🔑 Generative Agents: Interactive Simulacra of Human Behavior
Keynote | Michael Bernstein

Abstract: Believable proxies of human attitudes and behavior can empower applications ranging from immersive environments to social policy interventions. However, the last quarter century has seen a slow recession of human behavioral simulation as a method, in part because traditional simulations have been unable to capture the complexity and contingency of human behavior. I argue that modern artificial intelligence models allow us to re-examine this limitation. I make my case through generative agents: computational software agents that simulate believable human behavior. Generative agents enable us to populate an interactive sandbox environment inspired by The Sims, where end users can interact with a small town of twenty five agents using natural language. Our generative agent architecture empowers agents to remember, reflect, and plan — enabling them to wake up, cook breakfast, and head to work; artists paint, while authors write; they form opinions, notice each other, and initiate conversations; they remember and reflect on days past as they plan the next day. Extending this line of argument, I explore how proxying human behavior and attitudes can help us design more effective online social spaces, understand the societal disagreement underlying modern AI models, and better embed societal values into our algorithms.

Bio: Michael Bernstein is an Associate Professor of Computer Science at Stanford University, where he is a Bass University Fellow. His research in human-computer interaction focuses on the design of social computing systems. This research has won best paper awards at top conferences in human-computer interaction, including CHI, CSCW, ICWSM, and UIST, and has been reported in venues such as The New York Times, Science, Wired, and The Guardian. Michael has been recognized with an Alfred P. Sloan Fellowship, UIST Lasting Impact Award, and the Patrick J. McGovern Tech for Humanity Prize. He holds a bachelor's degree in Symbolic Systems from Stanford University, as well as a master's degree and a Ph.D. in Computer Science from MIT.