Two Workshops, a Report, and a 100-Year Study of Artificial Intelligence and Society

Lara Streiff

■ *The purpose of this article is to report on the findings of two workshops exploring the evolution of artificial intelligence technologies — specifically used in care-driven or predictive applications — and their impacts on society as a whole, organized as part of the AI100's 100-year-long study of artificial intelligence. Workshop participants concluded that care cannot be commodified or reduced into outcome-oriented tasks and can therefore not be encoded into technology; additionally, participants determined that regulation of predictive artificial intelligence technologies is required to maintain their benefit and trustworthiness. The aim of these workshops is to encapsulate both the ups and downs of these technologies. The study organizers and participants feel their role is to bring the integration of artificial intelligence technology into societal values to the forefront of discussions for its future.*

A 100-year-long study of artificial intelligence (AI) — known as the AI100 — is now working toward its second report to reflect on, and predict, the societal impacts of AI technologies.

When the project was launched in 2014, an interdisciplinary group of experts gathered to assess the effects AI has on its users and their communities, as well as the technology itself. The first report, titled "Artificial Intelligence and Life in 2030," is a reference for those in government and industry, as well as for the general public, on how to interact with AI (Stone et al. 2016). It covers eight sectors spanning from topics such as transportation and healthcare, to entertainment.

As we enter the next decade, a second report looms on the horizon. This follow-up report presents an opportunity to reflect on the booming changes to the industry and resultant impacts on society since the first study findings were released. While maintaining a level of continuity, this next report is expected to aim a broader lens on the influences of these technologies worldwide. It will also explore human-centric applications in greater depth, to touch upon the personal connections between individuals and AI technologies.

The human element is increasingly important as interactions with AI expand through applications like autonomous vehicles, increasingly capable search engines, and electronic personal assistants. Debating ethics, purpose, intention, and deployment of these technologies will remain an ongoing challenge for this study. To reflect these realities, the committee is expected to include scholars from disciplines such as philosophy, anthropology, sociology, and critical studies in addition to AI scientists and engineers.

A Tale of Two Workshops

The accelerating presence of AI use today means that the study also finds itself adapting. Deviating from the planned report timeline, the leadership of the AI100 commissioned two workshops to capture the rapidly evolving status of AI technologies in mid-2019.

Today, autonomous delivery robots and self-driving cars can roam city streets on a daily basis. The number of personal smart devices including watches and cell phones — handheld and AI-equipped components of a daily uniform seen on all ages and professions — now far outnumber people (Grosz and Stone 2018).

To pinpoint the most pressing topics that they should address, the AI100 leadership sent a call for proposals to the AI community and reviewed more than 100 submissions before settling on their final topics. They selected two for funding: taking on the roles and limitations of care technologies; as well as the decision-making systems in prediction technologies. The topics ranged from the court system to at-home healthcare technologies.

The aim of these workshops, reports, and even this longitudinal study, is to encapsulate both the ups and downs of these technologies — creating a long-term, realistic view of AI. The organization of the 100-Year Study makes it uniquely primed to capture the cyclical nature of public views and attitudes toward AI.

The outcomes of the two workshops identified the pitfalls of outsourcing problems for technology to solve rather than addressing the causes, and the possible consequences of outdated predictive modeling going unchecked. The AI100 recently released online summaries of these topics, which will inform the study panel and the upcoming second report. Key messages are outlined in sections below.

First Things First: The First Report

Even 100 years may not be enough time to cover all aspects of AI technologies. Keeping up with the rapid development poses the greatest challenge to this long-term study. To get things started in the first report, a selected study panel focused on eight prominent sectors wherein AI has a significant presence. These domains were transportation, home and service robots, healthcare, education, public safety and security, low-resource communities, employment and workplace, and entertainment.

The takeaways from the first report feed into both the workshops and the overarching goals of the study. For one, AI prediction tools were deemed to have great potential to reduce human bias, improve health outcomes, and address the needs of low-income communities in the first report. However, as the Prediction in Practice workshop concluded, there is an inherent risk to deploying these technologies without serious thought as to how the outcomes are interpreted.

The inaugural study panel also determined that society is underinvesting in the research of societal implications stemming from AI technologies. In a sense, the AI100 reaffirmed their own purpose, and the subsequent workshops reinforce this message. Society is welcoming AI technologies at ever-increasing rates; although these smart technologies have the potential to make our lives easier, they introduce a number of ethical, design, and policy challenges that will compound over time if not addressed. Discussion and deliberation from a diverse set of voices can help ensure that benefits from AI can be trusted and shared equitably — which the AI100 hopes to achieve.

