118th NOTE on Human-Centered AI

Feb 21, 2024, 12:03:21 AM

From: Ben Shneiderman
To: Human-Centered-AI on Google Groups

Dear HCAI Google Group,

A new book from Dennis Yi Tenen of Columbia University, Literary Theory for Robots, offers a delightfully fresh perspective on AI. He looks back over more than a thousand years into literary history to report on an astonishing variety of algorithmic strategies for writing poems, plays, novels, fables, and much more. He values these literary and linguistic templates, schema, scripts, methods, processes, and patterns, as the roots of computer science – "the hidden history of modern machine intelligence." By combining his literary knowledge with a background as a Microsoft software engineer, Tenen brings readers a human-centered way of appreciating technology.

Tenen sees dictionaries, thesauri, grammar books, and writing guides as tools. These tools are now implemented as spell checkers, auto-completion, and text generation applications, but that does not bring them to life. Rather, he celebrates the creative power of human effort and collaboration: "great art, like great cuisine, will always remain the purview of exceptional talent."

The book is short, so you should enjoy the history lessons, but the final chapter provides valuable guidance for how to think about AI. Among his nine ideas is that "AI is collective labor" and that "Intelligence is distributed." These two ideas remind us that AI is a deeply human endeavor that emerges from a long history and broad collaborations.

Tenen also worries that the metaphors around AI are misleading and that "metaphors obscure responsibility". He continues to focus on responsibility, writing that "machines alone cannot become moral agents." Then he closes by reflecting on human history: "what emerges . . . is not some mystical quality of intelligence, but cooperation, magical and deeply satisfying in its own right." His shifting from vague beliefs in intelligence (especially the vague notions of AI or AGI) to the power of human cooperation rings true to me.

I'm pleased to be starting a new project with Mona Sloane (Univ of Virginia) to co-chair four workshops for the U.S. National Academies on the topic of Human and Organizational Factors in AI Risk Management: A Workshop. Our 12-person committee will select themes and speakers for four one-day workshops in response to this task statement: "The National Academies of Sciences, Engineering, and Medicine will organize a workshop to identify and explore approaches to addressing human and organizational risks in AI systems. The emphasis will be on approaches that could be included in a more detailed guidance document complementing the recently issued 'AI Risk Management Framework' from the National Institute of Standards and Technology (NIST). Areas of interest in

the context of the NIST framework include human insights about AI-produced output and human oversight of AI systems and their operation in real-world environments." Committee member bios are available online.

Aneesh Raman and Maria Flynn write in the New York Times: When Your Technical Skills Are Eclipsed, Your Humanity Will Matter More Than Ever (February 14, 2024). They open with the usual fears that AI threatens many jobs, but then suggest that "A.I. could usher in a world of work that is anchored more, not less, around human ability." They ask "What are our core capabilities as humans?" but their answer is vaguely around communication, people skills, social abilities, and the "relationship economy." While I am warmly sympathetic to their argument, I wish they made stronger cases about (1) the remarkable abilities of people and (2) the potential for design changes in AI systems to give greater power to users. Like many critics they assume the technology is fixed, when it is easily redesigned to provide greater user control by way of improved user interfaces, which enable users to specify what they want.

(1) The remarkable abilities of people have been repeatedly discussed in these NOTES. The eight Exceptional Skills of Experts were mentioned in the 117th NOTE and discussed in detail in the 95th and 96th NOTES. These include "social skills to effectively work with peers, superiors, and staff" but include memory, perceptual, cognitive, motor, tool, and metacognitive skills. Finally, the skill of accepting responsibility for their actions, morally and legally, make for exemplary human performance.

(2) The potential for design changes in AI systems to give greater power to users is another repeated theme with the goal to amplify, augment, empower, and enhance human performance. Supporting user self-efficacy, creativity, responsibility, social connectedness, and human collaboration are clear paths to commercial success.

Stephen Marche writes in the New York Times: Everybody Is Talking About A.I. What the Heck Is It, Anyway? A guide to the best books about artificial intelligence. He makes questionable claims such as: "Before the launch of ChatGPT, a little more than a year ago, it was difficult to get readers to care about A.I.... After the launch of ChatGPT, everybody had an opinion, and nobody knew what they were talking about... after a decade during which Silicon Valley has demonstrated that it lacks any sense of social responsibility, it has become impossible to trust the creators of A.I." Sweeping statements are provocative, but more nuanced positions might be more helpful in suggesting what is needed from developers, business leaders, policy makers, and users. Marche's list of the 5 best books on AI are: (1) The Alignment Problem: Machine Learning and Human Values, by Brian Christian (2020), (2) Artificial Intelligence: A Guide for Thinking Humans, by Melanie Mitchell (2019), (3) The Algorithm: How AI Decides Who Gets Hired, Monitored, Promoted &Fired & Why We Need to Fight Back Now, by Hilke Schellmann (2024), (4) Progressive Capitalism: How to Make Tech Work for All of Us, by Ro Khanna (2022), and (5) AI 2041: Ten Visions for Our Future, by Kai-Fu Lee and Chen Qiufan (2021).

Readers of this Google Group suggested their best choices of book in the 38th NOTE to this group on December 15, 2021. The list contained 14 popular books on Human-Centered AI. It is interesting to see that similar issues are still under discussion. If readers send me their selections for the best recent popular books on AI, I will compile a new list.

Best wishes. . . Ben

Other items:

I was pleased to see the U.S. National Science Foundation Program Solicitation on Responsible Design, Development, and Deployment of Technologies (ReDDDoT). They emphasize multi-disciplinary and multi-sector (collaboration between academic, industrial, medical, and government) teams to "examine and demonstrate the principles, methodologies, implementations, and impacts associated with responsible design, development, and deployment of technologies in practice." The program managers are strongly oriented to human-centered and community-centered approaches that promote participation. They wish to fund design, development, and deployment of technologies that support "Accelerating pathways to societal and economic benefits while developing strategies to avoid or mitigate societal and economic harms; and Empowering communities, including economically disadvantaged and marginalized populations, to participate in all stages of technology development, including the earliest stages of ideation and design." The focus is on responsible technology co-design and knowledge co-production in AI, biotechnology, and disaster prevention or mitigation. The U.S. NSF continues to expand its emphasis on human-centered approaches that are multi-disciplinary, multi-sectoral, and oriented to implementation of technologies, in addition to traditional academic laboratory research.

The CHI 2024 Conference (May 11-16, 2024, in Honolulu, Hawaii) will include a hybrid Workshop on Theory of Mind in Human-AI Interaction.The six person organizing committee write that "Theory of Mind (ToM) refers to humans' capability of attributing mental states such as goals, emotions, and beliefs to ourselves and others." They seek papers on "Conceptual discussions of ToM in human-AI interaction, including how to define and measure ToM in both human and AI during interactions; Design approaches for imparting effective mental models of AI systems with ToM-like capabilities; Empirical evaluations of humans' mental state attribution behaviors (e.g., anthropomorphism) to AI" and related topics. Paper Submissions are due February 22, 2024 AoE

ABOUT: This Google Group on Human-Centered AI, run by Ben Shneiderman sends occasional notes devoted to ensuring human control while increasing the level of automation. The goal is to design AI-infused supertools that amplify, augment, empower, and enhance human performance. The Group, now with 3500+ members, is meant to carry forward the ideas in Ben's book on Human-Centered AI. Torrey Mortenson manages a Slack group to enable further discussions. Mengnan Du maintains our list of resources, which includes research centers, courses, events, etc. Chenhao Tan responds to requests for new members.

Thanks to them and others for helping to build our community.