Ph.D. Thesis

Divy Thakkar

Towards examining Human-AI collaboration across the AI pipeline

Advisors: Alex Taylor, Stephann Makri, Anoop Sinha

This thesis has been submitted (02/2024) to the Department of Computer Science, School of Science & Technology, City, University of London. I, Divy Thakkar, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The integration of Artificial Intelligence (AI) into decision-making workflows provides significant opportunities for productivity gains but also raises complex questions about its effectiveness and impact on society. Drawing upon rich qualitative studies from extensive fieldwork conducted in the Global South over five years, this thesis engages with the overarching question – *how do humans and AI collaborate across the AI pipeline, from design to deployment?*. This thesis provides an analysis of the Human-AI collaboration pipeline by examining it across stages and multiple stakeholders. I begin by addressing – *Who are the creators behind computing?*, focusing specifically on women's representation in India's computing industry. Next, I engage with growing challenges around data quality by exploring – *How does data powering AI come to be?*. This exploration is grounded in the study of datafication in India's public health sector, highlighting the complexities and challenges in ensuring high-quality data for AI systems. Further, I examine: *How does AI change the nature of work?* by studying a first-of-its-kind large-scale AI deployment in India. Lastly, I uncover: *How do people perceive and experience AI?* by examining AI perceptions and experiences of vocational technicians in India, a historically underserved community vulnerable to job loss through automation. Through a reflective analysis, I develop an understanding of Human-AI collaboration by examining human and non-human actors across the AI development pipeline. This perspective recognizes the interconnectedness of human and non-human elements, including social, cultural, and organizational factors, in shaping the development and impact of AI technologies. This thesis emphasises the study of AI's role in the Global South, particularly in high-stakes domains such as public health and future of work. It highlights the potential impacts of AI on historically underserved communities, underscoring the need for inclusive and context-sensitive approaches in AI development and deployment. By examining Human-AI collaboration across the entire pipeline and situating it within diverse contexts, this thesis contributes to a wider examination of AI's role in society and a path for future research in this area. In conclusion, this thesis contributes to the field of Human-Computer Interaction by examining Human-AI collaboration as a part of a broader pipeline of AI design to deployment.

Dedicated to

Mumma, Daddy and Deep

Table of Contents

Abstract i

Acknowledgements ix

1 Introduction **1**

2 Who are the creators behind computing? **13**

2.1 Introduction 13

2.2 Method 16

2.3 Related Work 17

2.3.1 Factors affecting CS careers 17

2.3.2 Employment and education in the Indian context 18

2.4 Findings 19

2.4.1 Undergrad: bright futures *through* CS 19

2.4.2 Master's degrees: recovering agency 21

2.4.3 Ph.D.: marriage pressures and isolation 21

2.4.4 Careers and practice: sharp drops due to family life 23

2.5 Discussion 24

3 How does data powering AI come to be? 28

3.1 Background and Motivation 28

3.2 Method 31

3.3 Related Work - Data Quality in ML 32

3.4 Findings 34

3.4.1 ML developers: Making Data Fit for Purpose 34

3.4.2 Data Stewards: Aggregating Data to Meet Reporting Requirements . 36

3.4.3 Data Collectors: Collecting Data as Part of Everyday Workflows . . . 38

3.5 Discussion 40

3.5.1 Valuing Across the Data Supply Chain 42

3.5.2 Transparency and Accountability in Data Transformation 43

4 How does AI change the nature of work? 45

4.1 Motivation 45

4.2 Background 48

4.2.1 Overview of the mCare Program 48
4.2.2 Integration of AI into the mCare Program 49
4.3 Related Work - Human-AI Collaboration 51
4.4 Methods 53
4.5 Findings 54
4.5.1 Who to Target: Identifying Beneficiaries 54
4.5.2 AI in the Background: Interactions between Callers and Beneficiaries 56
4.5.3 When to Call: Aligning Across Constraints 58
4.5.4 Why AI: Considering Program Goals 59
4.6 Discussion 60
4.6.1 AI as Actor 61
4.6.2 Configuring Human-AI Partnerships 62

5 How do people perceive and experience AI? 64

5.1 Background and Motivation 64
5.2 Related Work 67
5.2.1 Public Perceptions of Automation and AI 67
5.2.2 Algorithmic Fairness among Marginalized Groups 68
5.3 Method 69

5.4 Findings . . . 70
5.4.1 The Technological Promise of Vocational Training . . . 70
5.4.2 Local Visions of the Future of Work . . . 72
5.4.3 Algorithmic Fairness in Education, Skilling, and Jobs . . . 74
5.5 Discussion . . . 78

6 Summary and Discussion **81**
6.1 Key Contributions . . . 81
6.2 Analysis . . . 85
6.2.1 Interplay of Users and Creators . . . 86
6.2.2 Network of Human and Nonhuman Actors . . . 88
6.2.3 Configuration and Reconfiguration . . . 91
6.2.4 Diverse Context – Global South . . . 92
6.3 Limitations . . . 94
6.4 Future Work . . . 96
Bibliography . . . 102

List of Tables

3.1 Valuation of ‘good’ data across stakeholders 40

6.1 Summary of Academic Contributions. 101

Acknowledgements

I have, to my great bewilderment, been provided excellent opportunities, taken a chance on, mentored and nurtured by several people – I am perennially grateful.

I would like to begin by extending sincere appreciation to my advisors. Firstly, Alex Taylor, you are a force of good in this world. Your scholarship, kindness and rigour are values I aspire for. In many ways, this Ph.D. would not have been possible without your support and had you not taken a chance on me. I am grateful to my co-advisor, Anoop Sinha, for being an inspiration to pursue long-term high-impact research and for opening many doors for me. I am thankful to the wonderful academic and research staff at City, University of London, including Ann-Marie and Stephann Makri for supporting me throughout this journey. I am grateful to my thesis committee – Claus Bossen, Duncan Brumby, and Mai Elshehaly.

There are two mentors have been instrumental in shaping my career. Ashwani Sharma, thank you for constantly empowering me with opportunities, for being a brainstorming buddy and one of the kindest people. I would not be doing this Ph.D. without your support. Nithya Sambasivan, thank you who for being a life-long mentor, for spending hours teaching me the craft and science of HCI. I would not have pursued HCI if not for you. My Ph.D. began as soon as I started working with you, and I will be forever grateful for your mentorship, friendship, and kindness.

This ride has been fun because of several collaborators who have inspired me with their critical scholarship. Thank you so much, Neha Kumar, Azra Ismail, Kentaro Toyoma, Hari Subramonyam, Milind Tambe, Aparna Taneja, Balaraman Ravindran. I am grateful to Richard Anderson and Aditya Vashistha who first showed me what a path to doing a Ph.D.

could look like, during my visit to UW. I have reached so far due to the strong foundations of some excellent teachers - Alka Parikh, Madhumita Mazumdar, Ruppal Walia, Abha Dua, and Ranjana Ranjan. Deep gratitude goes to Maggie and Leslie for not just funding my Ph.D. but also nurturing my career at Google.

I am grateful to have incredible friends who have constantly cheered for me and shaped my thinking in many ways. A few lines cannot do justice to what your support, love, encouragement and presence has meant to me. Thank you so much – Anant Ranka, Artika Dixit, Ayush Sahay, Chirag Kavdia, Chinmay Chaudhary, Eram Fatima, Kavya Krishna, Kunal Jain, Koosh Doctor, Meghana Ratheesh, Nirav Devpura, Nikit Saraf, Nikhilesh Somani, Parth Nigam, Ravi Bhoot, Rishav Thakker, Sarthak Chand, Sanil Jain, Urvashi Gupta, Tanuj Maheshwari, Vishal Anand, and Vaibhavi Desai. A special thank you to Prateek Jain for being my family in Bangalore and a constant source of strength through this journey. I don't know what I'd do without you all.

I am deeply grateful to the hundreds of research participants who opened their hearts to me, thank you for trusting me to tell your stories and co-design technologies with you - I do not take this responsibility lightly.

It feels strange to thank an entire organisation - but I am truly grateful to Google Research, a place where I grew up. The culture of innovation, ambition and the deep sense of purpose this organisation embodies has uplifted me in myriad ways. I would not be here without my Google colleagues who have pushed me to think bigger every day. Thank you – Aishwarya PS, Adriana Budura, Anirudh GP, Abhirut Gupta, Ding Wang, Dheeraj Nagaraj, Dinesh Tewari, Gaurav Aggrawal, Kartikeya Badola, Laura Sasportas, Marie Efstathiou, Nitish Gupta, Pradeep Shenoy, Partha Talukdar, Manish Jain, Melanie Saldana, Michelle Huratado, Manish Gupta, Jay Yagnik, Jeff Dean, Rahul Singh, Rahul Sukthankar, Rainier Aliment, Shachi Dave, and Varun Gulshan.

I am thankful to my extended family (n=80 :)) for their constant support. I want to thank my late grandfather, Nanalal, who ingrained his love for academia. I am sure he's smiling from above knowing I got my doctorate. There are countless people who have inspired me, I have lost touch with some of them but I remain equally grateful to each of you.

Lastly, I am grateful to the best parents in the world – Jagruti and Hasmukh, who provided me with a childhood filled with love, books, science shows and a quest for independent thinking. Thank you for your countless sacrifices to provide me with a strong value system and for constantly believing in me. I owe everything to you both. Deep, thank you for being the best big brother one could ask for – I strive to embody your curiosity and empathy in my work and life. I would be lost without you. I am grateful for the support of my sister-in-law, Hiral, who joined our family a year ago and lit our lives with her kindness.

Chapter 1

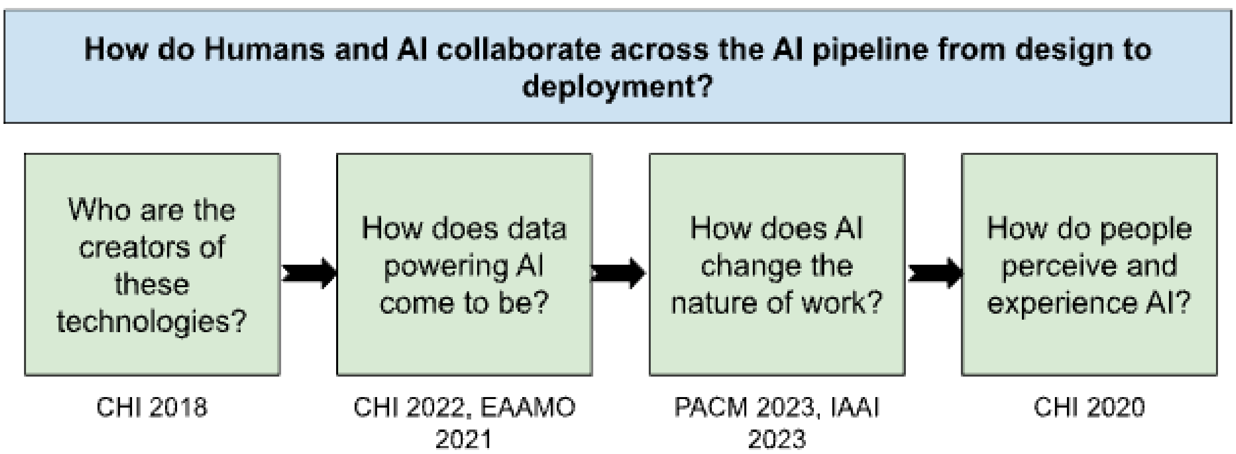
Introduction

The rapid proliferation of Artificial Intelligence (AI) into various facets of everyday life, particularly through AI-based decision-making systems, has transformed the landscape of human interaction and organizational workflows. There has been significant interest from industry, academia, governments and the non-profit sector to deploy AI in existing human workflows (Chui et al., 2018; *IBM Science for Social Good*, 2018; *MSR India Center for Societal impact through Cloud and Artificial Intelligence (SCAI)*, 2019; Murali & PK, 2019; *The Rockefeller Foundation Establishes Atlas AI – New Startup to Generate Actionable Intelligence on Global Development Challenges*, n.d.). AI is being touted as the next wave of productivity with applications ranging from helping developers write faster and better code, helping doctors with automated clinical notes, decision making systems to assess patient risk, helping financial analysts examine risks to assisting customer care agents with managing queries efficiently. McKinsey, a consulting firm advising Fortune 500 companies, estimates "current generative AI and other technologies have the potential to automate work activities that absorb 60 to 70 percent of employees' time today." (McKinsey, n.d.).

However, despite the rapid progress in the field, the adoption of these technologies has been cautiously met with concerns around data quality, job displacement, force-fitting, and safety. This has invited a plethora of questions concerning integration of AI into the complex tapestry of cultural, social, economic, and organizational ecologies. A notable example highlighting the complexities and challenges in AI deployment is the case of IBM Watson for Healthcare. The project was discontinued after incurring a significant financial loss of around USD 5 billion. This outcome was largely attributed to design flaws that overlooked the diversity of data and failure to accurately interpret the complexities of real-world decision-making contexts in healthcare. These shortcomings led to the provision of potentially harmful patient recommendations, ultimately rendering the system impractical for reliable use in real-world healthcare scenarios.

The field of Human-Computer-Interaction (HCI) is well suited to tackle these tensions. There has been a growing focus on systematically examining various aspects of computer-mediated and computer-supported collaborative workflows with a push for how humans can collaborate effectively with AI agents. However, there are overarching questions that preface responsible integration of AI in the society that this thesis will seek to address. Who are the humans creating, deploying, and using these systems? How does the introduction of AI in human workflows change the nature of work amongst humans? What is the impact of AI on communities where it is deployed? It is imperative to study the changing nature of human-human collaboration with the introduction of AI agents in multistakeholder teams across the value chain of AI creation and deployment. This can uncover the nuances of who adopts AI and why, how they work around/with AI and each other to meet their goals.

I seek to answer these questions through this thesis by situating them in a broader question — **how do humans and AI collaborate across the AI pipeline from design to deployment?**. To progress the field, this comprehensive approach is crucial as it provides a distinct and holistic perspective of Human-AI collaboration through the AI development pipeline. I examine how each stage influences and shapes the downstream impact of Human-AI collaboration on society.



I draw my analysis from rich qualitative data, from hundreds of participants, collected in the Global South over five years. Additionally, I draw attention to high-stakes domains such as public health and future of work.

My thesis provides an understanding of Human-AI collaboration across the AI pipeline by tracing it all the way from development to deployment. First, I uncover the diversity and representation of people creating computing technologies by situating it in the context of representation and challenges of women in computing in India. Next, I trace the origins of data powering advances in AI and discuss the ongoing debates around data quality in AI by understanding the valuation of 'good' data across the data supply chain. Less is understood about the impact of large-scale AI deployments when AI is integrated in multi-stakeholder teams with varying organisational workflows. My thesis answers this question by examining the collaboration between NGO staff and AI developers for a large-scale AI deployment. Lastly, through research with automation vulnerable technicians and community healthcare workers, I argue for the need to examine perceptions and experiences of humans when AI systems impact their workflows.

I view the study of Human-AI collaboration as a part of a broader AI development pipeline - a complex, interacting system of people, organizations, culture, practices, and technology. This lens allows us to handle the integrated complexity and dependencies of various elements in the overall production and utilisation of AI. The study of Human-AI collaboration requires a recognition of the diversity of human and non-human actors with varying agency, with evolving nature of relationships across elements. This allows us to study Human-AI collaboration through an appreciation of the broader embedding of AI in social, cultural, organisational and political tapestry.

Drawing on this exploration of the dynamic networks of AI systems, human actors, and organizational workflows, through my work I argue that it is critical to extend this understanding to the diverse contexts of the Global South. I suggest that recognizing the varied cultural and socio-economic landscapes in these regions is essential for developing a more inclusive, context-sensitive view of Human-AI collaboration. Such a shift in perspective is not just a theoretical exercise but a practical necessity to ensure that AI technologies serve diverse global populations equitably, aligning with the central thesis of this work which seeks to understand and enhance the collaborative dynamics between humans and AI across various contexts.

I situate my theoretical and empirical explorations within the broader area of Human-AI collaboration in HCI and bring attention to the importance of studying high-stakes domains, such as public health, in the Global South.

This thesis represents a culmination of my scholarly journey in HCI. During my tenure at Google Research, I have had the unique opportunity to contribute to impactful projects that have informed Google's products and services, while simultaneously advancing the academic discourse in HCI through several publications at leading peer-reviewed avenues such as Conference on Human Factors in Computing Systems (CHI) and Conference on Computer-Supported Cooperative Work & Social Computing (CSCW). Reflecting on this scholarship through a broader lens, I have synthesized these prior publications with additional reflective analysis and theoretical development, specifically focused on examining Human-AI collaborations throughout the broader pipeline of AI development and deployment. In my analysis, I underscore the criticality of looking across the AI pipeline by reflecting on the interplay of users and creators, networks of human and nonhuman agential actors, and the configuration and reconfiguration of Human-AI collaboration in Global South. While I am proud of the contemporary relevance and practical impact of my work, this thesis also marks a phase of critical self-reflection and academic maturation, opening avenues for future research in Human-AI collaboration, which will be even more critical as we enter the era of Large Language Models (LLMs) that are increasingly pervasive in our personal and professional lives. It is my hope that this thesis opens up further inquiry into examining the role of AI in high-stakes domains, while centering the needs of historically underserved communities in the global south.

I have organised this thesis by grouping key questions that I seek to answer through my work.

- RQ1 Who are the creators of computing technologies?
- RQ2 How does data powering AI systems, come to be?
- RQ3 How does AI change the nature of work?
- RQ4 How do people perceive and experience AI?

I contribute to advancing HCI research discourse around Human-AI collaboration by centering its study across the AI development pipeline. I break down the study of *How do humans and AI collaborate across the AI pipeline?* into four key questions.

Who are the creators behind computing?

Frequently, researchers behind theories and work in HCI have been those predominantly situated in the Global North. The affordances of people in the Global South have been thinly understood. By Global South I mean countries primarily located in Africa, Latin America, and Asia, often characterized by their emerging economies and complex socio-political histories. The Global South provides a rich context where AI systems are being increasingly deployed for end-users, but we frequently fail to ask - who creates the AI, who is it meant for, how is the system created, what are the affordances of communities where these systems are deployed, what is the agency of users, how do they contest technology outputs, how do these technologies interplay with work?

I begin by addressing the first part of the question - who are the creators behind computing in the Global South by studying the gender representation in computing in India. The weight of equitable gender representation (and of historically underserved communities) cannot be overstated – the systemic lack of women in computing has led to systems that have historically failed women – facial recognition and autonomous cars that do not recognise women (Buolamwini & Gebru, 2018) to violent gendered language on chat-bots (Bardhan, n.d.).

My contribution is in examining the foundational stage of the AI development pipeline by examining the impact of social, cultural and organisational factors on women's participation in computing in India.

How does data powering AI come to be?

Data is a foundational aspect of ML and is viewed as a critical resource to improve performance and model capabilities. Yet, data work and data collectors are frequently undervalued. HCI researchers have examined data quality issues by studying annotator diversity, data cascades from the view of ML developers. However, little is known about scenarios when there are multiple humans working on (versions of) the same data, as they go from field to function, i.e. from their collection to their use in ML models.

My contribution is in addressing the second stage of the AI pipeline by examining the creation, curation and usage of data. Data is a critical resource that powers advances in AI – yet AI research is significantly model-centric. Through a qualitative study with data collectors, data stewards and AI developers that work on public health data in India, I provide a nuanced understanding of the tensions in the valuation of 'good' data across the pipeline. I argue that tensions in valuation have a downstream impact on how, where and what data gets collected for AI systems.

How does AI change the nature of work?

Human-Computer Interaction (HCI) researchers have systematically examined various aspects of computer-mediated and computer-supported collaborative workflows across domains including healthcare, public welfare, accessibility et cetera. The emerging field of human-AI collaboration has helped shift the focus of HCI from examining only interactions towards goal understanding and shared progress tracking.

A large body of work within the field of Human-AI collaboration has focused on improving decision making for humans and AI by enhancing explainable AI systems, pushing for improved interpretability and building trust through artefacts such as model cards (Mitchell et al., 2019) that facilitate communication and collaboration across diverse user groups. A subset of the work has focused on situating technologies within day-to-day workflows of domain experts such as healthcare practitioners and law officers. These works have focused on AI mental models, designing around AI failures, human-in-the-loop and AI-in-the-loop for improved agency.

While recent studies have articulated methods for optimizing human-AI interaction through guidelines and auditing mechanisms, and by uncovering humans' mental models toward underlying AI system capabilities, much of this work overlooks real-world organizational workflows that are layered and dynamic. This is a critical oversight because, in real-world settings, it is rarely the case that a single human works in isolation with an AI agent. Rather, AI systems are integrated into organisational structures across stakeholders across multiple levels, distinct domain expertise and those with different organisational incentives and structures.

My contribution includes an analysis of the role of AI as an actor and examination of the configuration and reconfiguration of Human-AI collaborations due to the introduction of a large-scale AI systems introduced in a public health setting in India. Overall, my thesis challenges, through several novel works published at leading HCI and AI conferences, the fundamental idea that studying human-AI collaboration through individual interactions is sufficient to achieve optimal team performance.

How do people perceive and experience AI?

