from [Deceitful Media](#), 2021
- **Analysis:** Markus Krajewski, Excerpt from [The Server](#), 2018

Third assignment due.

8 Companions

From the very beginning, virtual assistants and similar AI products were gendered as women. We will examine the impact of fictional representations to technologists and the public.

- **Fiction:** Film [Her](#), 2013
- **Tech:** Scarlett Johansson's [Open Letter to OpenAI](#), 2024
- **Analysis:** DeepMind Regulation Report on [Advanced AI Assistants](#), 2024

DROP/ADD DEADLINE: this Wednesday

Week 5: Human Values and Social Bias

9 Value alignment

Value alignment is one of the central problems of AI. We will discuss how to instill values in machines justly and fairly, focusing on social bias and aligning common goals between humans and machines.

- **Analysis:** Brian Christian, *The Alignment Problem: Machine Learning and Human Values*, 2020.
- **Analysis:** Gebru, Timnit. "Race and Gender." *The Oxford Handbook of Ethics of AI*, 2020.

Fourth assignment due (on Wednesday).

Week 6: Writing Machines

10 Fictional Writing Machines

Writers have been famously afraid of their jobs being taken by machines. We will read two different scenarios of literature-generating machines, which both open questions of authorship, literary market and criticism and challenge our social and institutional order.

- **Fiction:** Roald Dahl, short story [The Great Automatic Grammatizator](#), 1954
- **Fiction:** Andrej Tomažin, short story from [Anonymous Technology](#), 2018

Fifth assignment due (on Monday).

11 Factual Writing Machines

Writing has now ceased to be a solely human activity. Who is the author in the age of AI? Are writing machines tools or creators on their own? Is co-creativity the best approach to writing with language models? How is AI changing the concepts of knowledge and the self?

- **Fiction:** Vauhini Vara, essay [Ghosts](#), 2021
- **Analysis:** Lila Shroff, essay [Datasets as Imagination](#), 2023

Week 7: Creativity

12 Creative Software

Human creativity has a certain signature. Can we call machines creative and what does machine creativity look like? Is machine creativity essentially human creativity processed through AI or is there a novel aspect of processing data, otherwise inaccessible to humans?

- **Analysis:** Agnieszka Kurant, [Collective intelligence](#), 2021
- **Analysis:** Blair Attard-Frost, [Queering Intelligence](#), 2024
- **Tech:** Guest lecture on generative adversarial networks (GANs)

Sixth assignment due.

13 Copyright Infringement + Debate

Digitization of (largely written, language-based, and partly also visual) human knowledge has been essential to knowledge distribution and organization in the digital age. Digital culture is a part of our cultural capital that often eludes societal laws of protection, intellectual property, and distribution. Copyright law, most often attributed to AI objections, is a complex landscape that

has predictably failed to catch up with AI advances, causing an uproar among creators. We will discuss best practices, solutions, and regulatory approaches.

- **Tech:** Ermira Murati, [Language and Coding Creativity](#), 2022.
- **Analysis:** Karla Ortiz, [Statement before US senate on AI and copyright](#), 2023.

A few students can choose to debate for or against AI creativity.

Week 8: Language Development

14 *Language Difference*

Human language in machines faces all kinds of technical as well as cultural hurdles. We will ask today how and how much language difference—translation, speech impediments, accent change—should we instill in machines.

- **Analysis:** Marc Shell, [Language Wars](#), 2001
- **Analysis:** Rebecca L Johnson et al., [The Ghost in the Machine has an American accent: value conflict in GPT-3](#), 2022
- **Tech:** Rime Blog, [How to Change Your Accent](#) (2024)
- **Fiction:** [Babel Fish translator](#) from *The Hitchhiker's Guide to the Galaxy*, 1978

Seventh assignment due.

15 *Voice and voicelessness*

How does machine language find a voice? What possibilities lie in mirroring an individual style and in generalizing human writing and speech? Besides that, is there something in machine language that eludes human abilities, a possibility to uproot this nonhuman production of language?

- **Analysis:** Michelle Huang, [Training an AI Chatbot on my Childhood Journal Entries](#), 2022.
- **Tech:** Claude Shannon, [A Mathematical Theory of Information](#). 1948. P. 1-7.
- **Fiction:** Stanisław Lem, SciFi philosophical essay [Golem XIV](#), 1981

Week 9: Therapy

16 *Myself*

Chatbots and virtual beings are being used for emotional support and mental health support. We will examine this unregulated use of AI and discuss its risks and benefits.

- **Analysis:** Zara Abrams, [AI is changing every aspect of psychology](#). The American Psychological Association, 2023
- **Fiction:** Sheila Heti's conversation with ELIZA in [The Paris Review](#), 2022
- **Analysis:** Excerpt from Hannah Zeavin, [The Distance Cure](#), 2021

Eighth assignment due.

17 Others

Technology is often introduced as a technical invention that can fix social problems. We will look at further aspirations introduced at the intersection of AI in medicine and evaluate them from the ethical perspective.

- **Fiction:** Black Mirror S2 episode 1 [Be Right Back](#), 2013
- **Analysis:** Gual-Montolio P et al. [Using Artificial Intelligence to Enhance Ongoing Psychological Interventions for Emotional Problems in Real- or Close to Real-Time: A Systematic Review](#), 2022
- **Analysis:** Minerva F, Giubilini A. [Is AI the Future of Mental Healthcare?](#) 2023

Week 10: Medicine

18 Neurotechnology

Neurotechnology is linked to AI products, such as LLMs and virtual beings, particularly through neurocommunicative devices. We will read neuroethical papers that deal with already existing treatments in parallel to a science fiction short story and a memoir by a completely paralyzed patient.

