Designing for Participation and Power in Data Collection and Analysis

by

Jonathan Zong

B.S.E., Princeton University (2018)
S.M., Massachusetts Institute of Technology (2020)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2024

Authored by:	Jonathan Zong Department of Electrical Engineering and Computer Science August 12, 2024
Certified by:	Arvind Satyanarayan Associate Professor of Electrical Engineering and Computer Science Thesis Supervisor
Accepted by:	Leslie A. Kolodziejski Professor of Electrical Engineering and Computer Science Chair, Department Committee on Graduate Students

Designing for Participation and Power in Data Collection and Analysis

by

Jonathan Zong

Submitted to the Department of Electrical Engineering and Computer Science
on August 12, 2024 in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

Technologies that mediate social participation are an increasingly important area for design, enabling people to create, share, and discuss information. While increased participation is generally considered empowering, it can also be a double-edged sword, as involuntary participation in systems can lead to disempowerment. In this dissertation, I apply the lens of participation and power to two problem domains: accessible data visualization and ethical data collection. First, existing approaches to accessible data visualization reinforce blind and low-vision (BLV) users' dependence on sighted assistance. In contrast, I design systems that empower BLV users to conduct self-guided data exploration and create non-visual representations without using visual idioms. Second, existing data ethics procedures are often designed to offer people more choices, but can serve to placate users and consolidate data collectors' power. I develop systems and frameworks that offer novel approaches to data protection by reframing people's non-compliance with data collection as a form of socio-technical design. Altogether, this work demonstrates how the lens of participation and power deepens our understanding of technology's social implications and inspires novel approaches to design.

Thesis supervisor: Arvind Satyanarayan
Title: Associate Professor of Electrical Engineering and Computer Science

Thesis reader: Daniel Jackson
Title: Professor of Computer Science

Thesis reader: Catherine D'Ignazio
Title: Associate Professor of Urban Science and Planning

Thesis reader: J. Nathan Matias
Title: Assistant Professor of Communication, Cornell University

Acknowledgments

I got my start in computer science in Don Adams's classroom in high school. As a kid, I'd spent a lot of time on the computer to an extent that probably worried my parents. Learning to program in Mr. Adams's class, I started to see how software was not only something that I could use, but also something I could create, shape, and express myself with. Mr. Adams encouraged me to explore this interest by unlocking the computer science room for me after school, so I could read books and documentation or just hang out. He spent a lot of his own time supporting the computer science club, even driving or flying with us to programming competitions on weekends. I'm so grateful for the way he enabled and amplified my curiosity and opened doors for me literally and figuratively.

Now, as I finish my Ph.D. at MIT, I can see that Mr. Adams was only the first in a long line of mentors who have gone out of their way to help me succeed. I'm grateful for Princeton mentors who pushed me to explore the creative, artistic, and humanistic sides of computing, including David Reinfurt, Jeff Snyder, and Aatish Bhatia. I'm also grateful for teachers who helped me broaden my horizons in the arts, like Jane Cox and Judith Hamera. I'm also deeply grateful for J. Nathan Matias, my thesis advisor in undergrad, who I've been lucky to continue to write and think with while I've been at MIT. I'm thankful for Nathan's belief in me, and the sense of perspective and clarity he brings to helping me and other junior scholars navigate academia.

At MIT, I've been incredibly lucky to have Arvind Satyanarayan as an advisor. Arvind has been one of my greatest supporters and conversation partners. He's encouraged me to become a more precise thinker, to value conceptual depth, and to think carefully about how I can influence my professional communities for the better. There are many things about how Arvind operates as a thinker and collaborator that I hope to emulate in the future: his skills as a writer and communicator, his knack for giving really great advice, and the way he has shepherded the Visualization Group into a uniquely supportive and uplifting space for junior researchers.

In the MIT Visualization Group, I want to thank the many past and former members who have shaped my growth in the Ph.D., including: Alan Lundgard, Aspen Hopkins, Crystal Lee, Wonyoung So, Josh Pollock, Dylan Wootton, Angie Boggust, Harini Suresh, Nava Haghighi, Ben Tang, Amy Fox, and Vineet Pandey. Thank you to the student collaborators I've worked with, including: Isabella Pedraza Pineros, Katie Chen, Evan Caragay, Shuli Jones, Matt Blanco, EJ Sefah, Katie Bacher, and Rupayan Neogy. Particular thank you to Crystal Lee for being a constant friend, supporter, and recipient of editing permissions on Google Docs. Finally, thank you to Daniel Hajas for our collaboration and friendship.

Thanks to my thesis committee for their valuable insights and feedback on this document:

Arvind Satyanarayan, Daniel Jackson, Catherine D'Ignazio, and J. Nathan Matias.

Around MIT, I'd like to thank the folks at the Office of the Arts: Sarah Hirzel, Heidi Erickson, Stacy DeBartolo, and Andrea Volpe. I'd also like to thank the leadership of MIT Morningside Academy for Design, for creating a wonderful space for strange people like me to thrive: John Ochsendorf, Maria Yang, Marion Cunningham, Adélaïde Zollinger.

I'd also like to thank these people for inspiring me and teaching me how to care for my community: Lilly Chin, Rodrigo Ochigame, Mani Mengiste, Beth Semel, Ed Bertschinger.

Beyond MIT, thank you to the Paul and Daisy Soros Fellowships for New Americans for not only supporting my work, but also creating a space where I could build lasting friendships with others. Thank you to Craig Harwood, Yulian Ramos, and Nikka Landau. Shoutout to Ariel Chu, Diana Delgado Cornejo, Samir Paul, Hilda Huang, Grace Pan, Grace Zhang, Shreyas Vissapragada, Sal Fu, and many others.

Perhaps the best part of this job is that academia connects us with people from all over who can share in our interests and values. I'm so thankful for the people I've gotten to know through conferences, group chats, and Twitter, including: Emily Tseng, Calvin Liang, Andrew McNutt, Frank Elavsky, Morgan Scheuerman, Pedro Reynolds-Cuéllar, Derya Akbaba, Jacob Ritchie, Jingyi Li, Sharon Heung, Jen Liu, Alicia DeVrio, Meg Young, and more.

Thank you to my dear friends: Drew Wallace, Kathleen Ma, Helen Lin, Eric Li, Sarah Wang.

Thank you to CW Taekwondo for being a home away from home, especially: Rachel Walker, Justin Durand, Amy Miao, Sam Sexton, Heeyun Choi Kim, KP Hunsinger, Anye Li, Avi Baral.