One of the lesser understood phenomena is how the introduction of AI agents leads to emergent behaviors in human-human interactions. Do these agents act as equalisers, do they bring transparency in workflows, do they lead to conflict, do they lead to improved goal alignment? Moreover, how do humans adapt their strategies for collaboration when an AI agent is part of the team? Understanding this could reveal new kinds of workflow optimizations or bring to light unintended consequences of AI adoption in multistakeholder teams.

My contribution provides an examination of later stages of Human-AI collaboration by examining perceptions and experiences with AI through the experiences of vocational technicians in India, who are susceptible to trends in automation. I describe the interplay of social capital with the choice of technical training and uncover their perceptions towards an AI-enabled future of work.

Thesis Organisation

In Chapter 2 (RQ1), I uncover the interplay of familial, cultural and organisational factors that influence the entry and a precipitous exit of women from computing in India. The advancement of human-AI collaboration literature is intricately tied to the diversity of perspectives that inform it. A Western-centric male focus in computing research can inadvertently shape AI systems that are culturally biased, thereby affecting their efficacy and ethical standing in non-Western settings. This work uncovers various critical insights such as the lack of relatable role models for families of girls, who are key decision makers and influencers in the pursuit of computing studies. These insights informed the creation of a highly influential program at Google to promote culturally relatable role models with messaging aimed at families. This chapter draws on my paper published at CHI 2018 (Thakkar, Sambasivan, Kulkarni, Kalenahalli Sudarshan, & Toyama, 2018).

In Chapter 3 (RQ2), we delve into the intricate challenges surrounding data quality in public health systems, drawing from firsthand experiences in deploying real-world AI solutions for societal impact. The oft-cited adage ”garbage in, garbage out” has become a convenient refrain in the AI community, frequently deflecting the issue of poor data quality onto the shoulders of undercompensated data workers. To dissect this complex issue, we conduct an in-depth study of data flow across multiple stages of processing within the public health sector in India. Our analysis encompasses a diverse range of actors, from frontline health workers to data stewards and machine learning developers.

Drawing upon valuation studies, I uncover that the notion of ‘good’ data is inconsistent across the data supply chain. Moreover, I identify inherent tensions between stakeholders, as what is deemed ‘good’ data by one party often conflicts with the criteria set by another. We discuss the tensions in valuing and how they might be addressed, as we emphasize the need for improved transparency and accountability when data are transformed from one stage of processing to the next. This chapter draws on my papers published at CHI 2022 and EAAMO 2021 (Karunasena et al., 2021; Thakkar et al., 2022).

In Chapter 4 (RQ3), I present an ethnographic study of a large-scale real-world integration of an AI system for resource allocation in a call-based maternal and child health program in India. This work uncovers complexities around determining who benefits from the intervention, how the human-AI collaboration is managed, when intervention must take place in alignment with various priorities, and why the AI is sought, for what purpose. This work offers takeaways for human-centered AI integration in public health, drawing attention to the work done by the AI as actor, the work of configuring the human-AI partnership with multiple diverse stakeholders, and the work of aligning program goals for design and implementation through continual dialogue across stakeholders. This chapter draws on my papers published in IAAI 2023 and PACM HCI 2023 (Ismail, Thakkar, Madhiwalla, & Kumar, 2023; Verma et al., 2023),.

In Chapter 5 (RQ4), I investigate the perceptions and practices surrounding the future of work, particularly in the context of automation, among vocational technicians in India. This work is highly relevant with the emergence of Generative AI technologies and presses upon urgency of this issue, especially in the Global South, where labor markets are highly susceptible to automation trends. This work aims to fill a gap in existing HCI for development and future of work research, which has been predominantly western-centric, by examining perceptions and practices around AI-powered futures for a vulnerable, socio-culturally marginalised and often excluded community. This chapter draws on my paper published at CHI 2020 (Thakkar, Kumar, & Sambasivan, 2020).

In Chapter 6, I outline the key contributions of this thesis, provide reflective analysis on emergent themes that necessitate an examination of Human-AI collaboration through the AI pipeline, including the interplay of users and creators, the network of human and non-human actors, and discuss limitations and future work. This thesis has broad relevance to HCI researcher and adjacent fields in Responsible AI and AI for Social Good.

In this thesis, I use the pronoun ‘we’ to denote collaborative efforts with fellow researchers on specific papers, including those where I led the work. This is a deliberate choice to acknowledge and honor the collective contributions of the team. Conversely, I use the pronoun ’I’ when discussing the overarching narrative and synthesis of the thesis.

Chapter 2

Who are the creators behind computing?

Creators behind Computing

2.1 Introduction

For about a decade now, there has been public scrutiny of gender inequality in the technology industry, and for valid reasons. On the one hand, the tech industry has made progress, especially in its attitudes toward gender equity. In 2007, claims of gender inequality were met with significant pushback from male leaders (Sterlicchi, n.d.). Many years later, Silicon Valley leaders have expressed a strong desire to work toward a more gender-diverse workforce and are quick to acknowledge missteps. To offer just one example, Uber co-founder Travis Kalanick was pressured into resigning by his board in 2017, in part for overseeing a toxic culture for women at the company (Isaac, n.d.).

On the other hand, the ground reality has yet to catch up to these aspirations. In the United States, only 24% of information technology professionals are women (Raghuram, Herman, Ruiz-Ben, & Sondhi, 2017), even though women make up 46% of the workforce overall (Group, 2017). The situation is not much better in other developed-world countries. In the United Kingdom, women hold about 25% of IT jobs (Hagel, Schwartz, & Bersin, 2017) (cf. 46% of workforce (Group, 2017)). In Sweden, it is 22% (cf. 47% of workforce (**?**)).

As even critics of Silicon Valley's gender inequality acknowledge, much of the challenge lies with the so-called "pipeline" that feeds into employment. Women make up only 18% of computer science majors at U.S. universities, and frustratingly, this number is down from a peak of about 37% in 1984 (Henn, n.d.). (Again, similarly low figures apply in other developed-world countries. For example, women's enrollment in computer science in the U.K. is 13% (Raghuram et al., 2017).) Though researchers have dissected this pipeline problem in the United States (Gregg, "2015 (accessed August 8, 2017)"), it turns out that it is not a universal phenomenon. In India, women are 45% of enrollees in computer science programs (Raghuram et al., 2017). And in Qatar, women make up the majority of computer science students—70% in the case of Qatar University (Mark, n.d.). Strikingly, the corresponding number of Indian women in the HCI community is estimated to be 25-30%, in a field that is one of the very few areas of computing to have achieved gender parity in developed countries (Dray et al., 2013). These figures are especially surprising because they occur in places where the larger social context is far less gender-equal across many measures than any Western European country or the United States (Mark, n.d.). While gender-sensitive design practices, processes and values have been a major area of inquiry in HCI (Bardzell, 2010;

Bardzell et al., 2011; Dray et al., 2014, 2013; Kotamraju, 2011), less attention has been paid to the representation and experiences of female computing researchers.

In our paper, "The Unexpected Entry and Exodus of Indian Women in CS and HCI", published at CHI 2018, we consider the potential causes of these apparent paradoxes in the case of India. We investigated the complexities of the phenomenon through qualitative research of Indian women scientists and practitioners in computing at various stages in their careers. Our inquiry was driven by critical questions: What factors enable India, despite its broader gender inequalities, to have relatively high enrollment in computer science? How long does the strong pipeline continue?, and can it offer lessons to other countries?

The novel contributions of this paper were: First, we explored how women in the Indian computing sector experienced gender-related issues across their lifetime, from undergraduates to accomplished researchers, and how gender discrimination and representation evolves through life stages. Against a backdrop of significant literature in developed-world contexts, we provided a range of culture-dependent differences. Second, we found a number of substantial differences in perceptions— by women and by Indian society as a whole —of HCI compared with computer science. To my knowledge, these had never been investigated formally anywhere in the world and have sparked a broader interest from HCI researchers since the publication. Finally, I argue that by understanding the technology industry itself as a socio-technical system, we can move towards a foundational understanding of the downstream impact of technologies such as AI, by examining *who are the creators behind computing?*

2.2 Method

We gathered data on Indian women and computing through 37 semi-structured interviews and 3 focus groups. We conducted a total of 47 hours of interviews and 8 hours of observations.

The semi-structured interviews were designed to include adult women in all stages in the tech sector and spanning industry IT jobs as well as academic research, and including a mix in terms of geographical origin and site of university education (India and abroad). The interviews were conducted with women undergraduate (UG) and graduate (PG/PhD) students in computer science, IT employees (IT), and both junior (JR) and senior (SR) professors, industry research scientists (IR) in computer science and UX Researchers (UX) in Industry. We additionally interviewed senior HCI professionals in India who pioneered HCI in India, both in academia as well as in industry. Participants were recruited through snowball sampling, Android developer relations, Women Techmakers' award winners, approaching through women in coding events, and contacting faculty and professionals via university websites. University students (undergraduate, Master's, and PhD) were those who attended top-tier national and regional universities in India, some of whom also had experience with graduate school in the United States.

I conducted almost all of the interviews were conducted in-person, though in a handful of instances, the interviews were conducted via video-teleconference. All interviews were audio-recorded and transcribed. Interviews were conducted in English, owing to the fluency and language of choice of our participants.

2.3 Related Work

Our work relates to several existing literatures. The largest of these is the significant literature examining the many factors leading to low representation of women in CS careers and education in the United States. A much smaller literature considers similar questions in the Indian context. Finally, we also review the feminist scholarship around science professionals, mostly with respect to developed-world contexts.

2.3.1 Factors affecting CS careers

Women in the United States face stiff headwinds throughout their lifetimes. The problems begin with socio-cultural factors such as gender stereotypes(Hill, Corbett, & St Rose, 2010) and subtle biases against girls in early education (Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012). These forces continue throughout formal schooling through university education (Cheryan, Master, & Meltzoff, 2015). Then at work, related socio-cultural factors and sexism lead women to disproportionately consider leaving the workplace (Cohoon, Wu, & Chao, 2009; Gregg, "2015 (accessed August 8, 2017)"). Issues such as a gendered work environment, childcare responsibilities, and a lack of female role models are all cited as prominent factors leading to the attrition of women in science professions more generally (Tapia & Kvasny, 2004), in research(Monroe, Ozyurt, Wrigley, & Alexander, 2008), and in academia (Wolfinger, Mason, & Goulden, 2008) (Finkel & Olswang, 1996).

2.3.2 Employment and education in the Indian context

Worldwide, there is wide variation in the participation of women in CS, with few clear patterns to explain the differences (Galpin, 2002). But apart from statistics, there has been limited work on women in CS in the Global South in context of epistemological histories, socio-cultural factors, and resource and capital concerns.

India has one of the lowest participation rate of women in the overall labor market at 27 % (Bank, 2016; Das, Jain-Chandra, Kochhar, & Kumar, 2015; Klasen & Pieters, 2015). Women were a negligible presence in IT industry in the 1980s, but their presence has grown dramatically since then, reaching approximately 30% of IT workers (Varma, 2010). That is to say, women in Indian IT have matched the rate of participation of women in the general workforce. The change has been attributed to market forces and changing social norms related to the booming IT economy within the constraints of a patrifocal society (Gupta, 2012). However, not all is rosy. Women participation is disproportionately in junior positions (80% by one measure (Raghuram et al., 2017)), and the same "leaky pipeline" afflicting U.S. tech companies appears to exist in India, as well (Gupta, 2015). Venkatesh et al. (Venkatesh, 2015) find that in STEM research careers, promotion depended more on gender than on productivity. Gupta (Gupta & Sharma, 2003) notes that a social prejudice against working women greatly limits women from top positions.

Though the above work addresses Indian women's technology careers in piecemeal fashion, no systemic analysis of their career trajectories has been conducted. Our study performs such an analysis using a lifestage model.

2.4 Findings

We analyzed factors affecting Indian women in their journeys through computing education (undergrad, grad and PhD) and careers (both IT and research or academia). A broad overview of our findings is as follows: Initially, through undergraduate and master's programs, socio-cultural pressures encourage women to enter CS as a field, with parents in particular expecting better financial and marital prospects. However, as women advance through further education and early careers, their ambitions meet familial duties and societal expectations. Marriage and childcare norms in India saddle women with a burden even heavier than that felt by their peers in the U.S.A. This weight leads to a drastic drop in the participation of women which starts unusually high: from 45% among undergraduates, to 30% in the IT workforce, to less than 10% of CS researchers, and less than 1% of C-level executives in the IT sector. The situation for women in HCI differs in some particulars, though the overall patterns are similar.

2.4.1 Undergrad: bright futures *through* CS

One of the surprising findings of our interviews was how infrequently participants mentioned any concerns about their ability to do technical work, while growing up, or a belief that they were less capable than their male peers in mathematics, science or other intellectual endeavors. The lack of biases at early stages leads to a level playing field. This perception is in sharp contrast to the U.S.A context (Correll, 2004; Nosek et al., 2009; Spencer, Steele, & Quinn, 1999).

India's booming IT industry with its high paying jobs, consistent growth and a promise of upward mobility is the prime employment destination for middle class India. Families are keenly aware of these economic prospects, which also affect partner selection in marriage—a process that is highly formalized in India.

CS is in strong demand at the undergraduate level. Studies have found that parents have significant say in their children's career decisions (Mukhopadhyay, 1996), our findings confirm this. Families look to their children to uphold and promote their social and economic status (Varma, 2010). Lower and middle-class families generally expect their children to provide them with security in old age. Participants pointed out that self-sufficiency and financial independence (even from their husbands and parents) was of importance to them to avoid the burden of any form of harassment in the future.

Providing context to gender imbalance in CS within elite institutes of India (Escueta, Saxena, & Aggarwal, 2013 (accessed August 20, 2017)), our participants reported not feeling "smart enough" as compared to their peers upon joining a premier institute. However, we did not observe any gendered self-confidence issues moving to Master's and Ph.D. in elite universities.

IT-related jobs were considered "safe" for women because they are white-collar jobs in which employees interact with a narrow, educated strata of society. Physical safety is a deep concern for women in India (every two minutes, there is a crime against a woman (Desk, 2015)), and particularly for their parents, who tend to imagine sexual assault as an ever present threat outside of their immediate communities. Parents thus saw in CS innocuous desk jobs, air-conditioned offices, and secured buildings with gated entrances. Even compared with other engineering disciplines, CS was a safe bet; mechanical engineers, in contrast, may be required to do physical tasks on a shop floor or to interact with people doing menial work.

Our participants reported that their parents discouraged or forbade any career involving physical hardships, late nights, or the perceived possibility of physical assault.

2.4.2 Master's degrees: recovering agency

Master's degrees in CS were seen to have the same advantages as CS undergraduate degrees, but with greater force: even better job and marriage prospects. Our participants reported that having a master's degree was a matter of honor for the family and a boost to the marital profile of the participant. (Arranged marriages in India are often preceded by an offline or online exchange of "profiles" – effectively marriage-related resumes.) Several participants noted that their degrees were prominently printed on wedding invitations, and in some cases the degree signaled the dowry paid by the daughter's family to the groom.

A Master's degree provided the ability to find more meaningful work through specialization. In many cases, the Master's degree offered women more freedom in a socially acceptable way to postpone marriage or work. Funding was crucial at this stage. Since funds were allocated by families for larger expenses like weddings, justifying the tuition fee was noted to be challenging for many (despite availability of student loans.) Scholarships enabled many participants to further Master's degrees, when other financial means were limited.

2.4.3 Ph.D.: marriage pressures and isolation

Only 32% of the Ph.D. cohort in CS-related fields in India are women (of Human Resource Development, 2015-16) as against 47% at the master's level. A common theme among our participants was that while there was encouragement to pursue undergraduate and master's degrees in CS, it did not extend to Ph.D. These concerns were largely related to issues around marriage and adult dynamics in India.

Families actively discouraged their daughters from pursuing Ph.D.'s out of fear that they would became "too educated to be married off". In many cultures, sociologists have identified a tendency for women to "marry up" (for Family Studies (IFS), "2016 (accessed August 23, 2017)"), but in India, this inclination is formalized by the norms of arranged marriages, in which it is assumed by all parties that the bride will be of lesser status than the groom. The same norms, of course, expand opportunities for men with more education. And, these norms are often internalized by women themselves.

In addition to the fear of over-education, there was an age-related pressure to marry at the "right time". Parents felt an obligation to ensure their daughters' marriage and it was expressed as an overt or covert pressure for their daughters to marry soon after obtaining bachelor's degrees. Considerable stigma was attached to single women past their mid-20s, with parental concerns ranging from social perception, diminished potential for marriage, or threats to future childbearing. Balancing the need to appease parents with the pursuit of academic responsibilities led to mental stress and anxiety, and depression in some cases.

In case of married participants, their life decisions were often determined by the husband and his family. The lack of authority over one's life led to complex negotiations across desires, value systems and priorities. Many of our participants reported the struggles to convince their in-laws about temporary relocation to finish their doctoral programs. Participants spoke about colleagues from their university who did not continue with a research career because of lack of support from in-laws.

The lack of aspirational role models (for upward mobility) outside family was a concern not only for oneself, but to show as an example to their family as well. Female role models were difficult to find, especially in families where the participant was the first person to break socio-cultural cycles or for participants hailing from rural and peri-urban areas. Our participants noted that the presence of a role model from a similar community/region as theirs would help them start a dialogue with their parents/partners about their independence and careers.

2.4.4 Careers and practice: sharp drops due to family life

Marriage and childcare have consequences for a woman's pursuit of her choices in India. In our study, women's career choices had multiple stakeholders (husband, parents, and in-laws) in the decision making. Practices and policies around maternity leave further led to discrimination.

In India's patriarchal society, the onus of domestic labor fell unequally on women, exacerbated by limited maternity leave policies. Despite juggling professional and familial duties, many expressed that credit was unfairly not attributed to them.

Structurally and culturally, the Indian IT industry presents disproportionate challenges for women. The emphasis on late night shifts was challenging for women with families; physical safety issues present obstacles to mobility and social acceptance. These factors attribute to the drop in the representation of women from 80% at entry level to 25% at managerial positions to less than 1% at the C-Suite (Raghuram et al., 2017).Despite India's IT industry boom, the structural design of the technology sector is antagonistic to women's success. Success in this work environment and team culture depends on physical presence in meetings, brainstorms and check-ins.

Many female researchers in our study were married to fellow researchers. In this two-body problem, women felt side-lined compared to their partners and were viewed as secondary intellectual contributors. They further encountered discrimination in their tenures compared to other equally qualified males.

2.5 Discussion

Analysis with central question

The genesis of the Human-AI collaboration is at the onset of the AI development pipeline and I do this by asking *who are the creators behind computing.* The lack of diversity in the development of technologies can lead to biases in technologies since a homogeneous group of creators might unintentionally induce their own biases within these systems and fail to represent diverse needs and perspectives. Hence, the study of moving towards effective Human-AI collaboration, begins at the start of the AI development pipeline.

Taking the case of women's representation in computing in India, I discuss the role of socio-cultural and organisational elements and it's direct impact on the women who participate in the development of computing (and AI) technologies. The diversity (or lack thereof), stemming from familial decisions and cultural narratives, in the pipeline significantly impacts the nature of AI technologies produced. Through this work, I aim to reflect on the broader arc of my thesis by refocusing downstream issues such as bias, at the very start of the development pipeline. This work is limited to examining women's participation in India and could benefit from broader engagement of historically underserved and marginalised communities across the global south.

Design and Policy Recommendations

Through this work, we suggest a number of ways in which technology companies, Indian policy-makers, and academic bodies & institutions could support women in computing careers.

Solutions used in developed countries to address the "leaky pipeline" in the workplace should be considered in Indian technology firms and universities, as the problems seem to be similar and at least as problematic. Everything from gender-equal parental leave policies, to diversity workshops, to systematic recruiting outreach to women is likely to have some impact. Re-integration back into work and fair, consensual assignment of work roles after long leaves are crucial to equitable treatment. Remote work is problematic without policies and a shared work culture on how to inclusively integrate remote employees. Everyday biases could be significantly reduced by regularly sensitizing employees about their unconscious biases, especially for those biases resulting out of ignorance and blind adherence to social norms (Google's bias busting workshops are a good example ((Manjoo, 2014)). Fair and inclusive environments further enable equal female participation in various aspects of HCI research and practice, such as fieldwork, brainstorming, design critique, or usability studies. Ultimately, gender parity of HCI researchers leads to raising the importance of gender in design practices, making gender a subject of research inquiry, and checking biases in design solutions—enabling better design for end-users as a whole ((Bardzell, 2010; Bardzell et al., 2011; Dray et al., 2013)).

To erode societal stigma against PhDs for women, parents of young adults and junior practitioners/researchers may be targeted in re-shaping deeply ingrained societal attitudes on women's gender roles. This may take the shape of easily accessible programs for parents to code, enabling 'take your parents to work' day more widely, and open house fora (such as the Indian Institute of Science's Open Day initiative, which introduces children and parents from all over the state to science). These initiatives expose parents firsthand to the benefits of being in senior management, instill empathy, and make computing more human-friendly. It may be helpful to run social marketing campaigns that highlight women with PhDs who also have "successful" families. Of course, any such messaging must take care not to further stigmatize unmarried women. There may be technology solutions at the margins, mediating between parents and their daughters. For two decades now, many arranged marriages in India begin online with matching sites such as Shaadi.com. Such sites could offer options for women's families to identify men who support prospective partners' further education or prefer highly qualified women.

