

Project Proposal:

Speculative blueprint for tending to your computer's memory.

In a world where we aim to have AI act closer and closer to humans, what happens when we give it the human trait of forgetting? In this project, I'm proposing a machine learning model that forgets its training over time. By doing so, I raise questions about the relationship between human and machine and how I can drive a wedge into technological determinism.

Questions:

- How can we become comfortable with a technology (that we expect to remember everything) "forgetting"?
- How do you tend and care for your machine that forgets? Do you speak to it often? How do you choose what to help it remember or what to let go?
- How does the relationship with AI change on a larger scale if everyone expects their machines to forget over time?

Key Topics:

- Technological Determinism
- Data Feminism
- Training machine learning models

Key Readings:

- "Algorithms of Oppression" - Safiya Noble
 - The nature of how Google is presented to the public is that of technological determinism by the public. It's the go to destination online for searching for anything on the web and the results are often "normalized as believable and often presented as factual" (Noble 24-25). OpenAI's development of ChatGPT currently falls into a similar sentiment of an all knowing being, guaranteed to have its "artificial general intelligence (AGI) aligned with human values and follow human intent," (Leike, Schulman, Wu). Though we're at a very early stage of what feels like an impending shift with our relationship to AI, the implication that AI is aligned with human values and intents feels similar to Google's existence as a credible source of information. So what happens if we can change the relationship with AI into something that is fallible, erroneous, an untrustworthy narrator, or more human?
- "Data Feminism" - Catherine D'Ignazio and Lauren F. Klein
 - I hope through this project I'll be able to reconstruct the direction that AI is moving towards. This project will be pushed forward through the lens of the seven principles of data feminism in order to question what memory means in the context of technology and how the idea of the "all knowing" machine can be questioned given new technological constraints.
- "¿Human ÷ (Automation + Culture) = Partner?" - Stephanie Dinkins

- “The increasing use of automated technologies will change us. The question is how much and to what end,” (Dinkins 296). When looking at Stephanie Dinkins work, I’d like to approach my project in a similar way from her “Conversations with Bina48” that led to “Not the Only One”, one a highly sophisticated model that couldn’t answer the question “who are your people” and another a multigenerational memoir created through small data and the voices of three generations of a black American family.

Additional Readings:

- “The Enduring Ephemeral, or the Future Is a Memory” - Wendy Chun
- [Replika’s software update](#) that had AI erase aspects of their relationship with users

Sources:

- Dinkins, Stephanie. “¿Human ÷ (Automation + Culture) = Partner?” *ASAP/Journal*, vol. 4, no. 2, 2019, pp294-297
- Leik, Jan, et al. Our Approach to Alignment Research, openai.com/blog/our-approach-to-alignment-research.
- Noble, Safiya Umoja. “A Society, Searching.” *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York University Press, New York, 2018.