- **Tech:** Seth Metzger et al, [A high performance neuroprosthesis for speech decoding and avatar control](#), 2023
- **EITHER Fiction:** Chapter from Jean-Dominique Bauby's memoir, *The Diving Bell and the Butterfly* (1997); alternatively, watch the [eponymous film](#) (2007)
- **AND Analysis:** Goering, Sara. 'Is It Still Me? DBS, Agency, and the Extended, Relational Me.' *AJOB Neuroscience* (2014): 50-51.
- **OR Fiction:** James Tiptree Jr. [Alice Sheldon], short story [The Girl Who Was Plugged In](#), 1973
- **AND Analysis:** Goering, Sara et al. 'Staying in the Loop: Relational Agency and Identity in Next-Generation DBS for Psychiatry.' *AJOB Neuroscience* (2017): 58-70.

Ninth assignment due.

19 Caretaking

As the world's population is aging, we will face a caretaking crisis. Is outsourcing work that requires human touch unethical? How can we improve human wellbeing through technology without dehumanizing them?

- EITHER **Fiction:** Martin Shoemaker, short story [Today I am Paul](#) (2015)
- OR **Fiction:** Isaac Asimov, [Robbie](#), 1940
- **Tech:** [Cynthia Breazeal](#)'s robots for pediatric hospitals, nursing home robots ([Paro](#))

Spring Break

Week 11: Humanities+AI Projects

20 Student Proposals Workshop

No readings. This week, focus on your proposal for the final paper and come to class prepared with questions for feedback.

You can propose your final project with a

- critical/analytic approach,
- creative/artistic approach,
- computational/technical approach.

A few students can choose to introduce their proposals to the whole class, followed by a general discussion.

21 Proposal Presentations

A few students can choose to introduce their proposals to the whole class, followed by a general discussion.

Final project proposal due.

Week 12: Search: Social Critique

22 Cybernetics, Cyberpunk, Cyberspace

How do we preserve objective and authoritative truth in the world of deepfakes? AI has amplified already known issues from the digital sphere that have begun with the invention of the Internet. We will look back at the development of the networked digital sphere and trace the philosophy that has developed alongside technological advancements.

- **Tech:** Konrad Zuse, Excerpt from *Calculating Space*, 1969.
- **Fiction:** William Gibson, *Neuromancer*, 1984.
- **Analysis:** Donna Haraway, Excerpt from [A Cyborg Manifesto](#), 1984.
- **OR Analysis:** Michael Polanyi. Excerpt from *Personal Knowledge*, 1958.

Tenth assignment due.

23 Surveillance Capitalism

The issues of power, access, and control have a long history around information technologies. As technologists developed search engines, based on the field of information retrieval (further based on library science), they also exacerbated core problems of the digital sphere: privacy, information accuracy and reliability, surveillance, labor exploitation, big tech monopolies.

- **Fiction:** [Star Trek computer](#), 2013
- **Tech:** Eric Schmidt, 'A Chance for Growth.' *Frankfurter Allgemeine*, 2014
- **Criticism:** Shoshanna Zuboff. Excerpt from [The Age of Surveillance Capitalism](#), 2018

Week 13: Risks and Benefits

24 Digital Labor

We will look at two sides of labor: companies and their employees. What technologists do and the vision they hold for work as opposed to the effects that their company is creating? What labor do they hide and what work environment do they promote? What is the future of work?

- **EITHER Analysis:** Irani, Lilly. 'Difference and Dependence among Digital Workers: The Case of Amazon Mechanical Turk.' 2015. P. 225-34.
OR Analysis: Bernard Geoghegan, 'Orientalism and Informatics: Alterity from the Chess-Playing Turk to Amazon's Mechanical Turk.' 2020. P. 45-90.
- **EITHER Analysis:** Kevin Scott, [Reprogramming the American Dream: From Rural America to Silicon Valley—Makin AI Serve Us All](#), 2020
OR Kevin Scott, [Do Not Think It Means What You Think It Means: Artificial Intelligence, Cognitive Work & Scale](#), 2022
- **Tech:** [Co-pilot](#)

Eleventh assignment due.

25 Distributed AI systems

With machine learning, computer engineering has changed in its essence, which pertains to the idea of human control. The lack of control over machines has been a persistent trope in AI narratives. How do we deal with existential risks of AI and why are these dystopian scenarios so prevalent both in our imagination and technology forecasting?

- **Fiction:** E. M. Forster, [The Machine Stops](#), 1909.
- **Analysis:** Nicole Starosielski, Excerpt from [The Undersea Network](#), 2015.

Week 14: Future: Speculations and Solutions

26 To the Planetary Future

In the final lecture of this course, we will discuss how to conceptualize AI as a part of the planetary system. The goal of the planetary philosophy is to embed AI systems more ethically for the humans and for the planet. At the same time, the planetary perspective fosters a distinct new philosophy, where technology is just one natural-technical subsystems of the planet. Can this philosophy help us build and employ technology better?

- **Fiction:** Kim Stanley Robinson. *Excerpt from Ministry of the Future*, 2020
- **Analysis:** Benjamin Bratton, *Planetary Sentience* (2021)
- **Analysis:** N. Katherine Hayles, Excerpt from *Bacteria to AI: Human Futures with Our Nonhuman Symbionts* (2025)

27 Final Debate: AI and Humanities

Six students can choose to debate the prompt as the final reflection of the course.

Week 15: RRR Week

Week 16: Exam Week

Final project due.