Thank you to my community at Highrock Cambridge (and by extension Josh's d&d group), including: Josh Chi, Alexis Drake, Esther Kim, Zach Clemente, Jamie Baik, Jenny Kim.

Shout out to my therapist!

And thank you to these streamers for playing games in the background while I wrote up a lot of this research: WolfeyVGC, Gawr Gura, Disguised Toast, Bynx_Plays, RegisKillbin.

Last, but certainly not least, thank you to my family: Helen Sun, Janet Zong York, Shuh Chang. I'm so grateful to have the unconditional support of my family in everything that I do, and thank you for being there both to help me through difficult times and celebrate my successes.

Biographical Sketch

Jonathan Zong was born in Houston, Texas after his parents immigrated there from China to pursue graduate school. He owes his curiosity and sense of social responsibility to his mother and maternal grandparents, who captivated him with stories about this and other worlds from an early age. In high school, he became an Eagle Scout, performed as a cellist in the orchestra, competed at the state and national level in programming competitions and hackathons, and created Android apps used by tens of thousands of people.

Jonathan received his B.S.E. degree in Computer Science and Certificate in Visual Arts from Princeton University in 2018. His computer science thesis, advised by J. Nathan Matias, investigated empirical methods for studying internet research ethics. His visual arts thesis, advised by David Reinfurt, was an exhibition exploring how his discomfort with authority and power—especially his own—shapes his identity. While at Princeton, he created research-based visual art that influenced discussions about technology in the New York Times and exhibited at the Centre National du Graphisme in Chaumont, France. He interned as a software engineer and graphic designer at companies including Coursera, Square, Linked by Air, and Google.

At MIT, Jonathan worked toward a Ph.D. in Computer Science with a minor in Science, Technology, and Society (STS). He was advised by Arvind Satyanarayan and worked closely with the MIT Visualization Group, a part of the Computer Science and Artificial Intelligence Laboratory (CSAIL). In his research, Jonathan partnered with blind collaborators and study participants to co-design interfaces for non-visual data exploration. These interfaces support the autonomy and agency of blind and low-vision users to conduct self-guided data analysis and help them establish common ground with sighted users with shared spatial data metaphors. In addition to his accessible visualization work, he collaborated with the Cornell Citizens and Technology Lab to design open-source software for managing the ethics of consent in large-scale social media research. He also developed conceptual frameworks for designers and activists to develop strategies for refusing mass data collection by powerful actors. While at MIT, his work was recognized with numerous awards and honors, including the NSF Graduate Research Fellowship, Paul and Daisy Soros Fellowship for New Americans, the MIT Morningside Academy for Design Fellowship, and Forbes 30 Under 30 Scientists. His artworks were exhibited at MIT's Wiesner Gallery and the Special Exhibitions Gallery of Harvard's Collection of Historical and Scientific Instruments.

Contents

Title page **1**

Abstract **3**

Acknowledgments **5**

Biographical Sketch **7**

List of Figures **13**

List of Tables **17**

I Introduction **19**

1 Introduction **21**
- 1.1 Thesis Contributions . . . 22
  - 1.1.1 Accessible Data Analysis . . . 22
  - 1.1.2 Ethical Data Collection . . . 24
  - 1.1.3 Participation and Power: A Unifying Conceptual Lens . . . 26
- 1.2 Interpreting the Contributions . . . 26
  - 1.2.1 Positionality . . . 26
  - 1.2.2 Key concepts . . . 27
  - 1.2.3 Designing for Participation and Power . . . 29
- 1.3 Prior Publications and Authorship . . . 31

2 Related Work **33**
- 2.1 Participation and Power in Design . . . 33
- 2.2 Accessibility as Empowered Participation . . . 34
  - 2.2.1 Screen Reader Accessibility for Blind and Low-Vision Users . . . 34
  - 2.2.2 Accessible Data Representations . . . 34
  - 2.2.3 Accessible Authoring Tools for Non-Visual Representations . . . 36
- 2.3 Refusal as Empowered Non-Participation . . . 36
  - 2.3.1 Data Refusal . . . 36
  - 2.3.2 Refusal and Standpoint . . . 37
  - 2.3.3 Design for Consent and Refusal . . . 39

II Empowered Participation in Data Analysis 43

3 Introduction 45

4 Rich Screen Reader Experiences for Accessible Data Visualization 47
4.1 Introduction . . . 47
4.2 Design Dimensions for Rich Screen Reader Experiences . . . 48
4.2.1 Structure . . . 49
4.2.2 Navigation . . . 51
4.2.3 Description . . . 53
4.3 Example Gallery . . . 55
4.4 Evaluation . . . 56
4.4.1 Evaluation Setup & Design . . . 56
4.4.2 Quantitative Results . . . 57
4.4.3 Qualitative Results . . . 59
4.5 Discussion and Future Work . . . 62
4.5.1 Enabling Richer Screen Reader Experiences Today . . . 62
4.5.2 Studying and Refining the Design Dimensions . . . 63
4.5.3 What are Accessible Interactions for Data Visualizations? . . . 63

5 Accessible Structured Editing of Multi-Modal Data Representations 65
5.1 Introduction . . . 65
5.2 Motivation: De-Centering the Visual Modality . . . 67
5.2.1 Design Goals . . . 67
5.2.2 Authoring Multimodal Representations Co-Equally: An Example Usage Scenario . . . 68
5.3 The Umwelt Editor . . . 71
5.3.1 Components of the Umwelt Editor . . . 71
5.3.2 Default Specifications and Heuristics . . . 74
5.3.3 Design Rationale . . . 79
5.4 The Umwelt Viewer . . . 81
5.4.1 Multi-Modal Data Representations . . . 81
5.4.2 Coordinating Modalities with Linked Interactions . . . 83
5.4.3 Design Rationale . . . 83
5.5 Evaluation: Example Gallery . . . 85
5.6 Evaluation: User Study . . . 85
5.6.1 Quantitative Results . . . 87
5.6.2 Qualitative Results: Multi-Modal Viewer . . . 90
5.6.3 Qualitative Results: Structured Editor . . . 92
5.7 Discussion and Future Work . . . 94
5.7.1 Designing Natively Non-Visual Data Representations . . . 94
5.7.2 Interdependence and Relational Dimensions of Access . . . 95

6 Discussion and Future Work 97
6.1 Participation and Power in Accessible Data Analysis . . . 97
6.2 Future Work . . . 97

