Communicating about the Social Implications of AI

Research Methods and Sample Composition

This supplement provides detailed information on the research that informs FrameWorks' strategic brief on reframing the social implications of AI. Below, we outline the research conducted with researchers, advocates, practitioners, and members of the public that provides the evidence base for the brief, describing the methods used and sample composition.

The Field Story of the Social Implications of AI

To develop an effective strategy for communicating about an issue, it's necessary to identify a set of key ideas to convey. For this project, these ideas were garnered from researchers and advocates in the field of the social implications of AI. Due to our interest in how these implications may vary by domain, we spoke not only to generalists in the field but researchers and advocates in three domains: policing, health care, and child welfare. FrameWorks researchers conducted 11 one-hour interviews with researchers and advocates and reviewed relevant literature. The interviews focused on predictive algorithms, and four experts were interviewed as generalists, while the remaining interviews were conducted with two experts on predictive algorithms as they apply to policing, three to child welfare, and two to health care. Interviews were conducted between August and September 2020 and were recorded and transcribed for analysis, with participants' permission. FrameWorks compiled the list of interviewees in collaboration with the MacArthur Foundation. To refine the field story, FrameWorks conducted a 90-minute feedback session with researchers and advocates in October 2020.

FRAME
WORKS

Interviews with members of the field consisted of a series of probing questions designed to capture their understandings about key aspects of the field. All interviews probed what artificial intelligence and predictive algorithms are and what the broader regulatory and policy context is around the use of predictive algorithms and other forms of AI. Experts in each of the three domains (policing, child welfare, and health care) were asked how predictive algorithms are used in their domain of expertise, including the uses' benefits and challenges, data sources, and what should be done to regulate them. In each interview, the conducting researcher used a series of prompts and hypothetical scenarios for members of the field to explain their research, experience, and perspectives; dissect complicated relationships; and simplify complex concepts. In addition to preset questions, FrameWorks researchers asked for elaboration and clarification and encouraged researchers and advocates to expand on concepts they identified as particularly important.

Analysis employed a basic grounded theory approach[1]. A FrameWorks researcher identified and inductively categorized common themes that emerged in each interview and across the sample. This resulted in a refined set of themes, which researchers supplemented with a review of materials from relevant literature.

Public Understandings of the Social Implications of AI

A primary goal of this research was to capture the various commonly held assumptions, or cultural models, that members of the public use to make sense of technology, computers, and artificial intelligence, as well as their social implications. Cultural models are cognitive shortcuts to understanding—ways of interpreting, organizing, and making meaning of the world around us—shaped through years of experience and expectations and by culturally embedded beliefs and values.[2] These ways of thinking are available to all members of a culture, although different models may be activated at different times. Individuals belong to multiple cultures, each of which include multiple models (e.g., people participate in public cultures at multiple levels, including national and subgroup cultures). In this project, our goal was to explore the models available in American public culture, but it is important to acknowledge that individuals also have access to models from other cultures in which they participate.

In exploring cultural models, we seek to identify *how* people think, rather than *what* they think. Cultural model findings thus differ from public opinion research, which documents people's surface-level responses to questions. By understanding the deep, often tacit assumptions that structure how people think about technology, computers, artificial intelligence, and their social implications, we can understand the obstacles that prevent people from accessing the field's perspective described in the field story. We can also identify opportunities that communicators can take advantage of—existing ways of thinking that can help people arrive at a fuller understanding of the issue.

To identify the cultural models that the public uses to contemplate issues related to technology, computers, artificial intelligence, and their social implications, FrameWorks researchers conducted a set of interviews with members of the public. FrameWorks researchers conducted 25 interviews over Zoom with people across the U.S. in October 2020. We recruited a diverse sample of participants, with variation along key dimensions such as race, ethnicity, and socioeconomic status (see below).

Cultural model interviews are one-on-one, semi-structured interviews lasting approximately two hours. They are designed to allow researchers to capture broad sets of assumptions, or cultural models, that participants use to make sense of a concept or topic area—in this case, issues related to technology, computers, artificial intelligence, and their social implications. Interviews consisted of a series of open-ended questions covering participants' thinking on those topics in broad terms, before focusing more specifically on their thoughts regarding the consequences of using AI in two of the three identified domains (policing, child welfare, and health care). Interviews examined what "artificial intelligence" means and how it is used broadly in our society, as well as how it is used specifically in the different domains and the associated social consequences. Researchers approached each interview with this set of topics but allowed participants to determine the direction and nature of the discussion. All interviews were recorded and transcribed with participants' written consent.