Relatable role models serve an important role for the *entire* family as much as they do for the scholar or student. However, as noted in the findings, local, aspirational role models are difficult to locate within social circles. (Near impossible if the female scholar is from a village and is the first person to break the patriarchal norms around education.) Technology platforms have an important role to play here in serving aspirational local figures and their stories, across a spectrum of gender roles, class, language, and sexuality, and improving discoverability of this content.

One outcome of our findings that deserves closer scrutiny is the apparent lack of gender-based ability disparities in India. Girls seem just as confident as boys in STEM subjects. More research is required, but the developed countries could undoubtedly learn from the strengths of India's approach, of early introduction of computing, lack of gender stereotypes in schools, and positive aspirations around the fields of computing. Other developing countries could learn from the upward mobility associated with computing, financial independence for women, and the creation of a prosperous IT job industry.

Chapter 3

How does data powering AI come to be?

When is Machine Learning Data Good?

3.1 Background and Motivation

Working with data, or data work, is of emergent interest to the field of Human-Computer Interaction (HCI), where recent research has actively investigated the challenges that arise around data procurement, organisation, management, visualisation, and more across a range of domains (Møller, Bossen, Pine, Nielsen, & Neff, 2020; Muller et al., 2019; Passi & Jackson, 2018; Sambasivan, Kapania, et al., 2021a). Many of these studies focus on the public sector, such as healthcare and public welfare, to foreground the design challenges that result from data work (Karusala, Wilson, Vayanos, & Rice, 2019; Mao et al., 2019; Sambasivan, Kapania, et al., 2021a). Møller et al. draw attention to the role of the human(s) behind the data who perform the work of data collection and processing (Møller et al., 2020). The workflows and collaborative practices of data scientists, machine learning (ML) developers, and data annotators are also topics of growing interest in HCI (Muller et al., 2019, 2021; A. X. Zhang, Muller, & Wang, 2020). I expand on this scholarship, investigating what happens when there are multiple humans working on (versions of) the same data, as they go from field to function, *i.e.* from their collection to their use in ML models.

I, with my collaborators, explored the second stage of the AI development pipeline by addressing the question *How data powering AI comes to be?*. This question is central to the study of Human-AI collaboration as it centers the data infrastructure powering AI, worked on by various stakeholders across organisations. We do this through an examination of datafication efforts in public health, augmenting a body of work within HCI that has progressively been engaging on topics around frontline health (e.g., (Batool, Toyama, Veinot, Fatima, & Naseem, 2021; DeRenzi, Dell, Wacksman, Lee, & Lesh, 2017; Ismail & Kumar, 2019)). This work, titled "When is ML Data Good? Datafication in Public Health in India " was published at CHI 2022.

Prior work has reported on the disconnects that exist between the locally relevant insights of frontline health workers (FHWs) and the information sought by state and healthcare authorities, leading the data collected by the FHWs to be viewed by other stakeholders as inaccurate, incomplete, or simply unreliable (Batool et al., 2021; Ismail & Kumar, 2018). I provide a deepened understanding of the disconnect in *valuing* by investigating the perspectives of multiple stakeholders or *data workers*, as data are collected by FHWs, passed on to *data stewards*, who prepared them for the *ML developers* putting them to use in ML models.

We presented findings from interviews held with the range of data workers employed in or contributing to the public health domain in India. We analysed the data collected from these interviews to arrive at an understanding of the data *supply chain* in public health, or how data changes hands from stakeholder to stakeholder. I draw from the field of valuation studies, particularly Heuts and Mol's discussion of valuing in a supply chain of tomatoes, and analyzed how data are valued by stakeholders differently at various stages of their collection, processing, and used (Heuts & Mol, 2013). Our participants shared how they know whether data are good, and the work they must do to make the data good.

We reflected on (a) the interdependencies of data work through the supply chain, (b) the values and priorities of different data workers participating in this supply chain, and (c) the need to make the labor involved in data work more visible.

I begin by situating our work in the context of prior research on data quality, frontline health and datasets, and valuing of data and data work. Our findings outline the data supply chain in our focus area of public health ecologies. Starting with the ML development stage, we describe how data changed hands from one stage to the next, the ways of valuing data across the supply chain, and the work undertaken to attain that value.

Drawing on these findings, we then discuss where tensions in valuing arise and are visible, and implications for the generation and curation of ML datasets. We argue for greater alignment in valuing through transparent and accountable processes and structures that empower data stewards and data collectors who are part of the supply chain. Our research insights seek to advance scholarship towards understanding valuing behaviors and practices in data work, paving the way for greater transparency and accountability overall in ML development.

3.2 Method

The goal of our research was to attain a deeper understanding of the origin and evolution of data in the public health system in India, before it comes to be used for ML. Our study draws from 46 semi-structured interviews with individuals involved in different parts of the data supply chain. Interviews were conducted from May 2020 to August 2020. Participants include data collectors, data stewards, and developers linked to ML applications being developed for public health such as healthcare resource allocation, improving public health program adherence, health outcome predictions, community health surveillance and healthcare worker evaluation.

I conducted interviews via phone calls and video calls. Interview questions with FHWs were focused on understanding their backgrounds in day-to-day workflows, public health data collection and analysis, incentives, system design, communication workflows, community interactions and their interactions with technology for their work. Interviews with data stewards and ML developers were additionally focused on capturing their experiences and challenges on working with data entry, processing, analysis, and modelling. Interviews lasted 45 - 60 minutes with all participants. We followed standard procedures of informed consent (in participant's local language) and anonmyization for data analysis. I carried out inductive qualitative data analysis to summarise and interpret the interview data and emergent themes were iteratively refined (Thomas, 2006).

3.3 Related Work - Data Quality in ML

The quality of training data has a significant impact on the quality of ML algorithms developed (Gudivada, Apon, & Ding, 2017). The aphorism 'garbage in garbage out' has frequently been used in relation to ML (Geiger et al., 2020). Recent research by Sambasivan et al. on data work among ML practitioners reports that data work is highly under-valued compared to model development, and as many as 92% of respondents reported experiencing data cascades, where issues with data collection led to even greater challenges on the development end (Sambasivan, Kapania, et al., 2021a). Our research augments this work by examining the supply chain of data for ML in the context of public health.

Data fields themselves can also be problematic, as Keyes points out, such as binary conceptualisations of gender in gender recognition algorithms (Keyes, 2018). Such challenges are further compounded when people share data across contexts, but taken-for-granted norms and standards of data from these contexts are not included (Neff, Tanweer, Fiore-Gartland, & Osburn, 2017). These broader questions around data gathering have long been the focus of the field of critical data studies, which has closely examined how the context around data—the people, institutions, instruments, policies, finances, and more—impact data collected and their use (Denton et al., 2020; E. S. Jo & Gebru, 2020; Katell et al., 2020; Kitchin & Lauriault, 2014).

Researchers have recently called for paying more attention to data gathering and processing, such as by documenting how data were labeled (Balayn, Kulynych, & Guerses, 2021; Geiger et al., 2020; Hutchinson et al., 2021; Muller et al., 2021). Denton et al. have outlined a research agenda for documenting the genealogy of ML data, investigating the origin of and values engaged in the collection of benchmark datasets such as ImageNet (Denton et al., 2020). We build on such prior work by investigating the origin and flow of data through its lifecycle comprising of several stakeholders.

Within HCI, emerging research has begun to explore the human-centered design of AI/ML systems in healthcare. For example, Beede et al. undertook an ethnographic study of a deep learning system for diabetic retinopathy in hospital settings (Beede et al., 2020). Okolo et al. have highlighted the need for explainability in AI systems to help communicate model outcomes to people with low literacy (Okolo, Kamath, Dell, & Vashistha, 2021). Others have studied how AI could support the routine work of clinicians (Pedersen & Bossen, 2024; Yang, Steinfeld, & Zimmerman, 2019), and help with collaborative decision making across medical experts (Cai, Winter, Steiner, Wilcox, & Terry, 2019). Despite the reliance of such interventions on large datasets, little research describes how public health data are collected, digitised, aggregated, and then processed into machine learning datasets.

HCI researchers have examined the role of collaboration between data science workers for data work (Batini, Cappiello, Francalanci, & Maurino, 2009; Kshirsagar et al., 2021; Muller et al., 2019; Neang, Sutherland, Beach, & Lee, 2021; Passi & Jackson, 2017, 2018; Piorkowski et al., 2021; A. X. Zhang et al., 2020). Discussions on collaborative data work have focused on collaborations among data science workers employing collaborative practices across data science workflows (A. X. Zhang et al., 2020). Data science workers also face challenges working with data and spend significant time data cleaning, data wrangling and working with 'dirty' data (Guo, Kandel, Hellerstein, & Heer, 2011; Sutton, Hobson, Geddes, & Caruana, 2018). We build on this work to discuss the collaboration practices within each stage of the data workflow and analyze the role of collaboration across the data workflow through the lens of valuation to make data 'good'.

3.4 Findings

In their study on the supply chain of tomatoes, Heuts and Mol describe the work performed on the tomato and the valuing by stakeholders at each stage to make the tomato good (Heuts & Mol, 2013). Taking inspiration from their analysis, we viewed data as part of a supply chain, where data are handed over from one set of stakeholders to the next.

3.4.1 ML developers: Making Data Fit for Purpose

Our interviews with ML developers revealed that they spent a substantial amount of time making data good for developing ML models and operations to make data fit for purpose. Participants noted the need to pay attention to how their data work could impact their models, and sought to contextualize the dataset to make appropriate data operations. They recognized their limited understanding of the context in which data were collected, and struggled to contextualize their models for solving relevant problems using ML.

3.4.1.1 Operations to make data good

Our ML developer participants reported facing several data quality issues for model training, and shared their concerns about lack of visibility into the data collection process. ML developers reported that the key constituents of good data included structured and standardised feature-rich data with validated ground-truth labels that would help them build robust ML models.

Missing data fields were a commonly reported issue among our participants. Other issues included large clusters of near-perfect values in a dataset and swapped entities (for instance, height with weight). Developers reported receiving crude datasets that had not been collected for ML applications, typically from non-profit organisations or public health agencies. ML developers noted the process of data transformation towards achieving better modelling.

> *"My goal is to try and reach the highest accuracy and noisy data makes that really difficult. I spend a large portion of my time in preparing the dataset even before getting to the modelling. If it were up to me, I would not do this task but it is a critical part given the quality of data."—P41, ML developer*

3.4.1.2 Contextualizing data and models

Participants worked with partner organizations to decide on data operations. They struggled to find reliable documentation on how data had been collected and formatted.Other challenges include connecting multiple datasets for improved model performance and evaluation. Additionally, they struggled with handling multiple languages in the dataset, since textual data, was collected in regional languages. Labelling datasets was reported to be critical yet challenging, requiring rules or heuristics to produce the correct labels and ground truth for validation, requiring acccess to domain expertise and an understanding of their biases and subjectivity. Anonymizing data is a frequent task for developers, often not performed by NGOs or health agencies due to technical limitations.

ML developers frequently assessed the feasibility of working on a specific project based on the availability of high-quality data. In other cases, ML developers preferred to work in collaborative teams that could potentially work with partners such as NGOs and public health agencies with access to high-quality data. They strived to understand if their models were solving the ‘right’ problems, learning the right data features, performing bias analysis, and validating model outputs and performance with domain experts.

All developers we interacted with shared that they had largely been taught to work with high-quality datasets during their training, starkly different from the realities of working with public health datasets. Data work was not directly incentivised, rather algorithmic innovations were critical for scientific publications and career growth.

3.4.2 Data Stewards: Aggregating Data to Meet Reporting Requirements

Before data reached the ML developer, they were aggregated and processed by data stewards. The data stewards we interviewed performed a range of operations on data collected by FHWs through paper forms, daily diaries and notebooks. These operations included data entry, processing, and analysis.

3.4.2.1 Making data legible and complete

Data stewards converted paper forms into structured data to meet reporting requirements for health information systems. They focused on completeness and readability. However, they reported to not receiving ‘’good’ data. They addressed issues like missing data fields, data fudging, data manipulation, and misrepresentative entries.

Data stewards, like ML developers, prioritized data privacy and handled sensitive fields (e.g., caste, religion, gender). They anonymized and aggregated data, ensuring unique IDs for inter-dataset connections. However, despite these efforts, ML developers often undertook further privacy measures due to varying procedures.

3.4.2.2 Organizational structures for meeting data quality

We found that there were several organizational structures in place to check the quality of data collected by FHWs, though there were also significant gaps and misaligned incentives that impacted data quality. Data stewards were geographically located far away from FHWs and without direct communication with FHWs, direct oversight or influence on the data work of FHWs.

We found that they spent significant time validating data from FHWs through manual checks, going through the data provided by a single FHW to develop a broad impression of the quality of the data they were collecting. They reported any specific instances of continued data quality issues with an FHW to their supervisors. Financial incentives for FHWs were perceived by data stewards to skew data focus to specific programs, creating gaps in reporting and influencing data operations for ML developers.

Data stewards performed practices to enable better health reporting mechanisms for state officials. However, they lacked incentives for good reporting, did not have the autonomy to define reporting structures, and were overburdened and had to operate amidst resource constraints. They also had to work with limited feedback and low visibiity into the impact of their data reporting work.

3.4.3 Data Collectors: Collecting Data as Part of Everyday Workflows

FHWs performed data collection alongside their primary role of healthcare outreach. Organizational structures and the local context shaped data collection by FHWs and their valuing of the data.

3.4.3.1 Organizational structures driving data collection

We found that data collection was performed by FHWs for two primary uses—generating and maintaining public health records of care coverage, and conducting specialised surveys. Data collection was perceived as a tedious and tiring process by our participants, spanning multiple forms, surveys, diaries. FHWs shared that the data collection process had several components which were opaque, repetitive, and redundant (such as multiple forms requiring similar data fields). They received financial incentives for each task completed, and data were used to demonstrate completion of their care work. As described by P4, *"…register should be tip-top (polished in presentation) and complete."* 'Good' data for FHWs thus reflected their performance on healthcare tasks, and were well presented and complete.

FHWs also received financial incentives to procure new data for specialised program surveys. However, low incentives (e.g. USD 14 per month for a reported program) may have impacted their motivation to collect high quality data in such cases.

The performance of advanced FHWs was also tied to data, though their work was not based on financial incentives. They received fixed salaries and held the perceived prestigious status of a ‘government employee’ (government jobs are frequently sought after in India). Advanced FHWs oversaw a cluster of villages (typically five to six) and were responsible for the overall performance of their cluster.

We found that FHWs received training at the start of their job on maintaining diaries and structured notebooks for data work, but lacked specificity for varied surveys, leading to challenges in data consistency

FHWs shared that they did not receive any regular feedback or communication on their data work which made it difficult for them to understand the use and monitoring of their data collection efforts. Despite not understanding immediate gains and value of their data work, they complied with the processes to avoid non-compliance repercussions

Despite limited feedback, FHWs found value in good data for care work, notably in creating ’due lists’ for routine health services and resource management, aiding in accurate distribution and accountability of resources (e.g. medicines).

3.4.3.2 Collecting and protecting sensitive data

FHWs were tasked with the goal of increasing access to health services in marginalised communities (e.g., caste minorities, religious minorities, and tribal communities). However, access to marginalised communities was restricted or limited due to social, cultural, and physical factors described below. They faced physical safety concerns in traveling to challenging terrains, geographical, as well as socially, in building trust with communities where their access was limited (e.g. indigenous communities, diverse religions and castes). This had a direct impact on sampling during data collection.

Stage	Valuation of 'good' data
ML Developers	- Structured and standardised data - Feature-rich data with validated labels - Improved model performance - Validation with field and domain experts - Contextualising bias and fairness
Data stewards	- Structure and completeness for complete entries on information management systems - Methods to decipher bad quality data - Ability to share feedback with data collectors - Synchronous data updation - Feedback on eventual data use
Frontline Workers (data collectors)	- Data work as a reflection of completion of tasks - Presentation of data registers - Training for new data workflows - Regular feedback on data work and use - Data to improve care work

Table 3.1: Valuation of 'good' data across stakeholders

Even before gathering data from a specific household, FHWs shared that they spent considerable time building their relationships and trust in that household. This was critical because FHW worked and collected data on topics considered to be sensitive and personal.

3.5 Discussion

Analysis with central question

The importance of high-quality or *good* data cannot be overstated in the AI development pipeline. Viewing data production as a part of the AI development pipeline to study Human-AI collaboration, helps with a deeper appreciation for how each stage can be taken apart, while also allowing for interactions between them to result in a synthesis of ML datasets. Data is not just a static element but part of a evolving process, where its value and meaning are continually negotiated and redefined by various actors, reflecting the evolving nature of Human-AI collaboration.

We found that ‘data work’ was often undervalued and there are tensions across the supply chain of data production and usage. Each stage in the data supply chain interacts and influences with others, yet each stage also has its distinct characteristics which are influenced by non-human actors such as organisational incentives, training, social hierarchy, and cultural norms around access. These unique characteristics at each stage showcase the varying and dynamic agency of human and non-human actors in the data supply chain – directly impacting what constitutes ‘good’ data – which is not just different at each stage but at tension with each other. Each stage of the data supply chain served a very specific purpose for human actors, from showcasing completion of care work for data collectors and meeting reporting requirements for data stewards to improved model performance for developers.

I further reflect on the affordances around greater transparency and accountability, enabled by examining Human-AI collaboration across the AI development pipeline. The interplay of human and nonhuman actors through the AI pipeline aids in understanding the formulation of what constitutes ‘good’ data, and the work performed by each stakeholder to attain ‘goodness’. I argue that this examination enables us to step back from the downstream application of AI, to reflect on how AI systems are created, by whom, and where. An analysis across the AI pipeline enables one to examine who bears the burden of creating transparency and who is accountable across the pipeline for the downstream effectiveness of Human-AI collaboration.

Our findings described the overlaps and conflicts in valuing by different stakeholders in the data supply chain. Below I discuss the tensions in valuing across stakeholders, and implications for work on ML datasets.

3.5.1 Valuing Across the Data Supply Chain

A lens of valuing allows us to move beyond arbitrary notions of data quality to focus on what are desirable and achievable data quality goals in a given context based on existing practices across the supply chain. Tensions in valuing occure across three key aspects —in data transformations performed to improve data quality, data contextualization, and as a result of organizational incentives.

Data Quality: ML developers and data stewards faced challenges with 'noisy' data from collectors, affecting ML development and reporting through information management systems. However, these issues are embedded in the ecological context (organizational, social, and cultural) of data collection where valuation of 'good' data is at tension across the supply chain. For instance, data collectors faced challenges in accessing all communities equitably due to safety risks, challenging geography, or social dynamics around caste or religion. These access challenges could be perceived as data distribution irregularities by developers. Hence, it is imperative to develop a shared understanding and valuation of 'good' data across the supply chain.

Data contextualisation: Building ML models far from the context where training data are collected and aggregated, as we found, can adversely impact ML development, leading to challenges with labelling data and validating results, and unclear algorithmic outcomes. Data collectors enabled data collection due to their community trust but were unaware of of downstream use which led to 'good' data as being a proof of care work, impacting their data operations. It is imperative to view data collectors as community leaders and engage them in the design of data collection workflows and ML deployment.

Organizational Structures: Data are valued differently based on the organizational structures and incentives for each stakeholder in the data supply chain. ML developers were motivated and incentivised to develop novel models, perceiving data work as mundane (Sambasivan, Kapania, et al., 2021a). Data stewards were constrained by limited organizational resources, and unavailability of statistical experts and automated data cleaning workflows. Even earlier in the data supply chain, data collectors were provided with limited training on data workflows and lacked feedback on data use, resulting in confusion about data collection workflows and eventually leading to data quality issues. Organizational structures could be designed to better support data work, such as through incentive structures that align with data needs, greater transparency around data workflows, and developing data literacy.

3.5.2 Transparency and Accountability in Data Transformation

We view transparency and accountability from a relational perspective, where the level of transparency and accountability desired is determined through negotiation across multiple stakeholders, based on what is possible and appropriate in a given context.

Prior work on transparency in ML datasets has largely focused on the development of artifacts and processes that document dataset development or summarize the dataset (Denton et al., 2020; Pedersen & Bossen, 2024). ML developers often lacked insight into data collection and previous operations, which could have been addressed through documentation. However, generating documentation is challenging since data are collected and aggregated over multiple cycles and real-time access to data is limited. I propose documenting stakeholder valuation as a more feasible approach.

Our findings demonstrated how data stewards and data collectors lacked visibility into data flows which resulted in conflicts in valuing. Additionally, receiving adequate and feedback on data work were reported as challenges. Engaging all stakeholders in developing a shared 'good' data taxonomy could align goals and motivations in the data supply chain.

Our findings underscores the need for better accountability structures to improve ML data quality. We build on prior work on accountability for model outcomes (*e.g.* (Hutchinson et al., 2021; Raji et al., 2020)), and extend focus to the development of ML datasets. We employ Bovens' definition of accountability as *"a relationship between an actor and a forum, in which the actor has an obligation to explain and to justify his or her conduct, the forum can pose questions and pass judgement, and the actor may face consequences"* (Bovens, 2007). To apply this definition effectively, we need to determine *who* is accountable *to whom*, *for what*, and *how*.