III Empowered Non-Participation in Data Collection 99

7 Introduction 101

8 Designing and Evaluating Research Ethics Systems 103
8.1 Introduction . . . 103
8.2 Debriefing Participants and Opting Out of Social / Behavioral Research . . . 104
8.3 Bartleby: Research Debriefing System . . . 106
8.3.1 Design Values . . . 106
8.3.2 Debriefing User Experience . . . 108
8.3.3 Researcher User Experience . . . 109
8.3.4 Supporting Multiple Studies . . . 110
8.3.5 Data Removal Procedure . . . 110
8.4 Field Deployment of the Debriefing System . . . 110
8.4.1 Debriefing an Observational Study on Twitter . . . 111
8.4.2 Debriefing a Field Experiment in a Reddit Community . . . 112
8.5 Discussion . . . 113
8.5.1 Procedural and Substantive Theories in Research Ethics . . . 113
8.5.2 Procedural Ethics Evaluation of Bartleby . . . 115
8.5.3 Substantive Ethics Evaluation of Bartleby . . . 116
8.5.4 Passive Non-Participation, Sovereignty, and Non-Alienation . . . 119
8.5.5 Debriefing and Spam Filters . . . 120
8.6 Conclusion . . . 121

9 Data Refusal as Design 123
9.1 Introduction . . . 123
9.2 Refusal as Design . . . 125
9.3 Methods: Design Dimensions in Computer Science . . . 126
9.4 Four Facets of Refusal . . . 127
9.4.1 Autonomy: Individual and Collective . . . 127
9.4.2 Time: Reactive or Proactive . . . 130
9.4.3 Power: Foreclosing Possibility or Creating Possibility . . . 133
9.4.4 Cost: Accepting Cost or Reducing / Redistributing Cost . . . 135
9.5 Discussion . . . 137
9.5.1 Uses of the Framework of Data Refusal from Below . . . 137
9.5.2 Lessons for Designers: Refusal as a Prompt for Design . . . 139
9.5.3 Working Toward Institutionalized Refusal . . . 139

10 Discussion and Future Work 143
10.1 Participation and Power in Ethical Data Collection 143
10.2 Future Work 143

IV Discussion 145

11 Discussion and Future Work 147
11.0.1 Summary of Thesis Contributions 147
11.0.2 Directions for Future Work 148

List of Figures

1.1 (a) An accessible visualization structure in the form of a tree and comprised of encoding entities. Solid magenta outlines indicate the location of the screen reader cursor. Solid blue arrows between labels indicate available next steps via keyboard navigability (up, down, left, right). (b) Three ways of navigating accessible visualization structures: structural, spatial, and targeted. 23
1.2 Example multi-modal representations constructed with Umwelt. Examples (e) and (f) demonstrate visual view composition via concatenated and layered charts. Examples (b) (e) (f) demonstrate multiple sonification tracks, where a user can choose which data field to play through. Examples (a) (e) (f) demonstrate textual structures with multiple sub-trees corresponding to visual and audio structures. 24
1.3 Bartleby's moral significance is best understood through both procedural and substantive lenses. Left: Bartleby's interface for opting out procedurally promotes individual autonomy. Right: Bartleby creates opportunities for participants to express their substantive values in a Reddit study. 25
1.4 Arnstein's ladder of citizen participation [17], which depicts a typology of eight levels of participation. The levels are arranged to convey an increasing extent of citizen power from bottom to top. 30
1.5 Casemajor et al.'s framework of participation and non-participation in digital media [56], which depicts states of intersecting active/passive intentionality and participation/non-participation. Arrows indicate possible transitions between states. 31
1.6 A figure with two subfigures depicting the two problem domains examined in this dissertation: (a) Accessible data analysis, which involves movement from disempowered non-participation to empowered participation; (b) Ethical data collection, which involves movement from disempowered participation to empowered non-participation. 32

4.1 (a) An accessible visualization structure in the form of a tree and comprised of encoding entities. Solid magenta outlines indicate the location of the screen reader cursor. Solid blue arrows between labels indicate available next steps via keyboard navigability (up, down, left, right). (b) Three ways of navigating accessible visualization structures: structural, spatial, and targeted. 50
4.2 Example structural and navigational schemes generated as part of the co-design process, and applied to diverse chart types. 55

5.1 An analyst's workflow in Umwelt: 1–6 illustrate an analyst's process of creating an initial multimodal data representation (shown in 7-9). 10 shows their initial exploration, before (11) making edits and then (12–14) continuing their analysis 69
5.2 The Umwelt interface. A) The data, visual, and audio tabs of the editor. B) The editor's fields tab, where users specify field definitions and encodings. C) The viewer, where users analyze data with interactive multimodal data representations. 72
5.3 A) Fragments of an internal declarative specification shown next to their corresponding Umwelt editor states. B) The output multimodal representation for that specification. 74
5.4 Prototypes of encoding- and field-oriented specifications of a scatterplot with concatenated sonification, illustrating the *role-inexpressiveness* [40] of field-oriented textual specification. Color-coded spans on the left side of each text prototype show the lines of code that pertain to each modality: green represents visual while blue represents audio. A) Encoding-oriented specification groups each modality into unit specifications. B) Field-oriented specification fragments each unit's encodings across the spec. 79
5.5 An example of linked interaction across modalities, driven by the textual modality. Navigating to a node in the textual structure emits a query predicate. The visualization reifies this predicate as a conditional encoding, and the sonification reifies it as a filter. 83
5.6 Example multi-modal representations created with Umwelt, expressing a range of key-value semantics. A) Two keys and one value. B) No keys and three values. Audio units represent the quantitative values' distribution in 2 dimensions. C) One key and two values. Audio units represent looking up each value with the same key. D) Concatenated visualization and sonification. E) Layered visualization with concatenated sonification. F) Three keys and one value. G) An alternate visual representation for the same key-value semantics as A. 86

8.1 Debriefing invitation messages sent to participants eligible for debriefing. The messages included a link to the debriefing login page for each study. 107
8.2 Experiment-specific login page for the debriefing software. People can log in with their social media account to verify their inclusion in the study and view debriefing information. 108
8.3 The components of the debriefing page. (A) Information about inclusion in the study and the purpose of the research. (B) Table showing what data was collected on the participant, with description of anonymization procedures. (C) Information on risks and benefits, with checkbox for opting out. (D) Details on following up about the study results. (E) Contact information for university review board. (F) Survey on improving the debriefing webpage. . 122