All participants were recruited by a professional marketing firm and selected to represent variation along several dimensions: age, gender, race and ethnicity, educational background, income, employment status, political views (as self-reported during the screening process), and family situation (e.g., married or single, with or without children). The sample of members of the public included 12 women and 13 men. Of the 25 participants, 12 identified as white, five as Black, four as Hispanic or Latinx, two as Asian, and two as mixed race. One participant had a high school degree or less; 10 had completed some college; 10 had graduated from college; and four had postgraduate degrees. Nine participants were between 20-39 years old, twelve participants were between 40-59 years old, and four were age 60+. Twelve participants described their political views as "Democrat"; seven as "Republican"; and six as "Independent/ Other".

To analyze the interviews, researchers used analytical techniques from cognitive and linguistic anthropology to examine how participants understood issues related to technology, computers, artificial intelligence, and their social implications.[3] First, researchers identified common ways of talking across the sample to reveal assumptions, relationships, logical steps, and connections that were commonly made but taken for granted throughout an individual's dialogue. The analysis involved discerning patterns in both what participants said (i.e., how they related, explained, and understood things) and what they did not say (i.e., assumptions and implied relationships). In many cases, participants revealed conflicting models on the same issue. In such cases, one conflicting way of understanding was typically found to be dominant over the other, in that it more consistently and deeply shaped participants' thinking

(i.e., participants drew on this model with greater frequency and relied more heavily on it in arriving at conclusions). To ensure consistency, researchers met after an initial round of coding and analysis, comparing and processing initial findings, then revisited transcripts to explore differences and questions that arose through the comparison. As part of this process, researchers compared emerging findings to the findings from previous cultural models research, as a check to ensure that they had not missed or misunderstood any important models. They then reconvened and arrived at a synthesized set of findings.

Analysis centered on ways of understanding that were shared across participants, as cultural models research is designed to identify common ways of thinking that can be identified across a sample. While there is fixed rule or percentage used to identify what counts as "shared," models reported are typically found in the large majority of interviews. Models found in a smaller percentage of interviews are only reported if there is a clear reason why they only appeared in a limited set of interviews (e.g., the model reflected the thinking of a particular subgroup of people).

While a sample of 25 participants is too small to ensure the sample is *statistically* representative, its demographic variability is adequate to ensure the identified patterns in thinking are shared across different groups *within* the U.S. While larger sample sizes are needed to investigate variability within a population, or to allow for statistically significant comparisons between groups, the goal of cultural models analysis is to describe common ways of understanding within a population. As a result, sample size for cultural models research is determined by the concept of *saturation*, wherein a sample is considered to be of a suitable size when new data do not reveal further underlying patterns of thinking within a population. For this project, our analyses confirmed that a sample size of 25 interviews was sufficient to reach a point of saturation for cultural models of technology, computers, artificial intelligence, and their social implications in the U.S.

Endnotes

1. Glaser, B. & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research* (observations). Aldine; Strauss, A. & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* SAGE publications.

2. Quinn, N. (Ed.). (2005). *Finding culture in talk: A collection of methods.* Palgrave Macmillan.

3. Quinn, N. (Ed.). (2005). *Finding culture in talk: A collection of methods.* Palgrave Macmillan.

About FrameWorks

The FrameWorks Institute is a nonprofit think tank that advances the mission-driven sector's capacity to frame the public discourse about social and scientific issues. The organization's signature approach, Strategic Frame Analysis®, offers empirical guidance on what to say, how to say it, and what to leave unsaid. FrameWorks designs, conducts, and publishes multi-method, multidisciplinary framing research to prepare experts and advocates to expand their constituencies, to build public will, and to further public understanding. To make sure this research drives social change, FrameWorks supports partners in reframing, through strategic consultation, campaign design, FrameChecks®, toolkits, online courses, and in-depth learning engagements known as FrameLabs. In 2015, FrameWorks was named one of nine organizations worldwide to receive the MacArthur Award for Creative and Effective Institutions.

Learn more at www.frameworksinstitute.org

Communicating about the Social Implications of AI: Research Methods and Sample Composition

August 2021

FRAME WORKS