At the individual level, valuation of data work for all stakeholders and thus, accountability was closely tied to organizational structures, incentives, and personal motivations. We suggest that accountability of ML outcomes be shared between ML developers and the organizations providing data, given the investment of these two groups in the outcome. ML developers could be held accountable by the intermediary organizations for the performance of their models, not just in the lab but also in the real world. Eventually, all actors should be accountable to the communities who are the target beneficiaries and are providing data. Protocols for handling sensitive data along with redressal systems for community members in case of a privacy breach could help increase community trust.We note that privacy norms are culturally situated, and we need to be careful about whose notion of privacy is imposed (Ahmed, Haque, Guha, Rifat, & Dell, 2017).

Chapter 4

How does AI change the nature of work?

Multistakeholder Collaboration in large-scale AI deployment

4.1 Motivation

Artificial Intelligence (AI) based systems are increasingly playing a role in decision-making and resource allocation in high-stakes settings, such as healthcare, public welfare, humanitarian crises, and more. The integration of AI in these contexts is frequently targeted towards supporting the efficient use of limited human and technical resources, and enabling more accurate and/or fairer decisions by stakeholders. Researchers in the field of Computer

Supported Cooperative Work (CSCW) have been investigating how such systems might be designed appropriately (Holten Møller, Shklovski, & Hildebrandt, 2020; Ismail & Kumar, 2021; Saxena, Badillo-Urquiola, Wisniewski, & Guha, 2020), and have drawn attention to implications for historically underserved populations (Brown, Chouldechova, Putnam-Hornstein, Tobin, & Vaithianathan, 2019; Suresh et al., 2022). Prior work has uncovered how AI can impact existing workflows, influence decision-making, and shape interactions across human actors (Ammitzbøll Flügge, Hildebrandt, & Møller, 2021; Panigutti, Beretta, Giannotti, & Pedreschi, 2022; Saxena, Moon, Shehata, & Guha, 2022). This body of work also highlights risks of limited transparency, reduced accountability, and bias in AI systems, which are amplified in high-stakes settings (Brown et al., 2019; Okolo, Dell, & Vashistha, 2022; Sambasivan, Arnesen, Hutchinson, Doshi, & Prabhakaran, 2021b).

As AI technologies make their way into public sector and health infrastructures, this work aims to uncover *How does AI change the nature of work* by offering an ethnographic perspective of a large-scale AI deployment. This work was published in the Proceedings of the ACM on Human-Computer Interaction, CSCW 2023 (PACM HCI) titled 'Public Health Calls for/with AI'.

Prior research has highlighted the paucity of appropriately designed AI interventions targeting global health (Ismail & Kumar, 2021), calling for a more meaningful, human-centric integration of AI in the quest for societal impact. This work provides a situated perspective of a real-world AI intervention to address this gap; and draws attention to key design decisions made when integrating AI into a complex maternal and child health ecosystem in a historically underserved context in Mumbai, India.

We focussed on mCare, one of HealthNGO's largest programs, which delivers voice-based messages on pregnancy and child care to more than 240,000 beneficiaries (pregnant women and mothers) every year. A persistent challenge faced by mCare was the drop in engagement over the 18-month involvement with each target beneficiary, attributed to an array of factors influencing listening behaviors, including intermittent access to phones, cultural norms, health literacies, challenges experienced in the care journey, and more. To increase engagement in mCare, HealthNGO employs and trains human callers who encourage and offer counseling to beneficiaries through phone calls. Given the massive scale of this program, HealthNGO can only conduct a limited number of such calls. The introduction of the AI system in mCare was aimed at increasing overall engagement, by helping identify beneficiaries who may be most at risk of dropping out of the program and could benefit from human intervention.

This paper presented an ethnographic investigation of a multi-stakeholder real-world AI-based public health intervention and examined *How does AI change the nature of work?*. We studied AI integration in the mCare program, conducting fieldwork across multiple sites in Mumbai between July to September 2022. Our research focused on three sets of stakeholders—callers (including call center executives and hospital supervisors) at HealthNGO who were responsible for calling beneficiaries who were predicted to drop out of the program by the AI model, program and IT staff at HealthNGO shaping the design of the AI intervention and managing callers, and the development team at TechOrg building the AI model and supporting digital infrastructure.

The goal of our research was to inform public health projects that rely on data-driven approaches, by identifying critical considerations for the design of human-centered AI integration and examining the changing nature of work due to the introduction of AI systems.

This paper presented one of the first ethnographic studies of an AI system deployed large-scale in a public health context. Our analysis detailed how different stakeholders attempted to determine the *what* or the program definition of the AI intervention, before uncovering complexities around determining *who* benefits, *how* the human-AI collaboration is managed, *when* calls must take place in alignment with various other priorities, and *why* the AI is sought, for what purpose.

I draw attention to the work done by the AI (as actor) and the work of building the human-AI partnership (with multiple, diverse stakeholders).

4.2 Background

Our research took place in the context of HealthNGO's mCare public health program and TechOrg's AI intervention (anonymized). Below is an overview of both stakeholders and their involvement, critical to understanding the contributions of our research.

4.2.1 Overview of the mCare Program

mCare is a free mobile call service operated by a maternal and child health non-profit organization called HealthNGO, headquartered in Mumbai, Maharashtra (India). The target beneficiaries of the program are pregnant women and new mothers. Women are enrolled—with written consent—by a health worker during a home or hospital visit. As part of the program, they receive timed recorded voice messages every week, corresponding to their gestational age or the age of their child (till they are one year-old). The voice messages provide information on breastfeeding, sanitary practices, nutrition, child development, and more. If the woman misses the call at the scheduled time, she receives a call the following day at the same time, and a third time on the day after that if she misses again. If the woman would like to listen at a time of her choice, she can call a number to hear the voice message for that week. mCare has reached over 2.6 million women across nine states in India as of 2022, and offers content in four languages. A recent evaluation of the mCare program through a three-year randomized control trial demonstrated that the calls have had a positive impact on infant birth weight, infant feeding practices, and immunization.

HealthNGO has set up a call center in their office in Mumbai where they run several programs to support mCare, mostly focusing on beneficiaries in the state of Maharashtra. The calls are placed by women Call Center Executives (CCEs). One of their largest programs is the *37-week program*, where *service calls* are placed by CCEs around the 37th week of pregnancy to determine if and when women have delivered their baby, based on which they activate the calls on child care. Once a month, HealthNGO sends a text to inform mothers that they can *opt out* of the program over a phone call, particularly in the case of miscarriage or child death. Finally, they also have a *missed call service* that beneficiaries can call to ask questions around maternal or child health, to report a delivery and activate the calls for child care, or to report a miscarriage or child death and stop the calls. All programs are free for beneficiaries.

4.2.2 Integration of AI into the mCare Program

AI was introduced into the mCare program with the goal of *predicting beneficiaries who were likely to drop out of the program* Verma et al. (2023).

Given the massive quantities and real-time nature of data, identifying beneficiaries manually was challenging. Also, due to limited human resources, only a certain number of live calls were possible every week. In 2019, HealthNGO began a partnership with TechOrg towards developing a machine learning (ML) system that could automate the process of identifying beneficiaries to provide a follow-up call. At the start of the collaboration, HealthNGO tested alternatives to calling, including text messages, and text messages followed by calls to only those beneficiaries whose engagement had not increased. They found live calls to have the greatest impact and decided to focus on this approach despite it being the most expensive and labor-intensive option. Through a collaborative and iterative process over two years, TechOrg developed, deployed, and evaluated several ML models.

4.2.2.1 Workflows Associated with the ML Intervention

The ML intervention has been deployed in mCare for over a year, and was being used by TechOrg *to generate a list of beneficiaries for callers at HealthNGO to reach out to every week.* Once research staff at HealthNGO receive the list, they use a web application to allocate calls to various callers, who receive them on a mobile application. Beneficiaries who had been registered at government hospitals or partner NGOs were assigned to their respective hospital supervisors or NGO staff, and the remaining were randomly assigned to CCEs. We use the term *callers* to refer to both CCEs and hospital supervisors. Callers received the list of the beneficiaries on the mobile application at the beginning of every week, and had to make three attempts to reach them that week. If an attempt was not "successful" because no one picked up the call, then they called again the next day. If no one picked up after three attempts, then the call was marked as "unsuccessful" on the app. The *call success* *rate*—or the percentage of beneficiaries that picked up the call—as well as the *call outcomes* were monitored by HealthNGO.

4.3 Related Work - Human-AI Collaboration

Several industry actors have proposed guidelines for the design of human-AI interaction, targeting AI practitioners and designers Amershi et al. (2019); Fuller (2019). Researchers have studied how AI practitioners think about fairness, and considered the role of checklists and visualization tools in informing AI practice Holstein, Wortman Vaughan, Daumé, Dudik, and Wallach (2019); Madaio, Stark, Wortman Vaughan, and Wallach (2020); Sambasivan, Arnesen, Hutchinson, Doshi, and Prabhakaran (2021a); Yan, Gu, Lin, and Rzeszotarski (2020). Hohman et al. developed a design probe to understand how data scientists understand machine learning Hohman, Head, Caruana, DeLine, and Drucker (2019), and have also studied how data scientists wrangle with and iterate on datasets in machine learning Hohman, Wongsuphasawat, Kery, and Patel (2020); Sambasivan, Kapania, et al. (2021a). We draw attention to how AI development must be iterative and collaborative rather than relying on checklists, and how concepts such as fairness must be conceptualized within the context in which AI is deployed.

Several design methods and theories have also been developed to assist designers, such as for explainable AI Radensky, Downey, Lo, Popovic, and Weld (2022); Wang, Yang, Abdul, and Lim (2019). Dove et al. and Yang et al. have outlined the unique challenges in designing for machine learning-based systems and human-AI interaction Dove, Halskov, Forlizzi, and Zimmerman (2017); Yang, Steinfeld, Rosé, and Zimmerman (2020). Researchers have explored the evolving nature of Human-AI collaboration due to the introduction of AI systems into organizational workflows (e.g. Datta, Sen, and Zick (2016); Oh et al. (2018); Rastogi et al. (2020); Wang, Weisz, et al. (2019); Yin, Wortman Vaughan, and Wallach (2019); R. Zhang, McNeese, Freeman, and Musick (2021).) Oh et al. conducted a user study on a co-creation AI tool, indicating user preferences for explanation Oh et al. (2018). Suh et al. examined the process of co-creating music with generative models and discovered the unique role of AI as a social glue in enabling collaboration between musicians M. Suh, Youngblom, Terry, and Cai (2021). Gal et al. developed new workflows to improve human-AI collaboration where humans aid the machine in solving difficult tasks with high information value and the machine can generate motivational messages that highlight different aspects of collaboration Gal et al. (2022).

Several studies have begun to also investigate the role of values in Human-AI collaboration. A study on the perceptions of algorithms used by Wikipedia, uncovered the need for these to be transparent and align with community values, and allow human actors to act as the final authority C. E. Smith et al. (2020). Prior work has also examined the role of values in AI systems and datasets Jakesch, Buçinca, Amershi, and Olteanu (2022); Scheuerman, Denton, and Hanna (2021); Thakkar et al. (2022), foregrounding the interplay of social, cultural and organizational factors that impact AI systems. There have also been experimental studies on how users perceive AI trustworthiness or credibility (e.g. Bansal, Nushi, Kamar, Horvitz, and Weld (2021); Jakesch, French, Ma, Hancock, and Naaman (2019); Y. Jo, Kim, and Han (2019); Papenmeier, Kern, Hienert, Kammerer, and Seifert (2022); Passi and Jackson (2018)). Researchers have examined the tradeoffs between team performance and model accuracy and found that optimizing for model accuracy is not sufficient to improve team performance, even in high-stakes settings Bansal, Nushi, et al. (2021); Bansal, Wu, et al.

(2021). We speak to the need to think about the role that might AI play when introduced into a context, and carefully define success metrics for AI interventions, given their potential lasting effects.

4.4 Methods

Our research objective was to gain a deep understanding of the integration of AI into broader care ecologies. To this end, we conducted a multi-sited ethnographic study (Marcus, 1995) of a machine learning system deployed as part of the mCare program in Mumbai (India) over six weeks in July-September 2022. We conducted observations, interviews, and focus group discussions, as well as content analysis of materials that had been generated over the history of the project, including documentation and study protocols.

We investigated AI integration from the perspectives of those implementing and using the AI system, including callers (CCEs and hospital supervisors), program staff at HealthNGO, and the development team at TechOrg. We conducted a total of 24 interviews, 2 focus group discussions, and approximately 90 hours of observation. Our research participants included 32 stakeholders at HealthNGO and 4 at TechOrg, at multiple levels of the organization. We analyzed our transcribed audio-recorded interview data, written observation notes, and other documents through an iterative inductive coding and memo-writing process, as recommended by Charmaz (Charmaz, 2006).

4.5 Findings

4.5.1 Who to Target: Identifying Beneficiaries

Who benefits from the AI intervention was a key question for our investigation. Our findings highlight that design decisions regarding who should be targeted are not always straightforward ones to make. How marginalized target beneficiaries were (and how marginalization was assessed to begin with), how equipped they were to partake of the intervention, and how great their need was for the intervention— all played a crucial role in determining the efficacy of the ML model at play, as we present below.

4.5.1.1 Identifying the Marginalized

In the context of our study, caste and religion often linked to marginalization, impacting access to health services (Baru, Acharya, Acharya, Kumar, & Nagaraj, 2010; Hamal, Dieleman, De Brouwere, & de Cock Buning, 2020; Paul & Chouhan, 2020). Program staff expressed concerns about collecting sensitive data due to potential misuse and privacy issues. It was also preferred that the callers not have this information so that they treated each beneficiary without bias. Details visible included only name, age, date of last menstrual period, gestational age, number of call attempts made, and the outcome of the calls. The ML model used income and education as markers of marginalization instead. Supervisors asked beneficiaries for their husband's occupation, and attempted an approximation of the income accordingly. Sometimes hospital supervisors put down their own perception based on their interaction with the beneficiary, or based on other demographic details collected such as education level and the type and ownership of the phone.

4.5.1.2 Identifying Potential for Impact

AI was intended to increase engagement from beneficiaries who were perceived to be "low listeners" or less engaged in the program. However, selecting these beneficiaries proved difficult due to low call success rates, due to connectivity issues, lack of mobile credit to receive calls, the phone number having changed, or the phone being with the husband or a family member. This was more likely to be experienced by women from marginalized backgrounds who might not have their own personal phone or reliable network access. Focusing on low listeners, however, came into conflict with the experience for callers, who experienced a drop in their motivation levels if their calls were not answered. The tradeoff was addressed by the ML model excluding beneficiaries who never answered a call.

4.5.1.3 Identifying Need

HealthNGO intended to identify women who were not in need of the information being provided. For example, Fatima (Hospital Supervisor) shared, *"There are some who are more educated or this is their second or third child so they feel like they already know everything and don't need to listen."* Reshma (Hospital Supervisor) also pointed out that if the beneficiary was more educated, they could find similar information on the internet such as on YouTube, and might find videos more appealing than voice calls.

HealthNGO was also interested in determining who was no longer engaged in the program, or numbers that were no longer active or repeatedly out of coverage. The motivation was to conserve resources that were already acutely constrained (time and money). Prioritizing informed consent however, HealthNGO did not switch off calls to anyone who had not actively opted out. A major concern was around identifying false positives, in case the woman was interested in the program but struggling to access it due to systemic barriers.

On the other hand, program officers also had ethical concerns with calls that were currently going to women who had miscarried, had an abortion, or lost their child. Many beneficiaries struggled to opt out later due to limited digital literacies. Our interviews revealed that the use of AI was particularly helpful in identifying and actively reaching out to such women. The caller could then provide emotional support, and also immediately switch off calls to them with their consent to prevent potential trauma.

4.5.2 AI in the Background: Interactions between Callers and Beneficiaries

The AI intervention's success (or not) relied on interactions and sustained relationships between callers and beneficiaries. Callers can be perceived as "humans in the loop", frequently discussed in AI systems. They aimed to engage low listeners, informed beneficiaries about mCare, and adapted to enhance system efficacy.

4.5.2.1 Enrolling and Informing Beneficiaries

Our interviews with callers confirmed that their main goal was to increase engagement among low listeners, in line with the goals of the ML intervention. A necessary outcome for beneficiaries to be engaged was awareness about mCare, starting with the hospital visit when they were enrolled in the program. Though most beneficiaries were aware, several had forgotten that they had enrolled or thought they were receiving spam calls. Often, beneficiaries forgot about their enrollment, particularly due to the stressful environment of hospital registrations, or mistook calls for spam, Consequently, ensuring beneficiaries were accurately informed during enrollment was critical.

4.5.2.2 Sustaining Beneficiary Engagement

The calls were key for strengthening the relationship between the HealthNGO's callers and beneficiaries. Hospital supervisors laid a solid foundation for long-term engagement with beneficiaries by building rapport, establishing trust during visits and answering questions. We found that calls by CCEs were typically shorter; they were also assigned more calls and did not have experience with repeated/in-person interactions with beneficiaries.

The nature of relationship that a beneficiary could nurture with a CCE was quite different from that with a hospital supervisor, but the latter was not part of the ML intervention, and any improvements in engaging beneficiaries that resulted from interacting with the supervisor would therefore not be factored into assessing the efficacy of the ML system.

4.5.2.3 Hustling to Connect

Our observations revealed that a key aspect of the call was not just informing the caller about the program, but uncovering why they were not engaging and addressing this lack of engagement. We found that common challenges included phones being with husbands or containing outdated information. Callers adapted call timings or updated beneficiary details to improve engagement. In some cases, women had provided their husband's number due to privacy concerns or had since gotten their own phone. Callers updated the delivery date for women still receiving pregnancy calls post-delivery. If a woman miscarried or got pregnant again, they updated the last menstrual period date to inform program delivery. Cultural practices around traveling to their parents' home, led to changes in beneficiaries' affecting phone access and preferred timings.

4.5.3 When to Call: Aligning Across Constraints

Our findings highlighted that key program decisions needed to be made around *when* calls were placed. The response rate on calls was also an important consideration from the perspective of designing the AI system, since calls took up time—a precious resource, and it was crucial to protect against failed call attempts. Below we discuss how calling needed to be brought in alignment with the care journey, availability, workflows, and adjacent programs.

4.5.3.1 Aligning with the Care Journey

Our discussions at HealthNGO revealed that the mCare program regularly saw a linear drop in engagement in the first few months after registration, and then engagement became fairly stable. There was a second significant drop after giving birth, before engagement became stable again.

Given these complexities, one of the aspects of program design that was discussed by HealthNGO and TechOrg was the "best time" to place calls. Calls were placed within three months of program registration, to prevent the initial drop in engagement. As most women were registered in their third to fifth month of pregnancy, the calls largely went to women who were pregnant. By calling earlier in the program, HealthNGO was also able to identify miscarriages earlier and prevent potential repeated trauma.

Callers worked to align with availability while navigating challenges such as phone usage and beneficiaries' varying schedules. Hospital supervisors, with more flexible calling times, adapted better than CCEs with fixed hours. Hospital supervisors had fewer people to call, they could more easily try to vary times for greater success in successive attempts. This was harder for CCEs, possibly due to fixed hours and the volume of calls they made every week that made it difficult to keep track. We find that workflows that were ideal for beneficiaries sometimes conflicted with caller workflows.

4.5.4 Why AI: Considering Program Goals

Increasing engagement of target beneficiaries with mCare was the goal of the AI intervention, which was developed to predict callers whose engagement was likely to drop. Program Staff at HealthNGO were interested in the prediction accuracy, but an additional metric at the program level – "call success rate". The challenge with focusing on this metric, was that it depended on beneficiary behavior that seemed to be out of scope for what AI could enable. A third success metric associated with the program that was of interest to the callers was the outcome of the calls that were picked up. We learnt that decisions on altering call timings didn't significantly increase engagement, possibly due to beneficiaries' unpredictable schedules and movement patterns.

The introduction of AI had brought HealthNGO to streamline data flows. For instance, the program staff introduced the ARN program as a result of the AI integration which is a unique ID generated during the registration process and helped verify that a beneficiary's number was operational.

4.5.4.1 Evaluating Impact

Program officers noted a significant increase in overall engagement across the mCare program since the ML system had been introduced. This was due to iterative changes in the workflows, such as incorporating feedback from beneficiaries, ARN data flow, and offline follow-up interactions in hospitals, which had strengthened the mCare program overall. The model showed a 30% increase in engagement compared to those who did not receive the AI intervention. To further increase model effectiveness, TechOrg was considering incorporating call outcomes into the AI model, based on results of a randomised controlled experiment.

4.6 Discussion

I argue that the study of Human-AI collaboration across the AI pipeline is particularly useful in understanding emergence of new work practices, roles, and power dynamics with the introduction of AI systems. I observed how AI can standardise and structure workflows for NGO staff, but it can also introduce changes in their existing workflows where their care work shifted to working through an AI generated list. Similarly, the alignment of multiple stakeholders towards shared goals around equity (ML fairness for AI developers and program distribution for NGO staff) showcases the reconfiguration of Human-AI collaboration across the AI pipeline. Next, I examine the dynamic role of AI as a non-human agential actor that acts as an advocate and mediator.