11.1 A diagram [263] conveying the conventional model of human-computer interaction as a self-contained feedback loop. 148

11.2 “How the computer sees us,” an illustration from a computing textbook [206] that depicts how a computer assumes the presence of a certain sensory-motor apparatus when modeling humans as information processors. 150

List of Tables

4.1 Rating scores for each prototype (Table, Multi-view, Targeted) on a five point Likert scale where 1 = Very Difficult (Very Unenjoyable) and 5 = Very Easy (Very Enjoyable). Median scores are shown in boldface, averages in brackets, standard deviations in parentheses. 58

5.1 Default specification heuristics based on a dataset's key and typings. T = temporal field, N = nominal field, Q = quantitative field. Each row represents a rule that matches a dataset's key and value tuples. The visualization, sonification, and textual description columns show default specifications for each rule, represented in YAML format for conciseness. 76
5.2 Participants' names, demographic information, and descriptions of their screen reader and data analysis experience. Participants are referenced by name with their consent. 88
5.3 Rating scores for the viewer and editor on a five-point Likert scale where 1 = Very Difficult (Very Unenjoyable) and 5 = Very Easy (Very Enjoyable). Median scores are shown in **bold**, averages in brackets [], standard deviations in parentheses (). 89

8.1 Accounts Participating in Twitter Debriefing 111
8.2 Accounts Participating in Reddit Debriefing 112
8.3 Direct Message Replies to Reddit Debriefing Invitation 112

Part I

Introduction

Chapter 1

Introduction

For decades, scholars in the field of human-computer interaction (HCI) have envisioned technology-mediated social participation as a promising frontier for research and design. For example, in 2006, Susanne Bødker identified a shift in HCI toward ideas of participation and sharing situated in the technology use of everyday life — a shift she termed "third-wave HCI" [53]. Soon after in 2011, Ben Shneiderman identified technology-mediated social participation as a central theme of the "next 25 years of HCI challenges" [245]. He envisioned a world in which technology was integrated into a variety of important social processes, including "national priorities and life-critical activities such as disaster response, community safety, health/wellness, energy sustainability, and environmental protection" [245]. Pervasive in these discussions was a sense of promise — the idea that technology can be designed better to facilitate human creativity, flourishing, and connection.

However, despite decades of research, there remain many challenges in the area of technology-mediated social participation that have prevented researchers' visionary ideas from coming to full fruition. To make progress on the shared goals of the field, it is important to understand their most salient obstacles. I argue that technology-mediated participation faces two core design challenges that lead to undesirable outcomes: exclusion and inclusion.

First, technology can be designed in ways that exclude people from important social processes. This idea is clearest in the domain of accessibility. Blind and low-vision (BLV) people are typically excluded from using websites that are not designed for the screen reader assistive technology that they use to narrate visual content as text-to-speech. Consequently, BLV people face barriers to their ability to participate in conversations and make decisions about about socially important issues. For example, researchers have found that about 97% of state websites in the U.S. that provided information about the COVID-19 pandemic had significant accessibility problems in 2021 [276]. It is important to note that the challenge here is not simply that designers are ignorant of the need for accessibility. For instance, companies have attempted to develop AI-powered tools to make web content more accessible. Yet, blind and low-vision users and disability advocates have pointed out that "the fix can be worse than the flaws" when these tools make websites even less usable [195]. When designers (and the broader HCI community) lack knowledge on how to meaningfully address accessibility issues, imperfect solutions can be worse than doing nothing. The challenge, especially for researchers, is to discover how to meaningfully address accessibility through design.

Second, technology can also include people in social processes that they find coercive or disagreeable. This is commonplace in the domain of data ethics. The public has become unfortunately accustomed to data scandals like Cambridge Analytica, in which many Facebook users' data was repurposed for political influence campaigns — sometimes without any action on their part [54]. Similarly, users' data is often repurposed and used to train facial recognition systems that are frequently sold to the police to power state surveillance [205], or directly sold by data brokers which can power discriminatory credit scoring or stalking and harassment [207]. As a consequence of these commonplace practices, people who do not agree with the repurposing of their data for certain uses do not have the power to influence the process or withdraw their participation. Here, challenges abound in addressing people's fundamental lack of agency over their data.

What's at stake in conversations about harmful forms of exclusion and inclusion? At worst, designers might create systems that exclude people from useful means of human interaction, reinforcing inequality; or, that conscript people into contributing to systems they don't agree with, without their consent. At best, designers might realize the promises of technology-mediated social participation by creating systems that expand opportunities for access to information and creativity in unprecedented ways, and empower people to shape large-scale uses of technology that affect their social and political lives.

1.1 Thesis Contributions

This thesis addresses the core challenges of exclusion and inclusion in technology-mediated social participation by contributing software systems and design frameworks in two motivating problem domains: *accessible data analysis* and *ethical data collection*. My research approaches designing empowering systems from two directions: (1) creating systems that broaden participation in empowering creative processes and (2) protecting users' individual and collective agency to reshape disempowering processes.

1.1.1 Accessible Data Analysis

Multi-modal interfaces that incorporate visualization, textual description, and sonification have become an important means of making data analysis accessible, but existing approaches often prioritize the visual modality — a critical limitation when it comes to BLV people's equal participation in data analysis. Tools often require an existing visualization to convert into an accessible format, making it difficult for a BLV user to independently create data representations or engage in rapid prototyping. Further, designers often approach non-visual representations using direct sensory re-mappings of existing visual chart forms, but this approach fails to leverage modality-specific, non-visual affordances.

To advance the design of natively non-visual data representations, I developed a framework of design dimensions for screen reader experiences (Chapter 4). Screen readers, an assistive technology that transforms text and visual media into speech, fundamentally redefine interaction for visualization. Because screen readers narrate elements one-at-a-time, they explicitly linearize reading a visualization. In contrast to sighted readers who can selectively shift their attention between specific elements and overall patterns in a visualization, screen reader users are limited to the linear steps made available by the visualization author and must remember the screen reader's prior output. To capture these modality differences, I worked with a blind co-designer to identify three design dimensions for enabling an expressive space of experiences (Figure 1.1): structure, or how the different elements of a chart should be organized for a screen reader to traverse; navigation, which describes the operations a user may perform to move through this structure; and, description, which specifies the semantic content, composition, and verbosity of text conveyed at each step. Prior to this work, best practices for accessible visualization involved providing textual description in a summary paragraph and providing a data table that a screen reader user could read row-by-row. Systems designed with these design dimensions in mind, such as my open-source library Olli, instead offer keyboard-navigable structures of multiple descriptions at varying levels of detail, affording user agency over the reading order and level of information granularity in each description. This work, which won Best Paper Honorable Mention at EuroVis 2022, has encouraged researchers and designers to reconceptualize textual description as a fundamentally interactive modality.

