

Human-Centered AI

Ben Shneiderman's Blueprint for an AI Age That Serves People

July 2026



Human-Centered AI (Oxford University Press, 2022). Cover: Oxford University Press.

Some books argue that artificial intelligence should serve humanity. Very few explain how to build the institutions that would make it do so. Ben Shneiderman's *Human-Centered AI*, published by Oxford University Press in 2022, is one of the rare works that does both — and it is for that reason a foundational text for everything AIWS is building.

Professor Shneiderman is among the founding figures of human-computer interaction as a discipline. He founded the Human-Computer Interaction Laboratory at the University of Maryland and directed it from 1983 to 2000; his innovations in information visualization, including treemaps and dynamic query interfaces, are now part of the everyday vocabulary of software design. He is a member of the U.S. National Academy of Engineering and a Fellow of the AAAS, ACM, IEEE, and the National Academy of Inventors. *Human-Centered AI* received the American Publishers Award for Professional and Scholarly Excellence in Computing and Information Sciences.

RISING ABOVE THE LEVELS OF AUTOMATION

The book's central intellectual move is deceptively simple, and it dissolves an assumption that has constrained thinking about AI for half a century.

For decades, engineers described automation along a single line: at one end, the human does everything; at the other, the machine does everything. Every design decision became a trade-off. More machine capability meant less human control. Less human control was simply the price of progress.

Shneiderman rejects the line and replaces it with a plane. Human control and computer automation are not opposing ends of one dimension but two independent axes. A system can be highly automated *and* leave people firmly in command — a modern camera, an elevator, a well-designed medical device. The failures we fear are not the consequence of automation itself; they occur when high automation is paired with low human control, and they are avoidable by design rather than inevitable by physics.

This reframing has a consequence that matters far beyond engineering. If human control and machine capability are not in conflict, then the choice between a powerful AI and a humane one is a false choice. Those who claim that safety must be traded against capability are not describing a law of nature. They are describing a design decision — one that can be made differently.

SUPERTOOLS, NOT REPLACEMENTS

The book's middle section examines the metaphors that shape AI research, and here Shneiderman is at his most useful for policymakers. The dominant metaphors — the intelligent agent, the artificial teammate, the machine that emulates the human — quietly smuggle in the assumption that the goal is to build a substitute for a person.

He proposes a different family: *supertools* that amplify human capability the way a microscope amplifies sight; *tele-bots* that extend human reach into places people cannot go; *control centers* that give operators comprehensive awareness of complex systems. The distinction is not a matter of taste. Choose the emulation metaphor and you will design toward autonomy and measure success by how little the human is needed. Choose the amplification metaphor and you will design toward mastery and measure success by what people become capable of.

THE BRIDGE FROM ETHICS TO PRACTICE

For AIWS, the most consequential portion of the book is its fourth part, on governance. Shneiderman observes what has become painfully obvious in the years since publication: the world has produced a great abundance of AI ethics principles and very little machinery for making them bind.

His answer is a layered structure, each layer operating on a different actor and a different timescale: sound software engineering practice inside development teams; a safety culture established through management strategy inside organizations; trustworthy certification carried out by independent oversight bodies outside the organization; and government intervention where the preceding three prove insufficient.

Readers of the Boston Declaration and the Tokyo Compact will recognize this architecture immediately, because it is the same one. AIWS Trust Standards define what verifiable trustworthiness requires at the engineering layer. The AIWS Automated Trust Monitor performs continuous independent observation. Trust Rating and certification supply the external verification that no organization can perform on itself. The AIWS Trust Order convenes the governments, institutions, and scientists whose agreement makes the whole structure hold.

Professor Shneiderman gave this argument its scholarly foundation. What AIWS is attempting is its construction.

AN HONEST BOOK

One further quality deserves mention. Each part of *Human-Centered AI* closes with what the author calls a skeptic's corner, in which he sets out the strongest objections to his own position and answers them without softening. It is a form of intellectual honesty that has become rare in writing about technology, where confidence is too often mistaken for insight.

That honesty is why the book has aged well while so much AI commentary from the same period has not. Shneiderman describes himself as an optimist, but he is an optimist of the demanding kind — one who believes a better outcome is achievable precisely because he has thought carefully about everything that could prevent it.

WITH GRATITUDE

Professor Shneiderman recently sent a copy of *Human-Centered AI* to Nguyen Anh Tuan, Co-Founder, Co-Chair, and CEO of the Boston Global Forum and Founder of AIWS. The Boston Global Forum records its sincere thanks — for the book itself, and for the generosity of spirit that accompanied it.

A gift of this kind carries more than its contents. It is an act of intellectual companionship, an offer to think together about questions that no institution and no discipline can answer alone. We receive it in that spirit.

The Boston Global Forum and AIWS are honored to count Professor Shneiderman among the thinkers whose work informs our own. *Human-Centered AI* is not merely a book we recommend. It is part of the foundation on which we are building.

Human-Centered AI by Ben Shneiderman is published by Oxford University Press (2022).