4.6.1 AI as Actor

Seeing AI as an *actor* with agency allows us to interrogate the work that it was doing in impacting program goals, design and human interaction. Here we draw a connection to Actor-Network Theory (ANT), an analytical and methodological approach that views humans and nonhumans as having agency and playing an *equal* role in acting or participating in a network of relationships (Latour, 2007; Walsham, 1997). There were several interconnected networks that AI acted on in our research, including (but not limited to) the organizational structures at HealthNGO and TechOrg, the relationships across callers and beneficiaries (which varied in strength across CCEs and hospital supervisors), and to a lesser extent, the relationships between beneficiaries and their families. Here we draw attention to some ways in which AI acted in this network of networks.

AI moved from optimising limited resources at the initiation to an *advocate* due to its improved learning of beneficiary behavior by drawing attention of HealthNGO to identifying beneficiaries who might have miscarried. AI also became a *mediator,* by facilitating dialogue among stakeholders about program goals and impacts, aligning HealthNGO's workflows with its objectives while considering fairness, privacy, and program sustainability. Future work could actively focus on AI as an embedded actor with agency and how we might explicitly design for AI to enable dialogue across stakeholders.

We build on prior work (Abebe et al., 2020; M. M. Suh, Youngblom, Terry, & Cai, 2021) that proposes AI's role for social change and social glue. We draw attention to the *power* that AI held in our study context, by shaping the attention of and conversations among stakeholders. On one hand, this effect was being leveraged towards addressing equity. I caution this by questioning the ethics of leveraging AI for such roles, when model performance may be uncertain, program goals may not be straightforward, and AI could have an undue influence on shaping decisions (Cao & Huang, 2022; Kapania, Siy, Clapper, SP, & Sambasivan, 2022).

4.6.2 Configuring Human-AI Partnerships

We have thus far repeatedly discussed AI as a single entity in the context we studied. In reality, what was perceived to be an "AI intervention" was a set of human and nonhuman/technological actors working together. This included the ML model in the background developed by TechOrg, the web application used by the research staff at HealthNGO to distribute calls, the mobile application used by callers, as well as the calls made to the phones of beneficiaries, access to which had to be negotiated with family members. Each of these human and nonhuman actors was embedded in the ecosystem.

A rich body of literature has emerged on human-AI collaboration in the HCI and FAccT (Fairness, Accountability, and Transparency) communities that examine the role of humans-in-the-loop (Shneiderman, 2020) or machine-in-the-loop (Clark, Ross, Tan, Ji, & Smith, 2018; Green & Chen, 2019), among other conceptualizations of humans and AI systems working together. We argue to consider the effect of the human and the technology in varying degrees within the partnership, rather than trying to determine what is "in the loop."

Our findings, indicated that the "hustling" the callers engaged in to connect with beneficiaries was complex and challenging, and differed across callers based on their role, experience, and expertise. To further strengthen AI's efficacy as advocate, the setup could be reconfigured for AI to have more influence on callers' decision-making, thereby supporting and easing their workflows. A perspective on configuration also allows us to creatively combine human and technological capacities.

Chapter 5

How do people perceive and experience AI?

5.1 Background and Motivation

Accelerated progress in computing power and artificial intelligence (AI) raises compelling, polarising questions about the impact of automation on jobs, skills, and wages. Job losses from automation are believed to be inevitable in many technology and policy circles (*Harnessing automation for a future that works | McKinsey & Company*, 2017; James, 2017; policy brief, 2017). For example, in firm conviction that a jobless future is imminent, technologists such as Elon Musk and Sam Altman have publicly advocated for social safety nets through Universal Basic Income (UBI) programmes (Chris, June 2017).

Labour markets in the Global South are especially susceptible to trends in automation. Entire industries built around rule-based jobs like call centres, technology outsourcing, and low-level factory jobs could face the risk of job destabilisation from automation (*e.g.* (Economist, 2018; Frey & Osborne, 2017; Goos & Manning, 2007; Manyika et al., 2017)), exaggerrated with permeation of Generative AI technologies (McKinsey, n.d.).

Despite the growing importance of implications of the Future of Work, research remains predominantly shaped on western populations (*e.g.* (Center, 2017; Goos & Manning, 2007; Manyika & Sneader, 2018)), with limited understanding of anticipated futures in other infrastructural, social, and economic realities. Understanding perceptions of promise and risk of automation and future-readiness among rule-based workers in the Global South is crucial in designing effective Human-AI collaboration experiences, that provide pathways to equitable futures. I argue effective human-AI collaboration can be designed by tracing the experiences of both – people building and using these systems, across the AI development pipeline. This examination provides an opportunity to understand the experiences of historically underserved communities by embedding their socio-cultural context and eventually designing AI experiences that are context-sensitive and human-centered.

Through this work, my collaborators and I, examine perceptions and practices around AI-powered futures among *vocational technicians*— a socioeconomically disadvantaged yet large community of rule-based workers in India and other emerging economies, projected to be susceptible to computerisation (Eichhorst, Rodríguez-Planas, Schmidl, & Zimmermann, 2012; Johanson & Adams, 2004). This work that seeks to examine *How do people perceive and experience AI?* was accepted at CHI 2020, titled 'Towards an AI-powered Future that Works for Vocational Workers'.

Vocational technicians constitute a critical labour force that transitions from high-school qualification to skilled technicians, specialising in fields like data entry operations, electrical wiring, and welding (of Trade, 2017). An estimated 1.5 million students are enrolled in over 15,000 vocational training institutes in India today (Ministry of Human Resource Development & Literacy, 2016).

I report findings from conducting participatory action research with 38 vocational technician students of low socio-economic status in Bangalore, India. We found that vocational technicians, socially and economically disadvantaged, viewed technical expertise as a powerful vehicle to break from boundaries of caste and class. Participants perceived automation as having limited impact on future livelihoods, expressing that their hard-earned vocational expertise was irreplaceable. Contrary to the dominant discourse on the future of work (e.g (Center, 2017; Hiringlab.org, 2019)), our participants were unfamiliar with the growing role of AI in rule-based tasks. Upon learning about automation, our participants then articulated an emic vision for the future of work, seeking legal protections, unionising, workplace collaboration, and accountability from employers and technology makers. However, as active users of technological platforms powered by AI-based recommendation systems, for skill-building and job-seeking, our participants were deeply excluded from gainful employment by the content and recommendations available to them. Learning platforms were reported to not recognise or validate informal sector skilling.

As up-skilling and re-skilling become paramount to employability, unmitigated algorithmic inequity further limits future preparedness of technicians. Based on these results, I present a manifesto for technical, policy, and ethical directions, moving towards an equitable future of work by designing effective Human-AI collaboration.

5.2 Related Work

5.2.1 Public Perceptions of Automation and AI

Public perceptions of AI in emerging challenges such as social justice, climate change, and other threats have been studied among western audiences (Cave et al., 2018; B. Zhang & Dafoe, 2019). Cultural factors shape how AI is portrayed in media, culture, and everyday discussion are previously studied in context of how it influences what societies find concerning or exciting about technological developments (B. Zhang & Dafoe, 2019). Media discourse around perceptions of General Artificial Intelligence highlights the perceived risks of autonomy provided to such technologies (Friend, 2018; *Understanding the public perception of AI*, 2019).

Prior work by Brookings has examined the perceptions of AI through the lens of Optimism, Impact on Humanity and Jobs, and Government regulation through a quantitative examination (West, 2018). Interestingly, while a majority of the respondents were sceptical about the impact of AI, most survey questions had over 30 per cent respondents who did not have a well formed opinion on AI. Quantitative studies focused on perceptions of job loss through automation focus on themes such as AI freeing up individuals from mundane work and show understanding among respondents about the effects of automation on blue-collar work and uncertainly of the effect on white-collar work, such as research by Smith . (e.g. (A. Smith & Anderson, 2014)). Our work enriches this growing body of research on AI perceptions by expanding the focus to non-western audiences and understanding the perceptions of a population vulnerable to trends in automation.

Prior research in HCI does not appear to examine how workers, in their own right, in the Global South perceive their futures of work through AI. Our research presented an empirical study on perceptions around automation by a worker community that is prone to risks of job instability and disruption. Through participatory engagement with the community, we identify a roadmap for globally equitable future of work designs and policies.

5.2.2 Algorithmic Fairness among Marginalized Groups

Research on algorithmic fairness has used various lenses, critiques and observations around the role of algorithms in reinforcing or widening biases. Ajunwa (Ajunwa, Friedler, Scheidegger, & Venkatasubramanian, 2016) study the disparate impact in hiring of deploying AI algorithms on protected classes. Chander (Chander, 2016) argue for algorithmic design and assessment in a race and gender conscious manner instead of gender and race neutrality/blindness. Lum (Lum & Isaac, 2016) find that predictive policing of drug crimes results in increasingly disproportionate policing of historically over-policed communities. Similarly, bias in image and textual search is discussed in prior work (*e.g.,* Otterbacher (Otterbacher, Bates, & Clough, 2017) and Buolamwini and Gebru (Buolamwini & Gebru, 2018)). Most relevant to us, algorithms are increasingly studied as vehicles of influence for social change on the society. Beer . argue that the power structures in algorithms need to be critically examined and understood given their role in everyday lives of people through social media platforms (Beer, 2009). Recent work in algorithmic fairness has started to expand the conceptualization of fairness beyond economistic calculations to sociotechnical framings, notably by Selbst (Selbst, Boyd, Friedler, Venkatasubramanian, & Vertesi, 2019)–we borrow from these broader framings of fairness in systems.

Recent research has also focused on examining perceptions of fairness through participatory action research. Lee . (Lee & Baykal, 2017) study the need for fairness through inculcating social behaviours that cannot be expressed in mathematical terms. Woodruff . (Woodruff, Fox, Rousso-Schindler, & Warshaw, 2018) understand the perceptions of (un)fairness among marginalised populations in the U.S., noting that while (un)fairness is experienced through vectors like racial prejudices and economic inequality, and negatively impacts user trust in systems. We extend the body of inquiry to examine the perceptions of algorithmic fairness of technical systems that potentially help uplift socio-economic opportunity for marginalised populations.

5.3 Method

Our study derived insights from 38 semi-structured interviews conducted in Bangalore, India, from June 2018 to May 2019, with students at vocational training institutes. We aimed to understand their aspirations and perceptions regarding vocational training and future work trends.

I recruited participants through snowball and purposive sampling (Biernacki & Waldorf, 1981; Seidman, 2006) from three institutes, ensuring a balance across various engineering trades. The study included 32 male and 6 female students, aged 18-24, mostly from low socioeconomic backgrounds (of India, 2011) and generally first in their families to attend college. We drew from the method proposed by Woodruff et.al. ((Woodruff et al., 2018)) to inquire on the perceptions of automation through a participatory action research by introducing automation trends from non-related fields to discuss potential impacts and gather perceptions. I carried out a structured, iterative qualitative data analysis (Thomas, 2006) to summarise and interpret the interview data by identified larger themes emerging from iterations.

5.4 Findings

We describe the aspirations of our participants in pursuing vocational programs, the digital platforms they engage with in these pursuits, and their perceptions of automated futures .

5.4.1 The Technological Promise of Vocational Training

Our participants pursued vocational training primarily for technological expertise, viewing vocational training as a vehicle for economic and social upward mobility. Most participants had migrated from rural areas in a quest for improved livelihoods and stable futures, particularly to move past struggles around finances, caste, domestic violence, and alcoholism among fathers.

5.4.1.1 Breaking from Caste Discrimination through Technology

Participants recognised the importance of the core technical skills from vocational training in breaking caste barriers. Most participants came from lower caste backgrounds, whose parents and relatives were employed in lower social status professions, like daily wage labourers. Our participants reported being at the receiving end of hierarchical treatment because of their caste (which determines their social standing).

Participant's aspirations to pursue vocational education over jobs in retail or commerce stem from the perceived social standing of technical skills. A technical background is associated with employability as well as a way to earn social capital including finding better matches in arranged marriage and boosting marriage resumes.

5.4.1.2 Economic Mobility through Accelerated Employability

Vocational training was perceived by participants as a path to a "good future". Participants felt assured that such training would provide them with a job, a sentiment they reported to be echoed by their instructors, social networks, and news articles (also found in prior work (Maithreyi, Padmanabhan, Menon, & Jha, n.d.; Okada, 2012)). The draw of computing for upward economic mobility in India has also been widely discussed in prior work ((Pal, 2008; Thakkar et al., 2018)). Participants perceived technical expertise as providing job security through managerial and leadership promotions.

Vocational training was seen as an accelerated path to employability. Instead of having to go through years of high school and university to find a job, participants saw vocational training as a quicker route to success, since the programs were 1-2 years long and highly application-oriented. Many participants felt that this training would not only ensure job readiness, but also help them secure (7x-8x) higher salaries. As these institutes required a minimum passing score from tenth/twelfth grade exams, they seemed within reach for participants who lacked social capital, or had a modest academic performance, in contrast to college degrees.

5.4.2 Local Visions of the Future of Work

5.4.2.1 Perceptions of Automation

We asked participants to reflect on the changing nature of work, given the onset of increased automation and growing popularity of the gig economy. Our participants disassociated computers from manufacturing and production jobs. Participants did not ascribe intelligence or cognitive ability with machines, using terms like ‘control’, ‘operate’, and ‘use’ in the context of machines. Participants had ’object-oriented’ relationships with machines and did not view them as collaborators, co-workers, or supervisors. Rather, machines and computing were viewed as efficiency tools that were controlled by operators. Participants perceived that technically complex tasks could only be performed by humans, not automation. In a socio-cultural context where technical careers were linked with prestige and respectable social status, participants harboured a belief that their hard-earned technical skills could not be automated.

Participants articulated the role of technicians as executors of perfection; performing polished, complex technical tasks. According to them, technicians brought leadership and creativity skills in addition to technical skills. Importantly, technicians were seen as having expertise and skill in correcting errors (in case of change in business or technical requirements on the job), whereas machines were seen as executors of repetitive tasks with no scope for correction without human input.

> *“If something goes wrong at work, my colleague will teach me how to do it. So will I for him. We learn from each other—how will we learn from a machine if*

> *it does my job?"—M, 21, Electrician*

In order to elicit reactions to increased computerisation, participants were introduced to past, present, and future trends in automation and professional domains where jobs were transformed. All participants were confident that their jobs were not under threat. While they understood what automation entailed, participants felt that the hard skills they had acquired could not be replicated by a machine.

5.4.2.2 Workplace Relationships and Bargaining

Participants viewed work as essential to their identity and social standing. We observed a deep sense of belonging to the workplace due to a loyal relationship with employers, linking personal growth with employer's.

> *"I would not ask for overtime generally because my boss will tell me to do extra only if it is really needed. If I work extra now, they will also help me when I need to take an off or need some money if something happens to me or my family members."—M, 20, Mechanic*

Coworkers were perceived as integral support systems and family away from home, since most participants migrated to urban areas in pursuit of better education and job opportunities. In general, participants envisioned an ideal workplace as one where they had a sense of belonging, good peer and mentorship networks, career growth opportunities, and adequate salaries.

5.4.2.3 Outlining a Future of Work that Works

Participants envisioned a future of work emphasizing legal protections and collective negotiation through unions, especially against automation-related job displacement. They stressed longer notice periods for job loss and employer-sponsored re-skilling. They advocated for worker participation in decision-making and accountability from upper management, highlighting limited access for lower-level workers in India's hierarchical society.

Participants expressed strong resistance to the concept of Universal Basic Income (UBI). As work was viewed as an integral part of their identities as well as a place for social interactions, participants viewed a life without work and free income as "going wrong", "going mad", and "wasting their life".

> *"I need to work—any form of work. If I do not and sit at my home even if I get free money, I will go mad if I do not do work and meet my colleagues."—M, 22, Mechanic*

5.4.3 Algorithmic Fairness in Education, Skilling, and Jobs

We now turn to results on present day usage of skilling and job search platforms by our participants, to understand how current experiences with online platforms and tools may enable preparedness for future skills. Participants frequently turned to online platforms for learning, complementing classroom training and enhancing practical skills. As active users of the Internet, all participants owned smartphones and routinely queried search engines, video repositories, and social networking sites to advance their careers.

5.4.3.1 Limited Discoverability of Contextually-Relevant Content

Participants discussed turning to the Internet for various learning intents, such as interview preparation, visualising jobs through videos, augmenting classroom learning and job readiness, project-based learning, learning English, and discovering past exam questions.

Although participants voiced a clear need for accessing content related to practical learning, they also frequently encountered challenges with the content served to them. One key challenge was that of language; participants were taught in Kannada and familiar with vernacular terms, but the content online was typically in English (or formal, unrelatable Kannada). Even when participants knew English, they struggled with instructional videos on account of cultural disconnects, such as the accents of instructors or their style of content delivery. They were unsuccessful in identifying videos featuring instructors that they could relate to.

Additionally, search terms—even if they were in English—often carried a different meaning online versus in participants' contexts. For example, participants reported using terms such as "welding job" or "grinding job" where "job" referred to an activity or a task as commonly used in the classroom, rather than an employment opportunity.

> *"If I look for 'grinding machine how to use' videos, first I will see video of mixer-grinder (blender) which is used in the kitchen and not related to my syllabus."*
> *—M, 20, Grinder*

5.4.3.2 Semi-Skilled Learning is Difficult on MOOCs

Participants reported that certificates were a key component of their interviews, and needed to be available in physical form (and preferably laminated). Physicality of certificates added another layer of credibility.

Most participants were unaware of online certification or MOOCs. After we explained the intent and operations of online certification to gauge the interest of participants, many showed interest in learning how these could improve their employment prospects. However, currently MOOCs only cater to college education or K-12 learning, and there are no platforms that cater to vocational training students that would provide validation of their learning.The three participants who knew about online certificates previously mentioned that instances of fraud experienced by relatives had made these participants more cautious.

5.4.3.3 Job Platforms Exclude Vocational Skills

Participants mentioned the various avenues they used for job searches, including internal social networks, online search, and institute placements. Government jobs were seen as being more reliable, stable, lucrative, and prestigious. However, they were limited in number, and required job-seekers to complete an additional round of competitive examinations. Those in the CS trade were more keen on private jobs and frequently turned to various online avenues in their search. Participants used portals such as Naukri and Govtjobsportal (*govtjobsportal*, 2019), in addition to directly searching for jobs on Google Search.

We observed high reliance on personal social networks for finding government jobs among non-CS participants (which can be limited, (Dillahunt, Lam, Lu, & Wheeler, 2018)). Many participants spoke about drawing on their familial relations for help with finding a job. Participants reported using WhatsApp groups for dissemination of information regarding upcoming institute placements and other job openings discovered on the Internet.

Many participants reported not finding relevant jobs for their skill-set online, and that the job postings they came across catered primarily to students with college degrees. Participants detailed instances where portals asked for a number of details that did not apply to them, such as high school and college grades, which confused them and resulted in "incomplete profiles", leading to a common perception that these portals do not cater to vocational training students. Our participants frequently looked for jobs that matched their skills/profession such as "grinder" or "grinding" on these filters and ended up not finding relevant skill-sets on the lists of postings.

5.5 Discussion

Analysis with central question

I observe the dynamic nature of the Human-AI collaboration where the interplay of human perceptions and experiences of AI capabilities shapes the adoption of technologies. The adoption (or not) of AI technologies, represents a dynamic inter-linkage. The downstream overplay of organisational practices (legal protections, opportunity to reskill) and socio-cultural norms (social capital acquired through work, workplace relationships) is constantly evolving and negotiated between multiple actors – human and nonhuman. Through this work, I argue that it's critical to examine perceptions and experiences across the AI pipeline to further effective Human-AI collaboration. I bring special attention to the needs of vocational workers in global south, vulnerable to trends in automation. This inquiry nudges for future work that encompasses participatory methods to design technologies that proactively respond to the needs of vulnerable populations.

Design and Policy Recommendations

Our findings revealed that vocational technicians perceived their futures to be unperturbed by automation, expressing great optimism over their human technical expertise. After learning about the growing power of automation, our participants outlined their visions of a future of work. As socio-economically disadvantaged workers, technicians reported being discriminated by current platforms for skilling and job-seeking—a bias gap that could worsen with time, if unchecked. Key questions that we might ask within the HCI community are how do we ensure that the future of work is less socially inequitable? How do we challenge our ways of thinking about algorithmic fairness? How do we design for fostering aspirations through skilling and employment?

The economic mobility of an individual is heavily shaped by a stratified societal structure, (Hnatkovska, Lahiri, & Paul, 2013). The rising complexity of machines threatens to reverse the perception that machines occupy a position lower than that of those from lower castes, impacting the social standing of the latter. It is crucial for technologists and policy makers to ensure that generations of social capital and social prestige accrued by those from lower castes in securing socially respected jobs is not impacted by automation.

Participants discussed the need for transparency and accountability while anticipating decisions concerning replacement of labour with automation. It is imperative to build an understanding of automation for workers to generate informed views, avoiding speculation and panic. Unless widespread concern on robots and computer replacing humans jobs by 72% of Americans in a 2017 Pew survey (Center, 2017), our participants displayed a lack of awareness about the impending impacts of automation on their livelihoods.

The gendered implications of automation on the workforce cannot be overlooked. In contexts like India, where female workforce participation is already low (Bank, April 2019), automation could disproportionately impact women in automation-prone sectors such as content moderation and data entry that witness higher female participation.