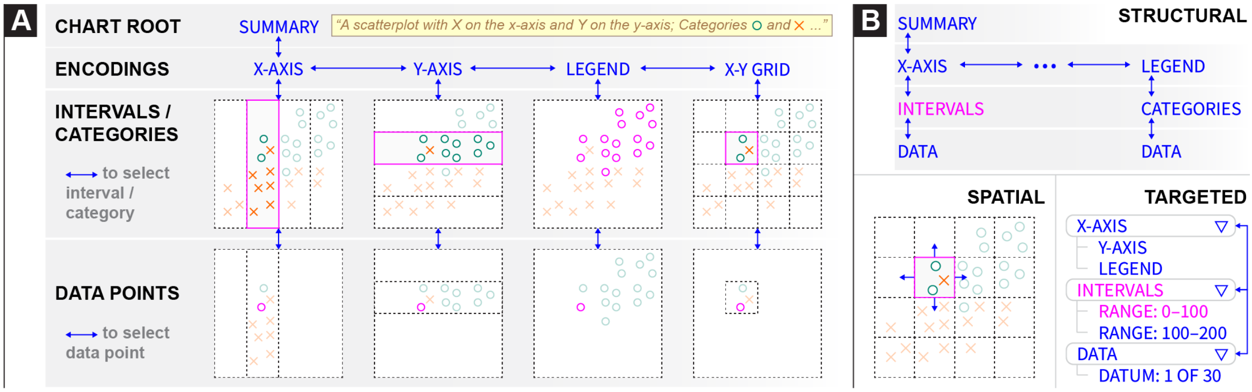
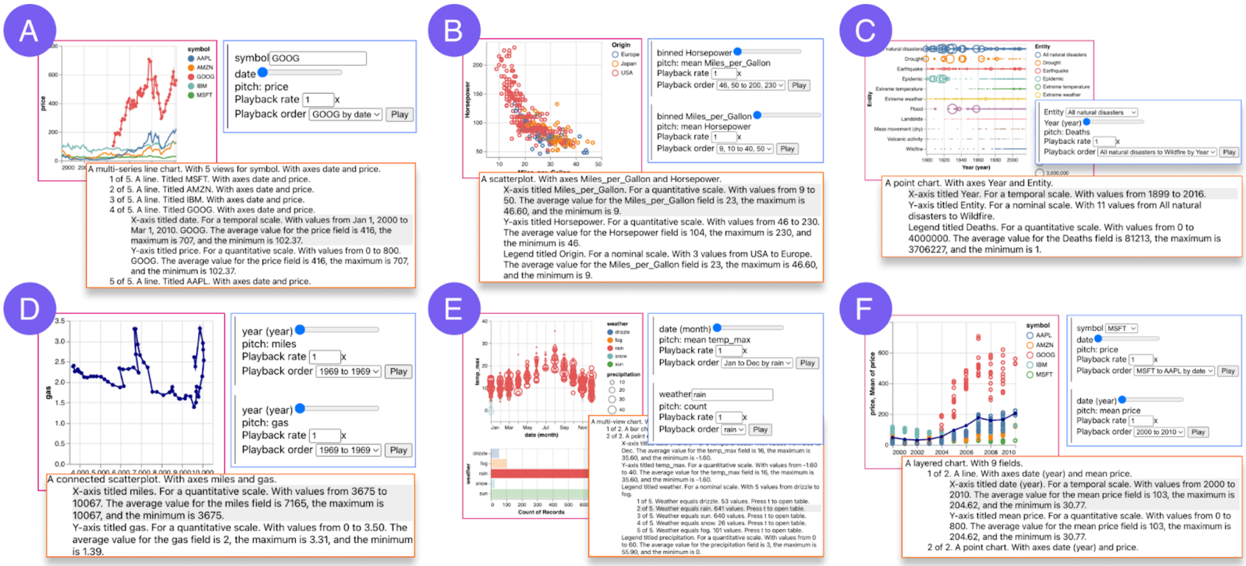


Figure 1.1: *(a) An accessible visualization structure in the form of a tree and comprised of encoding entities. Solid magenta outlines indicate the location of the screen reader cursor. Solid blue arrows between labels indicate available next steps via keyboard navigability (up, down, left, right). (b) Three ways of navigating accessible visualization structures: structural, spatial, and targeted.*

Guided by this conceptual approach to non-visual modalities, I created Umwelt: an accessible editor and viewer for interactive, multi-modal data representations incorporating visualization, textual structures, and sonification (Chapter 5). Umwelt presents abstractions and user experiences that treat all modalities equally. Prior systems that require a visualization to derive other modalities typically use the visualization's structure to control the other representations. Umwelt avoids placing undue emphasis on the visualization by deriving an independent structure and navigation for each modality directly from an abstract data model (Figure 1.2). Each structure is linked together interactively, enabling a user to select a subset of data in one modality and use it to filter another, or navigate between modalities while maintaining their selection. For example, a user can play through a sonification to understand the data's overall trend, then jump to the textual description to contextualize what they heard in concrete numeric values. Then, navigating down a level of detail in the textual structure, they can use their position in the text to filter the sonification, zooming in on a subset of the data to listen in more detail. Because interaction state is shared across modalities, a user can conduct freeform exploration by switching between complementary representations.



symbol GOOG
date
pitch: price
Playback rate 1 x
Playback order GOOG by date

A multi-series line chart. With 5 views for symbol. With axes date and price.
1 of 5. A line. Titled MSFT. With axes date and price.
2 of 5. A line. Titled AMZN. With axes date and price.
3 of 5. A line. Titled IBM. With axes date and price.
4 of 5. A line. Titled GOOG. With axes date and price.
X-axis titled date. For a temporal scale. With values from Jan 1, 2000 to Mar 1, 2010. GOOG. The average value for the price field is 416, the maximum is 707, and the minimum is 102.37.
Y-axis titled price. For a quantitative scale. With values from 0 to 800. GOOG. The average value for the price field is 416, the maximum is 707, and the minimum is 102.37.
5 of 5. A line. Titled AAPL. With axes date and price.

binned Horsepower
pitch: mean Miles_per_Gallon
Playback rate 1 x
Playback order 46, 50 to 200, 230

binned Miles_per_Gallon
pitch: mean Horsepower
Playback rate 1 x
Playback order 9, 10 to 40, 50

A scatterplot. With axes Miles_per_Gallon and Horsepower.
X-axis titled Miles_per_Gallon. For a quantitative scale. With values from 9 to 50. The average value for the Miles_per_Gallon field is 23, the maximum is 46.60, and the minimum is 9.
Y-axis titled Horsepower. For a quantitative scale. With values from 46 to 230. The average value for the Horsepower field is 104, the maximum is 230, and the minimum is 46.
Legend titled Origin. For a nominal scale. With 3 values from USA to Europe. The average value for the Miles_per_Gallon field is 23, the maximum is 46.60, and the minimum is 9.