Workshop 1: Coding Caring

The first workshop, entitled Coding Caring (Arnold et al. 2019), identified a specific trend of caring technologies. Although AI is already becoming widespread in healthcare applications, participants in this workshop concluded that care itself is uniquely human — not something that can be encoded into technology.

AI technologies in daily life, like smartwatches or smart speakers, can remind people to take medication or track health information. But they are limited in their ability to display empathy or provide emotional support. Care cannot be commodified or reduced into outcome-oriented tasks for smart personal assistant devices to take on.

As a result, participants agreed that new care technologies should be integrated into existing human-to-human care relationships, supplementing the interactions between a caregiver and a care-receiver, but never replacing them outright.

Workshop 2: Prediction in Practice

While the Caring Coding Workshop focused primarily on design and fixing issues from inside the industry, the second workshop, entitled Prediction in Practice (Barocas et al. 2019), emphasized both internal and external awareness. Its participants reinstated outside regulation as the key factor ensuring AI technologies remain beneficial and trustworthy.

As the name would suggest, this workshop centered on predictive technologies; specifically, case studies of AI-informed high-stakes decision-making in the public sector. This meant looking at AI applications that lead to impactful decisions in public institutions, including but not limited to modeling predictions and allocating resources. In the first AI100

Important Announcement!

**The 16th AAAI Conference on
Artificial Intelligence and Interactive Digital Entertainment**

October 19–23, 2020

Due to COVID-19, the AIIDE organizing committee has decided that AIIDE'20 will be held online. The conference dates remain the same, as shown above. As usual, accepted papers will be published by AAAI and at least one of the authors must register and attend the conference to present their work.

See aiide.org for more information.

report, prediction technologies were highlighted as an area where AI could provide great benefit by removing human bias, but this workshop's participants also identified some of the shortcomings of these algorithms.

The four case studies examined in the second workshop were: pretrial risk assessment in criminal justice cases; screening algorithms in child protection services; an algorithm that proposed new school start times; and prediction of high risk infections in healthcare. The level of impacts and risks associated with these scenarios varies, but each carries significant weight for the decision-makers. Outdated use in conjunction with these scenarios would affect decisions regarding imprisonment and childcare as well as public health.

Because predictive AI trains on data from the past — and both AI and society are rapidly changing — the technology will always need updates and reevaluations. The workshop suggests expiration dates for AI to ensure that base standards are maintained over time, especially in these high-impact applications.

Participants and organizers of the study and its workshops feel the AI100's role is to bring the integration of AI technology into societal values to the forefront of discussions for its future. The next report is slated to be released within a year.

Acknowledgments

References

Arnold, T.; Currie, M.; Elder, A.; Feldman, J.; Himmelreich, J.; and Niker, F. 2019. *Coding Caring Workshop Report*. Stanford, CA: Stanford University. ai100.sites.stanford.edu/sites/g/files/sbiybj9861/f/coding_caring_workshop_report_1000w_0.pdf.

Barocas, S.; Bogen, M.; Kleinberg, J.; Levy, K.; Nissenbaum, H.; and Robinson, D. 2019. *Prediction in Practice Workshop Report*. Ithaca, NY: Cornell University. ai100.sites.stanford.edu/sites/g/files/sbiybj9861/f/cornell_summary_report_public.pdf.

Grosz, B. J., and Stone, P. 2018. A Century-Long Commitment to Assessing Artificial Intelligence and Its Impact on Society. *Communications of the ACM* 61(12): 68–73. doi.org/10.1145/3198470.

Stone, P.; Brooks, R.; Brynjolfsson, E.; Calo, R.; Etzioni, O.; Hager, G.; Hirschberg, J.; Kalyanakrishnan, S.; Kamar, E.; Kraus, S., et al. 2016. *Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence (A100): Report of the 2015–2016 Study Panel*. Stanford, CA: Stanford University. ai100.stanford.edu/2016-report.

Lara Streiff is a freelance science writer and current graduate student in the Science Communication Program at the University of California, Santa Cruz. Her research background is in environmental science and innovations in climate solutions at the University of California, San Diego, and the Scripps Institution of Oceanography, and she has recently written for Stanford University, the Monterey Bay Aquarium Research Institute, and the Bay Area News Group.