Technology design can play a role in how aspirations are shaped, supported, and cultivated over time (Kumar et al., 2019; Toyama, 2018). Our findings also highlight the need for inclusive machine learning datasets, including vernacular content, to support diverse occupations and user behavior. This is particularly crucial given the barriers faced by technicians in accessing online resources such as learning and job search platforms due to language barriers and unique technology usage patterns. Recommendation systems can also be designed to surface content that creates learning trajectories for vocational technicians; this could address the problem of not having a clear learning path for cultivating and fulfilling aspirations.

Chapter 6

Summary and Discussion

6.1 Key Contributions

I make several key contributions to the field of Human-Computer Interaction (HCI) and in the adjacent fields of Responsible AI and AI for Social Good. The core argument of my thesis advocates for a shift in Human-AI collaboration, transitioning from examining individual interaction dynamics to viewing it as a part of a broader, interconnected pipeline of AI production. Through this work, I have not just situated my research in the global south but also attempted to change the kind of research questions the HCI community has been exploring by bringing special attention to the embedding of socio-cultural and organisational factors into the design and development of AI systems. Next, I note the specific contributions from each chapter, within the overarching question – *how do humans and AI collaborate across the AI pipeline from design to deployment?*.

In Chapter 2, I argue that the first stage of this pipeline begins by asking *AI created by* *whom?* – examining representation amongst technology creators. This stage forms the foundation of the pipeline, and this question enables us to reflect on the diversity of views and biases reflected in the creators of technologies. The emergence of AI technologies in various facets of our lives, especially in high-stakes domains in historically marginalised communities, necessitates reflection on the socio-cultural differences between the users and creators of these technologies. I shift the focus from a myopic view of technological fixes to the fundamental question at the very start of the pipeline by reflecting on the representation of the creators of these systems (Dame Vivian Hunt & Prince, 2015).

Through qualitative interviews with women across life-stages, I examined the first stage of the AI (and broadly technology) development pipeline. I discussed how cultural and economic norms around working in the software industry forged supportive environments for women at the undergraduate and Master's levels, but swiftly changed into stagnation, discrimination, and eventual exit from Ph.D., academia, and IT jobs (after entry level). This inquiry has provided rich insights into the role of culturally relatable role models for family and directed interventions at university and organisational levels. By bringing attention to the societal influence on women's education and careers, I aim to motivate future work that moves towards a fair and equitable computing industry for all.

In Chapter 3, I take a step back from the significant focus on model development to reflect on the second stage of the AI development pipeline - *how does data powering AI come to be?*. Sambasivan et al. (Sambasivan, Kapania, et al., 2021b) brought focus to the undervaluation of data work and the heightened focus on model development, leading to downstream effects on applications due to the presence of *data cascades* – compounding events that cause negative downstream issues in AI applications. My work takes this further by fundamentally examining the pipeline of data development through its creation, curation and usage across multiple stages. This work aims to address ongoing tensions across the AI development pipeline surrounding data quality by examining whether the taxonomy of 'good' data is consistent across the data supply chain.

I examined how data powering AI systems comes to be, and the valuation driving data work at various stages of the data supply chain. Taking the case of datafication of public health in India, I examined the movement of data through various stages, where the data workers included frontline health workers, data stewards, and ML developers. Through interviews with stakeholders across the data supply, I draw attention to the tensions in perception of 'good' data across its collection, creation, and usage. This work provides a new perspective in understanding data quality challenges in AI by understanding it from a lens of valuation of 'good' data and the work done by stakeholders to attain the 'goodness'. I offered recommendations for how data supply chains could be designed to bring transparency and accountability in the creation and use of data for ML development.

In Chapter 4, I explore the next stage of the Human-AI collaboration pipeline, examining how AI changes the nature of work, once introduced into organisational workflows across multiple levels of stakeholders. There are several speculations about the changing nature of work due to the introduction of AI agents – this work fills a critical gap by infusing a multi-stakeholder perspective into assessing the impact of AI as an actor with agency that influences, and in some cases aligns, goals across stakeholders. This line of work is a contemporary take on the existing discourse around AI and Future of Work that is seeking to explore the reconfiguration of the workplace due to AI. This research is also a fresh take separating the hype from on-ground realities of working with and around AI to advance organisational goals.

I provided an ethnographic study of a large-scale real-world deployment of an AI system for resource allocation in a call-based maternal and child health information delivery program in India. This is one of the first ethnographic accounts of a system at this scale. We began by presenting the *what* or the program definition of the AI intervention, before uncovering complexities around determining *who* benefits, *how* the Human-AI collaboration is managed, *when* intervention must take place in alignment with various priorities, and *why* the AI is sought, for what purpose. This work draws attention to the changing nature of work amongst multi-stakeholder teams due to the introduction of the AI system – signifying the role of studying Human-AI interactions and collaboration across multiple levels.

In Chapter 5, I examine the fourth stage of the pipeline and aimed to understand how people perceive and experience AI systems. Public narratives are known to shape the adoption of technologies and have been examined through longitudinal studies (Kelley et al., 2021). This is extremely timely for addressing the polarising narratives around job loss to increased ability to perform complex work. However, in the quest for societally responsible AI, it is critical to capture perspectives through a nuanced community-level understanding, especially for those communities that are historically underserved and vulnerable to impacts of AI (Okolo et al., 2021; Woodruff et al., 2023).

I presented our findings from participatory action research conducted with 38 vocational technicians in Bangalore, India, where we examined the perceptions and practices of vocational technicians who are projected to be at risk of job displacement in an AI-powered future. We described the aspirations of a socio-economically disadvantaged population who view gaining technical skills as a way to break the boundaries of caste and class. We analysed the perceptions of automation among vocational technicians, and outlined a ground-up vision for a future of work as described by our participants. We found that skilling and employment platforms display algorithmic bias against our participants who rely on these systems as a way to up-skill and re-skill in an AI-powered future. Based on our results, we identified implications for design, social policy, and ethical principles for an equitable future of work that meaningfully advances technicians.

Next, I provide a post-hoc reflective analysis of my findings along with outlining limitations of my work and provide directions for future work.

6.2 Analysis

This thesis presents an opportunity to synthesize findings from my prior research, focusing on the overarching question *How do humans and AI collaborate across the AI pipeline, from design to deployment?*. To address this, I segmented the central research question into four critical stages (RQ1-RQ4) of the AI development pipeline. While these questions are not exhaustive, they establish a foundation for a comprehensive examination of Human-AI collaboration.

I attempt to provide a perspective that shifts focus from individual components to examining the central question by analysing the interaction of human and non-human actors across the AI pipeline and reflecting on the interplay of users and creators across the pipeline. Next, I reflect on the configuration and reconfiguration of Human-AI collaboration across the AI pipeline in the Global South.

6.2.1 Interplay of Users and Creators

The HCI community increasingly acknowledges that Machine Learning algorithms often fail in real-world applications or face limited acceptability due to a lack of engagement with key stakeholders during the algorithmic design process. This oversight can lead to a disconnection from the tacit knowledge and insights of users and stakeholders, resulting in biases and ethical concerns in AI design. This thesis is a humble attempt to reflect on the interplay of creators and users across the development and deployment of AI systems.

HCI researchers have examined co-creation of technology through various analytical frameworks and methods such as value sensitive design (VSD) and participatory research. Notably, Zhu et. al. proposed Value-Sensitive Algorithm Design ((Zhu, Yu, Halfaker, & Terveen, 2018)) as a method that emphasizes integrating human insights to guide the creation of automated algorithms, which not only enhances the acceptability but also reduces potential biases in design choices. These frameworks are critical to capture values and context of users, and recognise traditional power dynamics between users and researchers. Building on these frameworks, I argue that creators and users play dynamic roles across the AI pipeline, significantly influencing the final creation, use, and perception of AI.

In Chapter 2, I noted the social, cultural and organisational factors that motivate women into computing and direct eventual exit of women from computing. The limited role of women, and by extension historically marginalised groups, as creators in technology has profound implications on their experiences as users. I highlight, for example, strategies women adopt for anonymity on online learning platforms to evade unsolicited advances. Similarly, in Chapter 3, I focused on a critical resource powering AI advances – *data*. I noted the interplay of AI developers as both creators of downstream applications but also users of data collected and curated by community health workers (data collectors) and data stewards. Similarly, data collectors are subject to data quality monitoring (Karunasena et al., 2021) by AI algorithms. AI system use data (provided by them as creators), transitioning data collectors to users of AI systems that are used to motivate them towards improved data creation practices. The interplay of various stakeholders as both creators and users is expounded by the tensions I noted across the value chain in the production of 'good' data. I argue towards a broader reflection of values embedded in the AI pipeline, extending beyond mere model-centric considerations to encompass data creation and usage. The interplay of usage and creation is further prominent in Chapter 4, where I showcase the role of healthcare workers in the background that determines the success of an AI system through their thoughtful engagement with beneficiaries of the mhealth program. Here, healthcare workers are part-creators of the AI system and part-users of the AI system as they receive AI predictions governing their care work.

My thesis makes an attempt to de-center the current paradigm in Human-AI collaboration that examines creators and users of technologies in isolation. I invite the research community to reflect on the AI development and deployment pipeline by examining and accounting the interplay of users and creators. The question of how effectively (or not) Humans and AI collaborate could be examined by reflecting who creates AI and data, what values do they embed onto these systems, who do they create these technologies for and who eventually uses it.

6.2.2 Network of Human and Nonhuman Actors

In this thesis, I aim to provide a nuanced understanding of Human-AI collaboration, tracing the pipeline of AI development to deployment and situating it within the social, cultural and organisational ecologies of the humans involved in the pipeline. This approach not only offers insights into the eventual use of AI but also highlights the profound impact of human and non-human elements in shaping AI technologies. I argue that the study of Human-AI collaboration necessitates a recognition of human and non-human actors, across the AI development pipeline. This perspective encourages a shift beyond the individual actors to consider how they each have agency to mutually influence, define and shape each other.

I underscore the role of the organisation as a pivotal actor that significantly influences Human-AI collaboration across the AI development pipeline. Organisations act as critical nodes that connect various human and non-human actors, influencing not only the technical development of data and AI, but also the socio-cultural context within which these technologies operate and evolve. I engage with 'duality of technology' (Orlikowski, 1992) in organizations which suggests that technology is both shaped by and shapes organizational structures and processes. In the context of this thesis, it means that an organization is not just a backdrop for AI development and deployment but an active actor that influences the creators of AI and is influenced by AI systems.

In Chapter 2, to examine the fundamental question *who are the creators behind computing?*, I espoused the role of various human and non-human actors like familial structures, cultural norms and organisational dynamics. These non-human actors had varying degrees of agency in shaping women's participation in computing. For instance, familial structures were highly conducive to encouraging participation of women at the undergraduate and masters level. However, socio-cultural norms around marriage and childcare exert greater agency and eventually lead to drop-off of women from research and academic careers. This exploration only scratches the surface in examining the diversity of participation of creators of technologies. It is my hope that it inspires deeper reflection into examining the efficacy of Human-AI collaboration for diverse users through an examination of how non-human actors shape participation of diverse creators, especially those that have been historically marginalised.

In Chapter 3, I examined *how does data powering AI come to be?*. I noted the socio-cultural norms around caste and class that governed data collectors' access to diverse data collection sites. On one hand, data collectors navigated influence of non-human actors such as their social status and personal safety to determine what sites they could access for their data work. On the other hand, AI developers frequently struggled with even data distribution across communities which impacted downstream application algorithmic design. Frequently, these issues are attributed as 'lazy work' on part of data collectors. I argue that viewing the agential role of non-human actors across the AI development pipeline, including 'data work', is crucial to uncover what data gets collected, how, by whom and in what context. This examination is important to examine the success or failure of Human-AI collaboration and shift from a model-focused narrative to centering the influence of human and non-human actors across the AI development pipeline.

In Chapter 4, I examined *how does AI change the nature of work* through the study of a large-scale AI deployment in a public health setting. I noted the role of AI as an agential actor. The role of AI as an actor is dynamic and evolves through the lifecycle of its deployment. First, I noted the evolution of AI from being deployed for optimising limited resources to its active role as an advocate facilitating conversation and alignment amongst AI developers and NGO staff to develop shared taxonomies around what it means for the program and by extension the AI to be equitable. AI also acted as a mediator by streamlining workflows and shining focus on process changes (such as the introduction of new data collection practices) that would improve the efficacy of the program.

In Chapter 5, I examined *how do people perceive and experience AI?*. This is a crucial part of the AI development and deployment pipeline as it views downstream Human-AI collaboration as a part of a broader network of human and non-human actors. I found how vocational technicians perceive the role of workplace, their colleagues and employer in a symbiotic way where they try to accrue social capital through technical training. They view technical work as a means to break social barriers and form communities of support at the workplace. I examined the impact of non-human actors such as educational and job platforms, which vocational technicians leaned on to promote their social standing, were insufficient in meeting their needs. This examination further exemplifies the role of considering the agency of non-human actors across the AI pipeline in shaping the experiences and perceptions of users in the examination of Human-AI collaboration.

I offer further explorations on the study of agential non-human actors that require deeper examination in future work. For instance, I initiate inquiry into socio-cultural and organisational factors as determining valuation of work. This raises pivotal questions about value representation: Whose values are being captured in this process, and whose are overlooked or marginalized? How can we effectively elicit and incorporate a diverse range of values across different stages of the AI pipeline and in varied contexts? These questions are essential for a more comprehensive and inclusive understanding of the dynamics at play in Human-AI collaboration.

6.2.3 Configuration and Reconfiguration

I focus on the configuration and reconfiguration of Human-AI collaboration to underscore the importance of looking across the AI development pipeline. I draw inspiration from Suchman's work that challenges the assumptions about human actions and explores how agencies can be imaginatively and materially reconfigured at the human-machine interface (Suchman, 2007). I draw on this argument to discuss the configuration and reconfiguration of human workflows across the AI pipeline. I refer to configuration as the existing state of workflows, before the introduction of an AI system. These workflows hold value for the human actors in the AI pipeline, independent of the AI system. In Chapter 3, I discussed the intrinsic value of 'good' data, underscoring its role beyond being an input for AI systems. It also served as a performance indicator for healthcare workers and a reporting tool for data stewards. These configurations are dynamic, yet independent of the AI system.

However, the integration of AI systems also brings about reconfiguration across the AI pipeline, altering human actors' workflows and the interaction between human and non-human actors. For example, in Chapter 4, I noted the introduction of new data fields during data collection and streamlining of data infrastructure by NGO staff due to the introduction of the AI system. These reconfigurations, while enhancing existing workflows, also hinge on the pre-existing configurations, such as the relationships healthcare workers have established with their beneficiaries, often formed at hospitals during registration.

The process of reconfiguration is dynamic and ongoing, requiring adaptability from human actors. As AI systems evolve and learn, the workflows and practices surrounding them also evolve. This dynamism can lead to both improvements and challenges in established configurations. The dynamic nature of the reconfiguration can be observed in Chapter 4, where the introduction of an AI-generated list freed up time for the NGO staff, which they used to engage in deeper questions around program equity and redesigning workflows to accommodate these new configurations.

I recognize that the exploration of the configuration and reconfiguration during Human-AI collaboration only begins to scratch the surface, suggesting a need for more extensive inquiry across various domains. A broader exploration is necessary to develop a general framework that can effectively analyze these concepts across different contexts.

6.2.4 Diverse Context – Global South

The common arc across all the chapters has been a focus on uncovering social, cultural and organisational factors that are unique to the Indian context. My work builds upon the efforts of several scholars who have tirelessly pivoted the focus of HCI to look beyond Euro-American contexts and embrace the diverse and rich landscapes of Global South. The term 'Global South' in my work refers to countries primarily located in Africa, Latin America, and Asia, often characterized by their emerging economies and complex socio-political histories.

The risks associated with AI – such as perpetuation of biases, invasion of privacy, and exacerbation inequalities – are often more pronounced in settings with limited resources and less regulatory oversight. I do not presume to represent the entirety of the Global South. Through this thesis, that reports from rich qualitative data in India, I intend to shine light on the intricacies of Human-AI collaboration across the AI development pipeline. Additionally, I have focused my work on high-stakes contexts where AI technologies are being increasingly deployed, such as public health.

A focus on diverse contexts such as global south, empowers the HCI research community to equitably design technologies, especially those that seek to enable effective Human-AI collaboration, by situating them in the socio-cultural and organisational context of their users. In Chapters 3 and 4, I discussed the critical role of community healthcare workers in enabling the collection of datasets and eventual success of AI systems by leveraging their community relationships. Given the broad interest in integrating AI technologies for public health, it is imperative for HCI researchers to consider the role of community healthcare workers as responsible community leaders and design partners, beyond data collectors or agents to permeate AI interventions. My work provides an initial inquiry into this space through an examination of India. Similarly, community healthcare workers play a vital role across many countries in Africa, Latin America where such an examination would be useful.

In Chapter 5, I uncovered how vocational technicians sought to gain social capital through work in technical fields and how they perceived machine to have limited agency. They viewed the agency of machines through their lens of social hierarchy where they had been historically underserved and hence viewed technical training as irreplaceable by AI. It is important for HCI researchers to consider designing for aspirations of vocational workers.

As I approach this work, I aim to be mindful of engaging with my work in a way that connects with, rather than distances from, the communities it aims to impact. Taylor refers to this phenomenon in his seminal paper "Out there" (Taylor, 2011) by punctuating HCI researchers working in ICTD in being thoughtful before pointing their analytical lenses to observe difference and complexity, while superimposing their own ways of ordering to the world. My work strives to embody this ethos, considering the unique challenges and opportunities within the Global South, and contributing to a more nuanced and empathetic understanding of Human-AI collaboration in these contexts.

6.3 Limitations

I recognise limitations in my work across the four key research questions. My research is qualitative in nature, which I believe is a key strength of this thesis to examine the complexity of the Human-AI collaboration pipeline. However, these studies could be subject to common limitations of qualitative studies, including recall bias, observer bias, participant self-censorship, and limited generalizability of the results.

RQ1: Who are the creators behind computing?

The focus on women's representation in computing serves as an initial step in studying the lived experiences and barriers that hinder the sustainable growth of careers in computing. However, there are several groups that have been historically marginalised and have not been represented in the design of technologies that impact their lives. Future work could explore the intersectionality of these various markers, as is relevant in a region's socio-cultural context, and make a cohesive effort to study and address the barriers to retaining underserved communities in computing and diversify the first stage of the AI development pipeline – the creators of computing.

RQ2: How does data powering AI come to be?

We discussed data quality issues in machine learning by tracing the data supply chain encompassing data creation, curation and usage. I situated my study design and findings in the domain of public health, a field increasingly integrating AI technologies. However, this work would benefit from broader generalisation. Our insights into the variances in the valuation of 'good' data and the inherent tensions within the supply chain, stemming from discrepancies in the perceived 'goodness' of data, offer a foundational basis for exploring these dynamics in other critical sectors such as insurance, social services, and education. Expanding this examination could enable more comprehensive generalisations of our findings and assist stakeholders across various domains in establishing context-specific definitions and standards for 'good' data.

RQ3: How does AI change the nature of work?

For the AI system we examined, there are many diverse stakeholders and perspectives. We investigated AI integration from the perspective of those implementing and using the AI system, including callers, program staff at HealthNGO, and the development team at TechOrg. We chose not to directly engage with beneficiaries since beneficiaries only interfaced with human actors over a phone call and had no interactions with the AI system. However, future work could examine the perceived impact of AI on end-users. I described implementation and design considerations around an AI system developed to support a specific public health program in India, this work has broader implications as similar systems for resource allocation are introduced in other resource-constrained settings. I hope our research can inform the ethnographic study of AI systems in other contexts, and that the discussion on the role and configuration of AI systems can shape how they are implemented in health settings and public sector programs.

RQ4 How do people perceive and experience AI?

During the study, we aimed for a gender-balanced sample, but this was challenging because of the skewed gender ratio in non-computing trades in the institutes we visited. We also tried to gain access to more stakeholders, including parents and alumni, but this was a challenge because they were geographically dispersed. Adding sites and stakeholders in future work could help towards a more holistic understanding of vocational technicians and their technology practices, perceptions and experiences with AI.

6.4 Future Work

This thesis lays the groundwork for a wide range of theoretical and empirical research in Human-AI collaboration. I propose future work in three broad areas and encourage scholars to critically examine the permeation of AI technologies in the society.

Individual –> Collective

A core contribution of this thesis is the shift in focus from analysing Human-AI collaboration through individual elements to examining it as a part of a broader, interconnected network that is influenced by socio-cultural and organisational factors alongside other human actors. Methodologically, my thesis showcased the value of studying AI systems as it evolves through multiple stakeholders and multiple stages in the AI pipeline. I hope this thesis inspires future work to take a comprehensive view in examining Human-AI collaboration and interactions. This has been a recurrent theme across my work – I report findings through a holistic multi-stakeholder and multi-stage study design. I believe the success of Human-AI collaboration hinges on integrating diverse perspectives while designing AI systems. The dynamic and constantly evolving interaction between human and non-human actors or elements observed in this thesis suggests a future research direction that focuses on developing iterative and dynamic interaction patterns within Human-AI collaboration systems. Prior work (Rong et al., 2023) in the domain of explainable AI and interpretable AI, such as that by Kawakami et. al. (Kawakami et al., 2023), is a good starting point to explore the dynamic nature of decision making through AI-mediated systems. This can be approached through longitudinal studies and participatory design processes that actively involve diverse stakeholders, including those from underrepresented groups.