Entity All natural disasters
Year (year)
pitch: Deaths
Playback rate 1 x
Playback order All natural disasters to Wildfire by Year

A point chart. With axes Year and Entity.
X-axis titled Year. For a temporal scale. With values from 1899 to 2016.
Y-axis titled Entity. For a nominal scale. With 11 values from All natural disasters to Wildfire.
Legend titled Deaths. For a quantitative scale. With values from 0 to 4000000. The average value for the Deaths field is 81213, the maximum is 3706227, and the minimum is 1.

year (year)
pitch: miles
Playback rate 1 x
Playback order 1969 to 1969

year (year)
pitch: gas
Playback rate 1 x
Playback order 1969 to 1969

A connected scatterplot. With axes miles and gas.
X-axis titled miles. For a quantitative scale. With values from 3675 to 10067. The average value for the miles field is 7165, the maximum is 10067, and the minimum is 3675.
Y-axis titled gas. For a quantitative scale. With values from 0 to 3.50. The average value for the gas field is 2, the maximum is 3.31, and the minimum is 1.39.

date (month)
pitch: mean temp_max
Playback rate 1 x
Playback order Jan to Dec by rain

weather rain
pitch: count
Playback rate 1 x
Playback order rain

A multi-view chart. W
1 of 2. A bar ch
2 of 2. A point c
X-axis titl
Dec. The average value for the temp_max field is 16, the maximum is 35.60, and the minimum is -1.60.
Y-axis titled temp_max. For a quantitative scale. With values from -1.60 to 40. The average value for the temp_max field is 16, the maximum is 35.60, and the minimum is -1.60.
Legend titled weather. For a nominal scale. With 5 values from drizzle to fog.
1 of 5. Weather equals drizzle. 53 values. Press t to open table.
2 of 5. Weather equals rain. 641 values. Press t to open table.
3 of 5. Weather equals sun. 640 values. Press t to open table.
4 of 5. Weather equals snow. 26 values. Press t to open table.
5 of 5. Weather equals fog. 101 values. Press t to open table.
Legend titled precipitation. For a quantitative scale. With values from 0 to 60. The average value for the precipitation field is 3, the maximum is 55.90, and the minimum is 0.

symbol MSFT
date
pitch: price
Playback rate 1 x
Playback order MSFT to AAPL by date

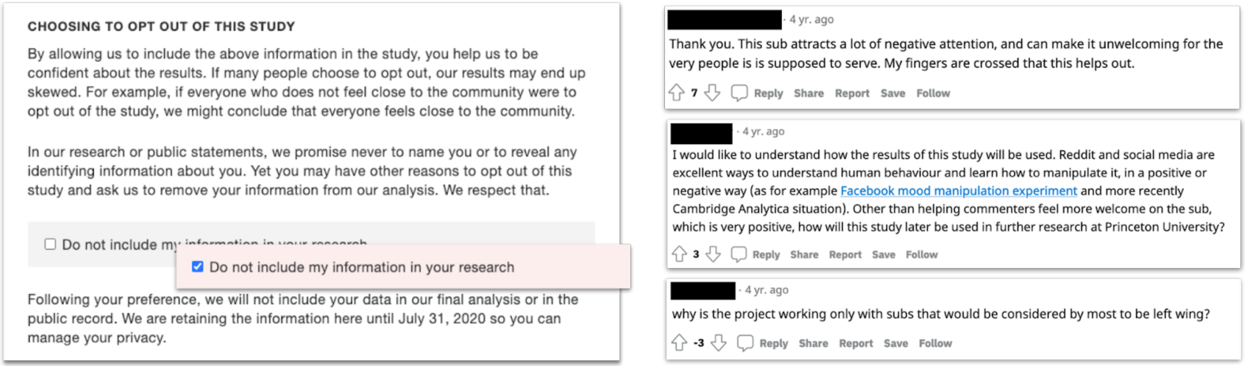
date (year)
pitch: mean price
Playback rate 1 x
Playback order 2000 to 2010

A layered chart. With 9 fields.
1 of 2. A line. With axes date (year) and mean price.
X-axis titled date (year). For a temporal scale. With values from 2000 to 2010. The average value for the mean price field is 103, the maximum is 204.62, and the minimum is 30.77.
Y-axis titled mean price. For a quantitative scale. With values from 0 to 800. The average value for the mean price field is 103, the maximum is 204.62, and the minimum is 30.77.
2 of 2. A point chart. With axes date (year) and price.

Figure 1.2: *Example multi-modal representations constructed with Umwelt. Examples (e) and (f) demonstrate visual view composition via concatenated and layered charts. Examples (b) (e) (f) demonstrate multiple sonification tracks, where a user can choose which data field to play through. Examples (a) (e) (f) demonstrate textual structures with multiple sub-trees corresponding to visual and audio structures.*

1.1.2 Ethical Data Collection

When social and behavioral research primarily happened in lab settings, people were generally always aware of their status as research participants; but as this research moved online to social media platforms, people are now included in data collection by default — frequently without their knowledge. This newfound possibility of passive participation in research has required researchers to apply ethics procedures in new ways. For instance, online studies reduce the viability of prior informed consent: a scientist studying online speech during a political debate cannot predict in advance who will participate in conversations during the event. I developed Bartleby, a system for automatically notifying data subjects after data collection and directing them to a web interface where they have a chance to opt out by deleting their data (Chapter 8). Bartleby takes a procedure known as debriefing, historically used in deception-based lab studies, and applies it in a new context online (Figure 1.3A). The design and use of software to manage the ethics of data collection in novel situations is clearly morally consequential; but how do researchers evaluate the moral outcomes produced by automating ethics? I developed a framework offering two lenses for understanding the moral significance of Bartleby and other data ethics software — procedural ethics and substantive ethics — and elaborated their implications through two field deployments of Bartleby following research data collection on Twitter and Reddit. Researchers applying a procedural lens consider data collection to be ethical when the right procedures (e.g. informed consent, debriefing, IRB review) are implemented to protect individual autonomy. Bartleby contributes procedurally by offering participants the choice to opt out; because this choice is not present in the counterfactual case where Bartleby is not used, Bartleby promotes the exercise of participant autonomy. Researchers applying a substantive lens consider data collection to be ethical when its outcomes align with participants' values. This perspective accounts for the idea that power dynamics can lead people to consent to things that cause them harm. Because individuals' subjective values are not always known to researchers or legible to software, Bartleby's core strengths are procedural. However, Bartleby contributes substantively by opening a communication channel between participants and researchers, enabling conversations that do not otherwise happen when people are not notified. In these conversations, participants might voice their values, ask questions, or challenge researchers' legitimacy (Figure 1.3B).