Furthermore, this thesis has highlighted the evolving nature of Human-Human collaboration due to the introduction of AI systems. The current paradigm of Human-AI collaboration primarily focuses on optimizing the interaction between individual users and AI. However, there is a significant opportunity to extend this research to examine the design of systems that proactively facilitate and improve collaboration amongst multi-stakeholder teams. This shift from an individual-level to a system-level focus in Human-AI collaboration requires the development of new methodologies (e.g explaining AI-decisions through user interfaces (Cheng et al., 2019) that emphasize collective goals and shared outcomes, considering the diverse roles and contributions of all team members.

Multi-stakeholder Co-Design

Next, future work should examine co-design of AI technologies, motivated by the findings of this thesis, to incorporate socio-cultural and organisational context into the design of AI technologies from the outset. To integrate these principles in human-centered design requires a deep partnership between HCI researchers with AI developers and the communities where they seek to impact. Hostein et. al.'s approach (Holstein, McLaren, & Aleven, 2019) is well motivated in this direction. They showcase a case-study of successful engagement of non-technical stakeholders in AI design through co-design a Teacher-AI collaboration experience through a learning analytics tool.

However, we need additional work in multi-stakeholder settings to build diverse perspectives in co-design of AI systems. This line of inquiry can particularly focus on uncovering mental models of end-users and design contextually relevant systems that account for historically underserved communities, especially those in the global south. Addressing potential challenges such as varying power dynamics in these partnerships will be crucial.

Additionally, co-design of AI poses novel challenges such as non-determinism of AI models that requires designers adapting to dynamic stakeholders. Designers and developers are facing challenges in balancing user agency and control with AI's emergent abilities. There is a broad opportunity for HCI researchers to reimagine the design opportunity in the multistakeholder setting. Emergent work by Zhang et. al. and Subramonyam et. al. showcases the value of data probes (Subramonyam, Seifert, & Adar, 2021; A. Zhang, Boltz, Lynn, Wang, & Lee, 2023) that use user-data as design probes for improved collaboration across stakeholders (e.g : designers and developers, gig workers and designers). This approach holds value and future work could extend the paradigm of data probes onto multi-stakeholder settings to co-design AI systems.

LLMs and Future of Work

Lastly, extending the line of inquiry of effective Human-AI collaboration will be timely to examine the future of work in the paradigm of large language models (LLMs). LLMs are increasingly more capable in tasks such as programming, contract drafting through tools such as Github copilot, chatGPT, Bard. What does the emergence of LLMs mean for entry-level workers that perform work (e.g. paralegals) that is vulnerable to emergent capabilities through LLMs (Forum, n.d.; McKinsey, n.d.)? Are there differences in productivity gains / loss across seniority levels in organisations? How does this dynamic evolve as we situate work across varying socio-cultural contexts? This shift is not just a displacement of labor but also a transformation of work nature, requiring a rethinking of skill sets and job training methodologies.

Researchers have started to measure and examine productivity gains from LLMs when integrated for knowledge work (Cambon et al., n.d.; Dell'Acqua et al., 2023; Noy & Zhang, 2023). Notably, Dell'Acqua et. al. (Dell'Acqua et al., 2023) ran a study with consultants from Boston Consulting Group and found that consultants using AI were significantly more productive and produced higher quality results. They noted the variance across adoption where some consultants performed equal division of tasks with AI versus others who engaged continually while completing a task. However, consultants were 19 percent more likely to produce poor quality work if the LLM made mistakes.

The role of HCI in this paradigm is crucial. Future research should focus on developing interfaces and systems that foster effective collaboration between (and amongst) humans and LLMs. This involves understanding and building upon the mental models of users, enabling them to recognize and leverage the capabilities and limitations of LLMs. Such an approach requires new research across trust calibration, algorithmic auditing, and a thorough understanding of workflow integration. As LLMs evolve and become more ingrained in professional roles, the need for upskilling workers and addressing potential ethical and societal challenges becomes increasingly significant.

My contributions to my papers have been attested by my co-authors: Azra Ismail, Milind

Tambe, Nithya Sambasivan, and Neha Kumar.

Table 6.1: Summary of Academic Contributions.

Primary Contributor		
Paper Title	**Publication Venue**	**Contribution**
The Unexpected Entry and Exodus of Indian Women in CS and HCI	ACM CHI Conference on Human Factors in Computing Systems (CHI 2018)	Co-led the research study across design, conduct, analysis, and publication
Towards an AI-powered Future that Works for Vocational Workers	ACM CHI Conference on Human Factors in Computing Systems (CHI 2020)	Led the research study across design, conduct, analysis, and publication
When is ML Data Good? Datafication in Public Health in India	ACM CHI Conference on Human Factors in Computing Systems (CHI 2022)	Co-led the research study across design, conduct, analysis, and publication
Public Health Calls with/for AI	Proceedings of the ACM on Human-Computer Interaction (PACM)	Co-led the research study across design, conduct, analysis, and publication
Collaborator		
Paper Title	**Publication Venue**	**Contribution**
Measuring Data Collection Quality for Community Healthcare	ACM conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '21)	Contributed to study design and analysis
Increasing Impact of Mobile Health Programs: SAHELI for Maternal and Child Care	The Thirty-Fifth Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-23)	Contributed to study design and analysis

References

Abebe, R., Barocas, S., Kleinberg, J., Levy, K., Raghavan, M., & Robinson, D. G. (2020). Roles for computing in social change. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 252–260).

Ahmed, S. I., Haque, M. R., Guha, S., Rifat, M. R., & Dell, N. (2017). Privacy, security, and surveillance in the global south: A study of biometric mobile sim registration in bangladesh. In *Proceedings of the 2017 chi conference on human factors in computing systems* (pp. 906–918).

Ajunwa, I., Friedler, S., Scheidegger, C. E., & Venkatasubramanian, S. (2016). Hiring by algorithm: predicting and preventing disparate impact. *Available at SSRN*.

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., … others (2019). Guidelines for human-AI interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (p. 3). ACM.

Ammitzbøll Flügge, A., Hildebrandt, T., & Møller, N. H. (2021, April). Street-Level Algorithms and AI in Bureaucratic Decision-Making: A Caseworker Perspective. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), 1–23. Retrieved 2022-09-03, from `https://dl.acm.org/doi/10.1145/3449114` doi: 10.1145/3449114

Balayn, A., Kulynych, B., & Guerses, S. (2021). Exploring data pipelines through the process lens: a reference model forcomputer vision. *arXiv preprint arXiv:2107.01824*.

Bank, T. W. (2016). *Labor force participation rate, female.* (https://goo.gl/rFRWkQ)

Bank, T. W. (April 2019). *Labor force participation rate, female.* https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS.

Bansal, G., Nushi, B., Kamar, E., Horvitz, E., & Weld, D. S. (2021). Is the most accurate ai the best teammate? optimizing ai for teamwork. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 35, pp. 11405–11414).

Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., … Weld, D. (2021). Does the whole exceed its parts? the effect of ai explanations on complementary team performance. In *Proceedings of the 2021 chi conference on human factors in computing systems* (pp. 1–16).

Bardhan, A. (n.d.). Men are creating ai girlfriends and then verbally abusing them =https://futurism.com/chatbot-abuse, year = 2022 [Computer software manual].

Bardzell, S. (2010). Feminist hci: taking stock and outlining an agenda for design. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1301–1310). ACM.

Bardzell, S., Churchill, E., Bardzell, J., Forlizzi, J., Grinter, R., & Tatar, D. (2011). Feminism and interaction design. In *Chi'11 extended abstracts on human factors in computing systems* (pp. 1–4).

Baru, R., Acharya, A., Acharya, S., Kumar, A. S., & Nagaraj, K. (2010). Inequities in access to health services in india: caste, class and region. *Economic and political Weekly*, 49–58.

Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. *ACM computing surveys (CSUR)*, *41*(3), 1–52.

Batool, A., Toyama, K., Veinot, T., Fatima, B., & Naseem, M. (2021). Detecting data falsification by front-line development workers: A case study of vaccination in pakistan. In *Proceedings of the 2021 chi conference on human factors in computing systems* (pp. 1–14).

Beede, E., Baylor, E., Hersch, F., Iurchenko, A., Wilcox, L., Ruamviboonsuk, P., & Vardoulakis, L. M. (2020). A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–12).

Beer, D. (2009). Power through the algorithm? participatory web cultures and the technological unconscious. *New Media & Society*, *11*(6), 985–1002.

Biernacki, P., & Waldorf, D. (1981). Snowball sampling: Problems and techniques of chain referral sampling. *Sociological methods & research*, *10*(2), 141–163.

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European law journal*, *13*(4), 447–468.

Brown, A., Chouldechova, A., Putnam-Hornstein, E., Tobin, A., & Vaithianathan, R. (2019). Toward Algorithmic Accountability in Public Services: A Qualitative Study of Affected Community Perspectives on Algorithmic Decision-making in Child Welfare Services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 41:1–41:12). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3290605.3300271` (event-place: Glasgow, Scotland Uk) doi: 10.1145/3290605.3300271

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77–91).

Cai, C. J., Winter, S., Steiner, D., Wilcox, L., & Terry, M. (2019). " hello ai": uncovering the onboarding needs of medical practitioners for human-ai collaborative decision-making. *Proceedings of the ACM on Human-computer Interaction*, *3*(CSCW), 1–24.

Cambon, A., Hecht, B., Edelman, B., Ngwe, D., Jaffe, S., Heger, A., … others (n.d.). Early llm-based tools for enterprise information workers likely provide meaningful boosts to productivity.

Cao, S., & Huang, C.-M. (2022, nov). Understanding user reliance on ai in assisted decision-making. *Proc. ACM Hum.-Comput. Interact.*, *6*(CSCW2). Retrieved from `https://doi.org/10.1145/3555572` doi: 10.1145/3555572

Cave, S., Craig, C., Dihal, K. S., Dillon, S., Montgomery, J., Singler, B., & Taylor, L. (2018). Portrayals and perceptions of ai and why they matter.

Center, P. R. (2017). Automation in everyday life. *Pew*.

Chander, A. (2016). The racist algorithm. *Mich. L. Rev.*, *115*, 1023.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis.* sage.

Cheng, H.-F., Wang, R., Zhang, Z., O'connell, F., Gray, T., Harper, F. M., & Zhu, H. (2019). Explaining decision-making algorithms through ui: Strategies to help non-expert stakeholders. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1–12).

Cheryan, S., Master, A., & Meltzoff, A. N. (2015). Cultural stereotypes as gatekeepers: increasing girls' interest in computer science and engineering by diversifying stereotypes. *Frontiers in psychology*, *6*.

Chris, W. (June 2017). *Business insider.* `https://www.businessinsider.com/sam-altman-mission-fix-american-democracy-2017-6`.

Chui, M., Harryson, M., Manyika, J., Roberts, R., Chung, R., van Heteren, A., & Nel, P. (2018). *Notes from the ai frontier: Applying ai for social good.* McKinsey Global Institute, Discussion Paper December.

Clark, E., Ross, A. S., Tan, C., Ji, Y., & Smith, N. A. (2018). Creative writing with a machine in the loop: Case studies on slogans and stories. In *23rd international conference on intelligent user interfaces* (pp. 329–340).

Cohoon, J. M., Wu, Z., & Chao, J. (2009). Sexism: toxic to women's persistence in cse doctoral programs. In *Acm sigcse bulletin* (Vol. 41, pp. 158–162).

Correll, S. J. (2004). Constraints into preferences: Gender, status, and emerging career aspirations. *American sociological review*, *69*(1), 93–113.

Dame Vivian Hunt, D. L., & Prince, S. (2015). *Why diversity matters].* McKinsey.

Das, S., Jain-Chandra, S., Kochhar, K., & Kumar, N. (2015). Women workers in india: why so few among so many?

Datta, A., Sen, S., & Zick, Y. (2016). Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. In *2016 ieee symposium on security and privacy (sp)* (pp. 598–617).

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., … Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of ai on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*(24-013).

Denton, E., Hanna, A., Amironesei, R., Smart, A., Nicole, H., & Scheuerman, M. K. (2020). Bringing the people back in: Contesting benchmark machine learning datasets. *arXiv preprint arXiv:2007.07399*.

DeRenzi, B., Dell, N., Wacksman, J., Lee, S., & Lesh, N. (2017). Supporting community health workers in india through voice-and web-based feedback. In *Proceedings of the 2017 chi conference on human factors in computing systems* (pp. 2770–2781). ACM.

Desk, F. N. (2015). *Crimes against women.* http://tinyurl.com/crm-wmn-2022.

Dillahunt, T. R., Lam, J., Lu, A., & Wheeler, E. (2018). Designing future employment applications for underserved job seekers: A speed dating study. In *Proceedings of the 2018 designing interactive systems conference* (pp. 33–44).

Dove, G., Halskov, K., Forlizzi, J., & Zimmerman, J. (2017). Ux design innovation: Challenges for working with machine learning as a design material. In *Proceedings of the 2017 chi conference on human factors in computing systems* (p. 278–288). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3025453.3025739 doi: 10.1145/3025453.3025739

Dray, S. M., Busse, D. K., Brock, A. M., Peters, A. N., Bardzell, S., Druin, A., … others (2014). Perspectives on gender and product design. In *Chi'14 extended abstracts on human factors in computing systems* (pp. 53–56).

Dray, S. M., Peters, A. N., Brock, A. M., Peer, A., Druin, A., Gitau, S., … Murray, D. (2013). Leveraging the progress of women in the hci field to address the diversity chasm. In *Chi'13 extended abstracts on human factors in computing systems* (pp. 2399–2406).

Economist, T. (2018, April). *A study finds nearly half of jobs are vulnerable to automation.* https://www.economist.com/graphic-detail/2018/04/24/a-study-finds-nearly-half-of-jobs-are-vulnerable-to-automation.

Eichhorst, W., Rodríguez-Planas, N., Schmidl, R., & Zimmermann, K. (2012). A roadmap to vocational education and training systems around the world.

Escueta, M., Saxena, T., & Aggarwal, V. (2013 (accessed August 20, 2017)). *Women in engineering:a comparative study of barriers across nations* (Tech. Rep.). Aspiring

Minds India. (https://goo.gl/kVLVEh)

Finkel, S. K., & Olswang, S. G. (1996). Child rearing as a career impediment to women assistant professors. *The Review of Higher Education*, *19*(2), 123–139.

for Family Studies (IFS), I. ("2016 (accessed August 23, 2017)"). Better-educated women still prefer higher-earning husbands [Computer software manual]. (https://goo.gl/7KZRyL)

Forum, W. E. (n.d.). Jobs of tomorrow: Large language models and jobs https://www3.weforum.org/docs/wef_jobs_of_tomorrow_generative_ai_2023.pdf, year = 2023 [Computer software manual].

Frey, C. B., & Osborne, M. A. (2017). The future of employment: how susceptible are jobs to computerisation? *Technological forecasting and social change*, *114*, 254–280.

Friend, T. (2018, May). *How frightened should we be of ai.* https://www.newyorker.com/magazine/2018/05/14/how-frightened-should-we-be-of-ai.

Fuller, J. (2019, May). *Here are the grantees of the google ai impact challenge.* Google. Retrieved from https://www.blog.google/outreach-initiatives/google-org/ai-impact-challenge-grantees/

Gal, Y., Segal, A., Kamar, E., Horvitz, E., Lintott, C., & Walmsley, M. (2022). A new workflow for human-ai collaboration in citizen science. In *Proceedings of the 2022 acm conference on information technology for social good* (pp. 89–95).

Galpin, V. (2002). Women in computing around the world. *ACM SIGCSE Bulletin*, *34*(2), 94–100.

Geiger, R. S., Yu, K., Yang, Y., Dai, M., Qiu, J., Tang, R., & Huang, J. (2020). Garbage in, garbage out? do machine learning application papers in social computing report where human-labeled training data comes from? In *Proceedings of the 2020 conference* *on fairness, accountability, and transparency* (pp. 325–336).

Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in britain. *The review of economics and statistics*, *89*(1), 118–133.

govtjobsportal. (2019, August). http://www.govtjobsportal.in.

Green, B., & Chen, Y. (2019). Disparate interactions: An algorithm-in-the-loop analysis of fairness in risk assessments. In *Proceedings of the conference on fairness, accountability, and transparency* (pp. 90–99).

Gregg, M. ("2015 (accessed August 8, 2017)"). The deficiencies of tech's 'pipeline' metaphor [Computer software manual]. (https://goo.gl/SPKyBu)

Group, T. W. B. (2017, September). *Trouble in the making? the future of manufacturing-led development.* http://www.worldbank.org/en/topic/competitiveness/publication/trouble-in-the-making-the-future-of-manufacturing-led-development.

Gudivada, V., Apon, A., & Ding, J. (2017). Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations. *International Journal on Advances in Software*, *10*(1), 1–20.

Guo, P. J., Kandel, S., Hellerstein, J. M., & Heer, J. (2011). Proactive wrangling: Mixed-initiative end-user programming of data transformation scripts. In *Proceedings of the 24th annual acm symposium on user interface software and technology* (pp. 65–74).

Gupta, N. (2012). Women undergraduates in engineering education in india: A study of growing participation. *Gender, Technology and Development*, *16*(2), 153–176.

Gupta, N. (2015). Rethinking the relationship between gender and technology: a study of the indian example. *Work, employment and society*, *29*(4), 661–672.

Gupta, N., & Sharma, A. K. (2003). Gender inequality in the work environment at institutes of higher learning in science and technology in india. *Work, employment and society*, *17*(4), 597–616.

Hagel, J., Schwartz, J., & Bersin, J. (2017, July). *Navigating the future of work: Can we point business, workers, and social institutions in the same direction?* https://www2.deloitte.com/insights/us/en/deloitte-review/issue-21/navigating-new-forms-of-work.html.

Hamal, M., Dieleman, M., De Brouwere, V., & de Cock Buning, T. (2020). Social determinants of maternal health: a scoping review of factors influencing maternal mortality and maternal health service use in india. *Public Health Reviews*, *41*(1), 1–24.

Harnessing automation for a future that works | mckinsey & company. (2017, January). https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works. ((Accessed on 08/18/2018))

Henn, S. (n.d.). When women stopped coding =https://goo.gl/xqky72, month = October, year = 2014 [Computer software manual].

Heuts, F., & Mol, A. (2013). What is a good tomato? a case of valuing in practice. *Valuation Studies*, *1*(2), 125–146.

Hill, C., Corbett, C., & St Rose, A. (2010). *Why so few? women in science, technology, engineering, and mathematics.* ERIC.

Hiringlab.org. (2019, January). *Who's afraid of automation?* https://www.hiringlab.org/2019/01/10/automation-ai-fears-jobs/.

Hnatkovska, V., Lahiri, A., & Paul, S. B. (2013). Breaking the caste barrier intergenerational mobility in india. *Journal of Human Resources*, *48*(2), 435–473.

Hohman, F., Head, A., Caruana, R., DeLine, R., & Drucker, S. M. (2019). Gamut: A design probe to understand how data scientists understand machine learning models.

In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1–13).

Hohman, F., Wongsuphasawat, K., Kery, M. B., & Patel, K. (2020). Understanding and visualizing data iteration in machine learning. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–13).

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–ai complementarity. *Journal of Learning Analytics*, *6*(2).

Holstein, K., Wortman Vaughan, J., Daumé, H., III, Dudik, M., & Wallach, H. (2019). Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 600:1–600:16). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3290605.3300830` (event-place: Glasgow, Scotland Uk) doi: 10.1145/3290605.3300830

Holten Møller, N., Shklovski, I., & Hildebrandt, T. T. (2020, October). Shifting Concepts of Value: Designing Algorithmic Decision-Support Systems for Public Services. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (pp. 1–12). Tallinn Estonia: ACM. Retrieved 2022-09-03, from `https://dl.acm.org/doi/10.1145/3419249.3420149` doi: 10.1145/3419249.3420149

Hutchinson, B., Smart, A., Hanna, A., Denton, E., Greer, C., Kjartansson, O., … Mitchell, M. (2021). Towards accountability for machine learning datasets: Practices from software engineering and infrastructure. In *Proceedings of the 2021 acm conference on fairness, accountability, and transparency* (pp. 560–575).

IBM Science for Social Good. (2018). Retrieved from https://www.research.ibm.com/science-for-social-good/

Isaac, M. (n.d.). Uber founder travis kalanick resigns as c.e.o. =https://goo.gl/w7wz1h, month = June, year = 2017 [Computer software manual].

Ismail, A., & Kumar, N. (2018). Engaging solidarity in data collection practices for community health. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 76.

Ismail, A., & Kumar, N. (2019). Empowerment on the margins: The online experiences of community health workers. In *Proceedings of the 2019 chi conference on human factors in computing systems* (p. 99).

Ismail, A., & Kumar, N. (2021). Ai in global health: the view from the front lines. In *Proceedings of the 2021 chi conference on human factors in computing systems* (pp. 1–21).

Ismail, A., Thakkar, D., Madhiwalla, N., & Kumar, N. (2023). Public health calls for/with ai: An ethnographic perspective. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW2), 1–26.

Jakesch, M., Buçinca, Z., Amershi, S., & Olteanu, A. (2022). How different groups prioritize ethical values for responsible ai. *arXiv preprint arXiv:2205.07722*.

Jakesch, M., French, M., Ma, X., Hancock, J. T., & Naaman, M. (2019). Ai-mediated communication: How the perception that profile text was written by ai affects trustworthiness. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1–13).

James, M. (2017, May). *Technology, jobs, and the future of work | mckinsey & company.* https://www.mckinsey.com/featured-insights/employment-and-growth/

technology-jobs-and-the-future-of-work.

Jo, E. S., & Gebru, T. (2020). Lessons from archives: strategies for collecting sociocultural data in machine learning. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 306–316).

Jo, Y., Kim, M., & Han, K. (2019). How do humans assess the credibility on web blogs: Qualifying and verifying human factors with machine learning. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1–12).

Johanson, R. K., & Adams, A. V. (2004). *Skills development in sub-saharan africa.* The World Bank.

Kapania, S., Siy, O., Clapper, G., SP, A. M., & Sambasivan, N. (2022, April). "Because AI is 100% right and safe": User Attitudes and Sources of AI Authority in India. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–18). New York, NY, USA: Association for Computing Machinery. Retrieved 2022-09-03, from https://doi.org/10.1145/3491102.3517533 doi: 10.1145/3491102.3517533

Karunasena, R., Ambiya, M. S., Sinha, A., Nagar, R., Dalal, S., Abdullah, H., … Tambe, M. (2021). Measuring data collection diligence for community healthcare. In *Equity and access in algorithms, mechanisms, and optimization* (pp. 1–12).

Karusala, N., Wilson, J., Vayanos, P., & Rice, E. (2019). Street-level realities of data practices in homeless services provision. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–23.

Katell, M., Young, M., Dailey, D., Herman, B., Guetler, V., Tam, A., … Krafft, P. (2020). Toward situated interventions for algorithmic equity: lessons from the field. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 45–55).

Kawakami, A., Guerdan, L., Cheng, Y., Lee, M., Carter, S., Arechiga, N., … Holstein, K.

(2023). Training towards critical use: Learning to situate ai predictions relative to human knowledge.

Kelley, P. G., Yang, Y., Heldreth, C., Moessner, C., Sedley, A., & Woodruff, A. (2021). " mixture of amazement at the potential of this technology and concern about possible pitfalls": Public sentiment towards ai in 15 countries.

Keyes, O. (2018). The misgendering machines: Trans/hci implications of automatic gender recognition. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 88.

Kitchin, R., & Lauriault, T. (2014). Towards critical data studies: Charting and unpacking data assemblages and their work.

Klasen, S., & Pieters, J. (2015). What explains the stagnation of female labor force participation in urban india? *The World Bank Economic Review*, *29*(3), 449–478.

Kotamraju, N. P. (2011). Playing stupid, caring for users, and putting on a good show: Feminist acts in usability study work. *Interacting with Computers*, *23*(5), 439–446.

Kshirsagar, M., Robinson, C., Yang, S., Gholami, S., Klyuzhin, I., Mukherjee, S., … others (2021). Becoming good at ai for good. *arXiv preprint arXiv:2104.11757*.

Kumar, N., Wong-Villacres, M., Karusala, N., Vishwanath, A., Kumar, A., & Ismail, A. (2019). Aspirations-based design. In *Proceedings of the tenth international conference on information and communication technologies and development* (p. 2).

Latour, B. (2007). *Reassembling the social: An introduction to actor-network-theory*. Oup Oxford.

Lee, M. K., & Baykal, S. (2017). Algorithmic mediation in group decisions: Fairness perceptions of algorithmically mediated vs. discussion-based social division. In *Proceedings of the 2017 acm conference on computer supported cooperative work and social computing* (pp. 1035–1048).

Lum, K., & Isaac, W. (2016). To predict and serve? *Significance*, *13*(5), 14–19.

Madaio, M. A., Stark, L., Wortman Vaughan, J., & Wallach, H. (2020). Co-designing checklists to understand organizational challenges and opportunities around fairness in ai. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–14).

Maithreyi, R., Padmanabhan, S., Menon, N., & Jha, J. (n.d.). Skills development, social mobility and educational change: A sociological analysis of the effects of the national policy on skills development in india.

Manjoo, F. (2014, September). *Exposing hidden bias at google - the new york times.* https://goo.gl/EmaqNq. ((Accessed on 09/02/2017))

Manyika, J., Lund, S., Chui, M., Bughin, J., Jonathan, W., Batra, P., … Sanghvi, S. (2017, September). *Jobs lost, jobs gained: What the future of work will mean for jobs, skills, and wages.* http://www.worldbank.org/en/topic/competitiveness/publication/trouble-in-the-making-the-future-of-manufacturing-led-development.

Manyika, J., & Sneader, K. (2018, June). *Ai, automation, and the future of work: Ten things to solve for.* https://www.mckinsey.com/featured-insights/future-of-organizations-and-work/ai-automation-and-the-future-of-work-ten-things-to-solve-for.

Mao, Y., Wang, D., Muller, M., Varshney, K. R., Baldini, I., Dugan, C., & Mojsilović, A. (2019). How data scientistswork together with domain experts in scientific collaborations: To find the right answer or to ask the right question? *Proceedings of the ACM on Human-Computer Interaction*, *3*(GROUP), 1–23.

Marcus, G. E. (1995). Ethnography in/of the world system: The emergence of multi-sited ethnography. *Annual review of anthropology*, 95–117.

Mark, G. (n.d.). Women in cs in qatar: It's complicated =https://goo.gl/aa1uqn, month = May, year = 2010 [Computer software manual].

McKinsey. (n.d.). The economic potential of generative ai: The next productivity frontier =http://tinyurl.com/mckinsey23genai, year = 2023 [Computer software manual].

Ministry of Human Resource Development, D. o. S. E., & Literacy. (2016, February). *Educational statistics at a glance.* http://mhrd.gov.in/sites/upload_files/mhrd/files/statistics/ESG2016_0.pdf.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., … Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency* (pp. 220–229).

Møller, N. H., Bossen, C., Pine, K. H., Nielsen, T. R., & Neff, G. (2020). Who does the work of data? *Interactions*, *27*(3), 52–55.

Monroe, K., Ozyurt, S., Wrigley, T., & Alexander, A. (2008). Gender equality in academia: Bad news from the trenches, and some possible solutions. *Perspectives on politics*, *6*(2), 215–233.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479.

MSR India Center for Societal impact through Cloud and Artificial Intelligence (SCAI). (2019). Retrieved from https://www.microsoft.com/en-us/research/theme/scai/

Mukhopadhyay, C. (1996). The cultural context of gendered science: a theoretical model derived from studies in india. *Gates*, *3*(1).

Muller, M., Lange, I., Wang, D., Piorkowski, D., Tsay, J., Liao, Q. V., … Erickson, T. (2019). How data science workers work with data: Discovery, capture, curation, design, creation. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1–15).

Muller, M., Wolf, C. T., Andres, J., Desmond, M., Joshi, N. N., Ashktorab, Z., … others (2021). Designing ground truth and the social life of labels. In *Proceedings of the 2021 chi conference on human factors in computing systems* (pp. 1–16).

Murali, A., & PK, J. (2019, April). *India's bid to harness ai for healthcare.* Retrieved from `https://factordaily.com/ai-for-healthcare-in-india/`

Neang, A. B., Sutherland, W., Beach, M. W., & Lee, C. P. (2021). Data integration as coordination: The articulation of data work in an ocean science collaboration. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW3), 1–25.

Neff, G., Tanweer, A., Fiore-Gartland, B., & Osburn, L. (2017). Critique and contribute: A practice-based framework for improving critical data studies and data science. *Big data*, *5*(2), 85–97.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., … others (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593–10597.

Noy, S., & Zhang, W. (2023). *Experimental evidence on the productivity effects of generative artificial intelligence [working paper].* MIT.

of Human Resource Development, M. (2015-16). *All india survey on higher education* (Tech. Rep.). Government of India,Ministry of Human Resource Development, Department of Higher Education, Statistics Division.

of India, M. R. S. (2011, May). *The new sec system.* http://mruc.net/uploads/posts/b17695616c422ec8d9dadafc1c3eec26.pdf.

of Trade, D. G. (2017, May). *Trade syllabus.* http://dget.nic.in/content/innerpage/trade-syllabus.php.

Oh, C., Song, J., Choi, J., Kim, S., Lee, S., & Suh, B. (2018). I lead, you help but only with enough details: Understanding user experience of co-creation with artificial intelligence. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1–13).

Okada, A. (2012). Skills development for youth in india: Challenges and opportunities. *Journal of International Cooperation in Education*, *15*(2), 169–193.

Okolo, C. T., Dell, N., & Vashistha, A. (2022, June). Making AI Explainable in the Global South: A Systematic Review. In *ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS)* (pp. 439–452). Seattle WA USA: ACM. Retrieved 2022-09-03, from https://dl.acm.org/doi/10.1145/3530190.3534802 doi: 10.1145/3530190.3534802

Okolo, C. T., Kamath, S., Dell, N., & Vashistha, A. (2021). "it cannot do all of my work": community health worker perceptions of ai-enabled mobile health applications in rural india. In *Proceedings of the 2021 chi conference on human factors in computing systems* (pp. 1–20).

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization science*, *3*(3), 398–427.

Otterbacher, J., Bates, J., & Clough, P. (2017). Competent men and warm women: Gender stereotypes and backlash in image search results. In *Proceedings of the 2017 chi conference on human factors in computing systems* (pp. 6620–6631).

Pal, J. (2008). Computers and the promise of development: aspiration, neoliberalism and „technolity in india s ictd enterprise. *A paper presented at confronting the Challenge of Technology for Development: Experiences from the BRICS*, 29–30.

Panigutti, C., Beretta, A., Giannotti, F., & Pedreschi, D. (2022, April). Understanding the impact of explanations on advice-taking: a user study for AI-based clinical Decision Support Systems. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–9). New York, NY, USA: Association for Computing Machinery. Retrieved 2022-09-03, from `https://doi.org/10.1145/3491102.3502104` doi: 10.1145/3491102.3502104

Papenmeier, A., Kern, D., Hienert, D., Kammerer, Y., & Seifert, C. (2022). How accurate does it feel?–human perception of different types of classification mistakes. In *Chi conference on human factors in computing systems* (pp. 1–13).

Passi, S., & Jackson, S. (2017). Data vision: Learning to see through algorithmic abstraction. In *Proceedings of the 2017 acm conference on computer supported cooperative work and social computing* (pp. 2436–2447).

Passi, S., & Jackson, S. J. (2018). *Trust in data science: Collaboration, translation, and accountability in corporate data science projects. 2 (2018), 136: 1–136: 28. issue cscw.*

Paul, P., & Chouhan, P. (2020). Socio-demographic factors influencing utilization of maternal health care services in india. *Clinical Epidemiology and Global Health*, *8*(3), 666–670.

Pedersen, A. M., & Bossen, C. (2024). Cultivating data practices across boundaries: How organizations become data-driven. *computer supported cooperative work*(CSCW).

Piorkowski, D., Park, S., Wang, A. Y., Wang, D., Muller, M., & Portnoy, F. (2021). How ai developers overcome communication challenges in a multidisciplinary team: A case study. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), 1–25.

policy brief, O. (2017). Going digital: The future of work for women. *OECD*.

Radensky, M., Downey, D., Lo, K., Popovic, Z., & Weld, D. S. (2022). Exploring the role of local and global explanations in recommender systems. In *Chi conference on human factors in computing systems extended abstracts* (pp. 1–7).

Raghuram, P., Herman, C., Ruiz-Ben, E., & Sondhi, G. (2017). Women and it scorecard-india.

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., … Barnes, P. (2020). Closing the ai accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 33–44).

Rastogi, C., Zhang, Y., Wei, D., Varshney, K. R., Dhurandhar, A., & Tomsett, R. (2020). Deciding fast and slow: The role of cognitive biases in ai-assisted decision-making. *arXiv preprint arXiv:2010.07938*.

The rockefeller foundation establishes atlas ai – new startup to generate actionable intelligence on global development challenges. (n.d.). Retrieved from `https://www.rockefellerfoundation.org/about-us/news-media/rockefeller-foundation-establishes-atlas-ai-new-startup-generate-actionable-intelligence-global-development-challenges/`

Rong, Y., Leemann, T., Nguyen, T.-T., Fiedler, L., Qian, P., Unhelkar, V., … Kasneci, E. (2023). Towards human-centered explainable ai: A survey of user studies for model explanations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Sambasivan, N., Arnesen, E., Hutchinson, B., Doshi, T., & Prabhakaran, V. (2021a). Re-imagining algorithmic fairness in india and beyond. In *Proceedings of the 2021 acm conference on fairness, accountability, and transparency* (pp. 315–328).

Sambasivan, N., Arnesen, E., Hutchinson, B., Doshi, T., & Prabhakaran, V. (2021b, March). Re-imagining Algorithmic Fairness in India and Beyond. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 315–328). New York, NY, USA: Association for Computing Machinery. Retrieved 2022-09-03, from `https://doi.org/10.1145/3442188.3445896` doi: 10.1145/3442188.3445896

Sambasivan, N., Kapania, S., Highfill, H., Akrong, D., Paritosh, P., & Aroyo, L. (2021a). ”everyone wants to do the model work, not the data work”: Data cascades in high-stakes ai. In *proceedings of the 2021 chi conference on human factors in computing systems.*

Sambasivan, N., Kapania, S., Highfill, H., Akrong, D., Paritosh, P., & Aroyo, L. M. (2021b, May). “Everyone wants to do the model work, not the data work”: Data Cascades in High-Stakes AI. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). New York, NY, USA: Association for Computing Machinery. Retrieved 2022-09-03, from `https://doi.org/10.1145/3411764.3445518` doi: 10.1145/3411764.3445518

Saxena, D., Badillo-Urquiola, K., Wisniewski, P. J., & Guha, S. (2020). A human-centered review of algorithms used within the us child welfare system. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–15).

Saxena, D., Moon, S. Y., Shehata, D., & Guha, S. (2022). Unpacking invisible work practices, constraints, and latent power relationships in child welfare through casenote analysis. In *Chi conference on human factors in computing systems* (pp. 1–22).

Scheuerman, M. K., Denton, E., & Hanna, A. (2021). Do datasets have politics? disciplinary values in computer vision dataset development. *arXiv preprint arXiv:2108.04308*.

Seidman, I. (2006). *A guide for researchers in education and the social sciences.* New York, NY: Teachers College Press.

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency* (pp. 59–68).

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, *36*(6), 495–504.

Smith, A., & Anderson, J. (2014). Ai, robotics, and the future of jobs. *Pew Research Center*, *6*.

Smith, C. E., Yu, B., Srivastava, A., Halfaker, A., Terveen, L., & Zhu, H. (2020). Keeping community in the loop: Understanding wikipedia stakeholder values for machine learning-based systems. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–14).

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of experimental social psychology*, *35*(1), 4–28.

Sterlicchi, J. (n.d.). Sexist attitudes prevail in silicon valley = `https://goo.gl/gxltfr`, month = November, year = 2007 [Computer software manual].

Subramonyam, H., Seifert, C., & Adar, E. (2021). Towards a process model for co-creating ai experiences. In *Designing interactive systems conference 2021* (pp. 1529–1543).

Suchman, L. A. (2007). *Human-machine reconfigurations: Plans and situated actions.* Cambridge university press.

Suh, M., Youngblom, E., Terry, M., & Cai, C. J. (2021). Ai as social glue: uncovering the roles of deep generative ai during social music composition. In *Proceedings of the 2021 chi conference on human factors in computing systems* (pp. 1–11).

Suh, M. M., Youngblom, E., Terry, M., & Cai, C. J. (2021). Ai as social glue: Uncovering the roles of deep generative ai during social music composition. In *Proceedings of the 2021 chi conference on human factors in computing systems.* New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3411764.3445219 doi: 10.1145/3411764.3445219

Suresh, H., Movva, R., Dogan, A. L., Bhargava, R., Cruxen, I., Cuba, A. M., … D'Ignazio, C. (2022, June). Towards Intersectional Feminist and Participatory ML: A Case Study in Supporting Feminicide Counterdata Collection. In *2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 667–678). New York, NY, USA: Association for Computing Machinery. Retrieved 2022-09-03, from https://doi.org/10.1145/3531146.3533132 doi: 10.1145/3531146.3533132

Sutton, C., Hobson, T., Geddes, J., & Caruana, R. (2018). Data diff: Interpretable, executable summaries of changes in distributions for data wrangling. In *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining* (pp. 2279–2288).

Tapia, A. H., & Kvasny, L. (2004). Recruitment is never enough: Retention of women and minorities in the it workplace. In *Proceedings of the 2004 sigmis conference on computer personnel research: Careers, culture, and ethics in a networked environment* (pp. 84–91).

Taylor, A. S. (2011). Out there. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 685–694).

Thakkar, D., Ismail, A., Kumar, P., Hanna, A., Sambasivan, N., & Kumar, N. (2022). When is machine learning data good?: Valuing in public health datafication. In *Chi conference on human factors in computing systems* (pp. 1–16).

Thakkar, D., Kumar, N., & Sambasivan, N. (2020). Towards an ai-powered future that works for vocational workers. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–13).

Thakkar, D., Sambasivan, N., Kulkarni, P., Kalenahalli Sudarshan, P., & Toyama, K. (2018). The unexpected entry and exodus of women in computing and hci in india. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1–12).

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation*, *27*(2), 237–246.

Toyama, K. (2018). From needs to aspirations in information technology for development. *Information Technology for Development*, *24*(1), 15–36.

Understanding the public perception of ai. (2019, February). `https://medium.com/goodai-news/understanding-the-public-perception-of-ai-a14b0e6b6154`.

Varma, R. (2010). Forms of social asymmetry and cultural bias: Of gender and science in india and the world. *Journal of Women and Minorities in Science and Engineering*, *16*.

Venkatesh, S. (2015). Forms of social asymmetry and cultural bias: Of gender and science in india and the world. *Transcience (2015)*, *6*(1).

Verma, S., Singh, G., Mate, A., Verma, P., Gorantla, S., Madhiwalla, N., … others (2023). Increasing impact of mobile health programs: Saheli for maternal and child care. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 37, pp. 15594–15602).

Walsham, G. (1997). Actor-network theory and is research: current status and future prospects. *Information systems and qualitative research*, 466–480.

Wang, D., Weisz, J. D., Muller, M., Ram, P., Geyer, W., Dugan, C., … Gray, A. (2019). Human-ai collaboration in data science: Exploring data scientists' perceptions of automated ai. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–24.

Wang, D., Yang, Q., Abdul, A., & Lim, B. Y. (2019). Designing theory-driven user-centric explainable ai. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1–15).

West, D. M. (2018, May). *Brookings survey finds worries over ai impact on jobs and personal privacy, concern u.s. will fall behind china.* `https://www.brookings.edu/blog/techtank/2018/05/21/brookings-survey-finds-worries-over-ai-impact-on-jobs-and-personal-privacy-concern-u-s-will-fall-behind-china/`.

Wolfinger, N. H., Mason, M. A., & Goulden, M. (2008). Problems in the pipeline: Gender, marriage, and fertility in the ivory tower. *The Journal of Higher Education*, *79*(4), 388–405.

Woodruff, A., Fox, S. E., Rousso-Schindler, S., & Warshaw, J. (2018). A qualitative exploration of perceptions of algorithmic fairness. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1–14).

Woodruff, A., Shelby, R., Kelley, P. G., Rousso-Schindler, S., Smith-Loud, J., & Wilcox, L. (2023). How knowledge workers think generative ai will (not) transform their industries. *arXiv preprint arXiv:2310.06778*.

Yan, J. N., Gu, Z., Lin, H., & Rzeszotarski, J. M. (2020). Silva: Interactively assessing machine learning fairness using causality. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–13).

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining whether, why, and how human-ai interaction is uniquely difficult to design. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–13).

Yang, Q., Steinfeld, A., & Zimmerman, J. (2019). Unremarkable ai: Fitting intelligent decision support into critical, clinical decision-making processes. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1–11). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/3290605.3300468 (event-place: Glasgow, Scotland Uk) doi: 10.1145/3290605.3300468

Yin, M., Wortman Vaughan, J., & Wallach, H. (2019). Understanding the Effect of Accuracy on Trust in Machine Learning Models. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 279:1–279:12). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/3290605.3300509 (event-place: Glasgow, Scotland Uk) doi: 10.1145/3290605.3300509

Zhang, A., Boltz, A., Lynn, J., Wang, C.-W., & Lee, M. K. (2023). Stakeholder-centered ai design: Co-designing worker tools with gig workers through data probes. In *Proceedings of the 2023 chi conference on human factors in computing systems* (pp. 1–19).

Zhang, A. X., Muller, M., & Wang, D. (2020). How do data science workers collaborate? roles, workflows, and tools. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW1), 1–23.

Zhang, B., & Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. *Available at SSRN 3312874*.

Zhang, R., McNeese, N. J., Freeman, G., & Musick, G. (2021). " an ideal human" expectations of ai teammates in human-ai teaming. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW3), 1–25.

Zhu, H., Yu, B., Halfaker, A., & Terveen, L. (2018). Value-sensitive algorithm design: Method, case study, and lessons. *Proceedings of the ACM on human-computer interaction*, *2*(CSCW), 1–23.