CHOOSING TO OPT OUT OF THIS STUDY

By allowing us to include the above information in the study, you help us to be confident about the results. If many people choose to opt out, our results may end up skewed. For example, if everyone who does not feel close to the community were to opt out of the study, we might conclude that everyone feels close to the community.

In our research or public statements, we promise never to name you or to reveal any identifying information about you. Yet you may have other reasons to opt out of this study and ask us to remove your information from our analysis. We respect that.

☐ Do not include my information in your research

☑ Do not include my information in your research

Following your preference, we will not include your data in our final analysis or in the public record. We are retaining the information here until July 31, 2020 so you can manage your privacy.

[redacted] · 4 yr. ago

Thank you. This sub attracts a lot of negative attention, and can make it unwelcoming for the very people is is supposed to serve. My fingers are crossed that this helps out.

[redacted] · 4 yr. ago

I would like to understand how the results of this study will be used. Reddit and social media are excellent ways to understand human behaviour and learn how to manipulate it, in a positive or negative way (as for example Facebook mood manipulation experiment and more recently Cambridge Analytica situation). Other than helping commenters feel more welcome on the sub, which is very positive, how will this study later be used in further research at Princeton University?

[redacted] · 4 yr. ago

why is the project working only with subs that would be considered by most to be left wing?

Figure 1.3: *Bartleby's moral significance is best understood through both procedural and substantive lenses. Left: Bartleby's interface for opting out procedurally promotes individual autonomy. Right: Bartleby creates opportunities for participants to express their substantive values in a Reddit study.*

Computer scientists have historically lacked a framework for understanding opt-outs as an expression of user voice and agency; instead, computer scientists often assume that non-users will become future users with improvements to technology design. A consequence of this assumption is that non-users are considered either passive resources to be mined for design ideas, or problems to be solved through design. My work puts forth an alternate view, arguing that refusal can be understood as a generative act of socio-technical design (Chapter 9). Because people refuse to change an existing situation into a preferred one, refusal in the form of active non-participation in a system is actually an important way to influence a design process toward desired outcomes. Crucially, the idea of refusal as design empowers computer scientists to apply the methods of design to understand refusal. To that end, I developed four design dimensions for data refusal: autonomy, time, power, and cost. These dimensions help researchers understand seemingly-disparate actions — from opt outs, to collective data management, to the creation of new policy and institutions — through the shared vocabulary of refusal. Though refusal might initially seem to be the opposite of generative design, computer scientists equipped with an understanding of refusal as design can see refusals as a healthy and necessary part of a socio-technical design process that can inspire directions for systems-level change.

1.1.3 Participation and Power: A Unifying Conceptual Lens

Though these two domains of accessible data analysis and ethical data collection are seemingly separate, the idea of *participation and power* provides an underlying conceptual grounding that unites my approach to both sets of problems. Conventional wisdom is that increased participation in social processes — from design to democracy — is empowering, and that making these processes inclusive of more diverse groups of people leads to better individual and social outcomes. However, participation is a double-edged sword. The proliferation of computing throughout all aspects of society has also resulted in systems that disempower people through involuntary participation, resulting in scandals about data misuse. As this dissertation will demonstrate, my research in accessible data analysis and data ethics approaches designing empowering systems from two directions: (1) creating systems that broaden participation in empowering creative processes and (2) protecting users' individual and collective agency to reshape disempowering processes.

1.2 Interpreting the Contributions

This dissertation, like many others in HCI, makes contributions in the form of software systems, interface designs, and design frameworks. But given the interdisciplinary scope of the problems this work addresses, how should a reader interpret these contributions and understand their implications? In this section, I provide the necessary context needed to understand which literatures I am drawing from in this work, and introduce key concepts for interpreting the significance of the work.

1.2.1 Positionality

To understand the contributions of my work, it is helpful to understand my background as a researcher and the disciplines I am working across. In HCI, researchers have included positionality statements as a way of providing this context to readers. Writing a positionality statement is a practice of what Bardzell calls *researcher self-disclosure*, or being transparent about "the researcher's position in the world, her or his goals, as well as the researcher's position in her or his intellectual and, to an appropriate extent, political beliefs" [29]. Bardzell and others, such as Erete et al., have advocated for self-disclosure as a research methodology because "self-disclosing information about aspects of our identity and positionality, and potential impacts to the design research process, also helps improve the transparency and understandability of our research" [92]. Here, my goal is to provide sufficient context to readers for understanding which disciplinary perspectives I have drawn from in my work, what key ideas from those disciplines I am employing, and how I situate my use of those ideas in existing scholarly conversations.

I am trained as a computer scientist and visual artist, and the work in this thesis engages with design as both a method and community of practice in HCI. In other words: on one hand, some of my contributions use the act of designing software systems as a way to advance knowledge about behavior in socio-technical systems (Chapter 8, Chapter 5) — an intellectual move inspired by Simon's idea of design as a "science of the artificial" [246]. On the other hand, my other contributions advance the practice of design by developing conceptual frameworks that help designers think systematically about how to approach a problem that people face (Chapter 4, Chapter 9). I contextualize the motivation and goals of my work as *use-inspired basic research* [81], which is oriented toward advancing fundamental knowledge while grounded in the need to solve practical problems facing people today.

Due to my interdisciplinary background, my work frequently draws from disciplines outside of computer science that have shaped my thinking. These disciplines include media studies and science and technology studies (STS), which are fields that think about the development of technology together with the interaction between the social forces that shape technology, and how technology shapes society in return. My work is also shaped by intellectual traditions in feminist and disability scholarship, which both include ways of thinking about structures of power and oppression in society.

1.2.2 Key concepts

Because I am drawing from interdisciplinary scholarship, the meaning of important terms can be contingent on their use in a particular discipline (and can be distinct from their meaning in colloquial use). In this section, I summarize the scholarly conversations from which I am deriving my working definitions of *power* and *participation*.

Power

In my work on accessible data analysis, my understanding of power is informed by *disability studies*. In my work on ethical data collection, my understanding of power is informed by *feminist standpoint theory*. Here, I summarize these two perspectives on power.

Power in disability studies. My work in accessible data analysis draws on research and advocacy in disability studies, a field in which scholars seek to understand the interaction between people's bodies and their environments, and the value system (known as *ableism*) that confers social and material advantages (power) to certain kinds of bodies over others. The idea of power in disability studies informs the way scholars interpret the experiences of disabled people and the way they think about addressing the challenges of access and inclusion.

Disability scholars have drawn contrasts between two common models of disability. The first is the medical model, which treats disability as something that emerges when a body's abilities diverge from the norm. Scholars have critiqued the medical model because it "frames atypical bodies and minds as deviant, pathological, and defective" [161]. As a result, this framing locates disability as an individual problem and suggests interventions that are primarily medical.

In contrast, the social model of disability argues that disability emerges from social and environmental barriers that are designed to privilege certain kinds of bodies over others [161]. In this framing, addressing disability means changing environments and architectures to make them more inclusive of a multitude of body abilities, so that people can achieve the same things despite bodily differences.

Building on the social model, Alison Kafer introduces a political/relational model that acknowledges both medical and social perspectives, while emphasizing that both embodied and social factors must be understood in a cultural and political context [161]. For example, the political/relational model acknowledges that impairments and chronic pain are important and may not be solved by purely social approaches, but also understands that political questions like structural barriers to medical care exist beyond an individual medical frame.

Disability movements have used these concepts to articulate goals for advocacy and empowerment. For example, much work in disability advocacy and accessibility has focused on the issue that disabled people are often dependent on non-disabled people for assistance to participate in society [37]. Advocates have placed emphasis on the idea of independence as something that supports the power and autonomy of disabled people, and full access and participation in society.

However, in addition to independence, disability scholars have also advanced interdependence as a complementary idea [37]. Interdependence acknowledges that all people constantly depend on others, and so a focus on relationships is necessary to understand how access is socially created in practice.

Power in feminist standpoint theory. Feminist theorists have argued that people's ways of understanding the world depend on their position within social structures [123]. *Standpoint* is an important concept for describing one's social position. Standpoint is relational in that it is described relative to an existing hierarchy of power. The language of "above" and "below" are sometimes used to "emphasize and articulate the tensions between the wants, needs, and knowledge of unequally empowered stakeholders" [282]. Feminist scholar Donna Haraway, for example, argues that scientists should value knowledge "from below," which begins from the experiences of those "subjugated" under hierarchies of power [122].

For standpoint theorists, power is frequently tied to the idea of knowledge. Scholars have argued that multinational corporations' domination over digital ecosystems grants firms "direct power over political, economic and cultural domains of life," an arrangement scholars have described as digital colonialism [73, 173]. Their power — or ability to intervene in these domains of life — frequently stems from their control over large-scale systems of surveillance and capture, which grant access to data and information. The power to act is a result of the power to create and define knowledge about others.

As HCI scholars of tactics from below have noted, "above" and "below" are oversimplifications — researchers are often simultaneously situated "above" and "below" in different contexts [282]. However, standpoint is an important concept for thinking about power because it equips scholars with language to explicitly identify unequal social structures and talk about relative position within them.

Participation

In this thesis, my use of *participation* draws on existing conversations in digital media studies that explore participation and non-participation as forms of mediated political action within digital systems. In this section, I first summarize the scholarly conversation about participation in which I am engaging. Then, I articulate the working model of participation that the rest of this thesis will adopt.

A common starting point in conversations about participation is Sherry Arnstein's "ladder of citizen participation," a foundational conceptual model in participatory democracy. Arnstein referred to citizen participation as "a categorical term for citizen power" [17]. In other words, participation is defined as citizen control over decision-making processes in a democratic society — or, the power to influence decision-making. The ladder is a diagram (Figure 1.4) depicting a typology of eight levels of participation, which are ordered to convey an increasing level of citizen participation (and power) from bottom to top. The bottom rungs represent non-participation, which represent the absence of "citizen power" — an undesirable state of affairs. Arnstein developed this model in the context of conversations in urban planning, as a critique of practices that appeared to enable participation without actually delegating power to participants. An important contribution of this work was to demonstrate that participation and non-participation are not binary states — instead, there are gradations to participation.

However, this conventional account of participation has been challenged in by scholars like Casemajor et al., who saw non-participation as something more than passivity or deprivation. Instead, scholars have proposed alternate conceptual models that more centrally consider non-participation as a form of political action. Casemajor et al. advances a framework of participation with two orthogonal dimensions: a dimension of *active* vs *passive*, and a dimension of *participation* vs *non-participation* [56]. Here, active and passive describe the presence or absence of political intention applied to (non-)participation to distinguish cases where "participatory strategies lack or devolve their political character, and others where non-participatory practices become politically willful" [56]. In this framework, people inevitably shift between combinations of intention and (non-)participation as they reconfigure their relationship to digital media (Figure 1.5). Casemajor et al.'s key insight is to provide a framework that highlights non-participation as a meaningful expression of agency, and to recognize that people's engagement with socio-technical systems moves between participation and non-participation over time.

1.2.3 Designing for Participation and Power

How does HCI and design relate to ideas of participation and power? As Casemajor et al. have argued in their model of participation and non-participation as mediated political action, a person's relationship to participation and power are not static. Rather, they can shift their engagement in various socio-technical systems via the active and passive dimensions of mediated participation.

Building on these conversations, this thesis formulates the relationship between participation and power as two orthogonal dimensions that I use to interpret my contributions in the two problem domains of accessible data analysis and ethical data collection (Figure 1.6).

In accessible data analysis, I observe that most BLV users are in a state of disempowered non-participation due to the design of their software tools — they are fully excluded from conversations about data due to a lack of knowledge in the design community about how to design effective non-visual data representations and interactions. In this domain, my work aims to help people move toward a state of empowered participation (Figure 1.6A). I accomplish this by developing a design framework for rich screen reader user experiences, and introducing software tools that prioritize self-guided, independent data exploration without

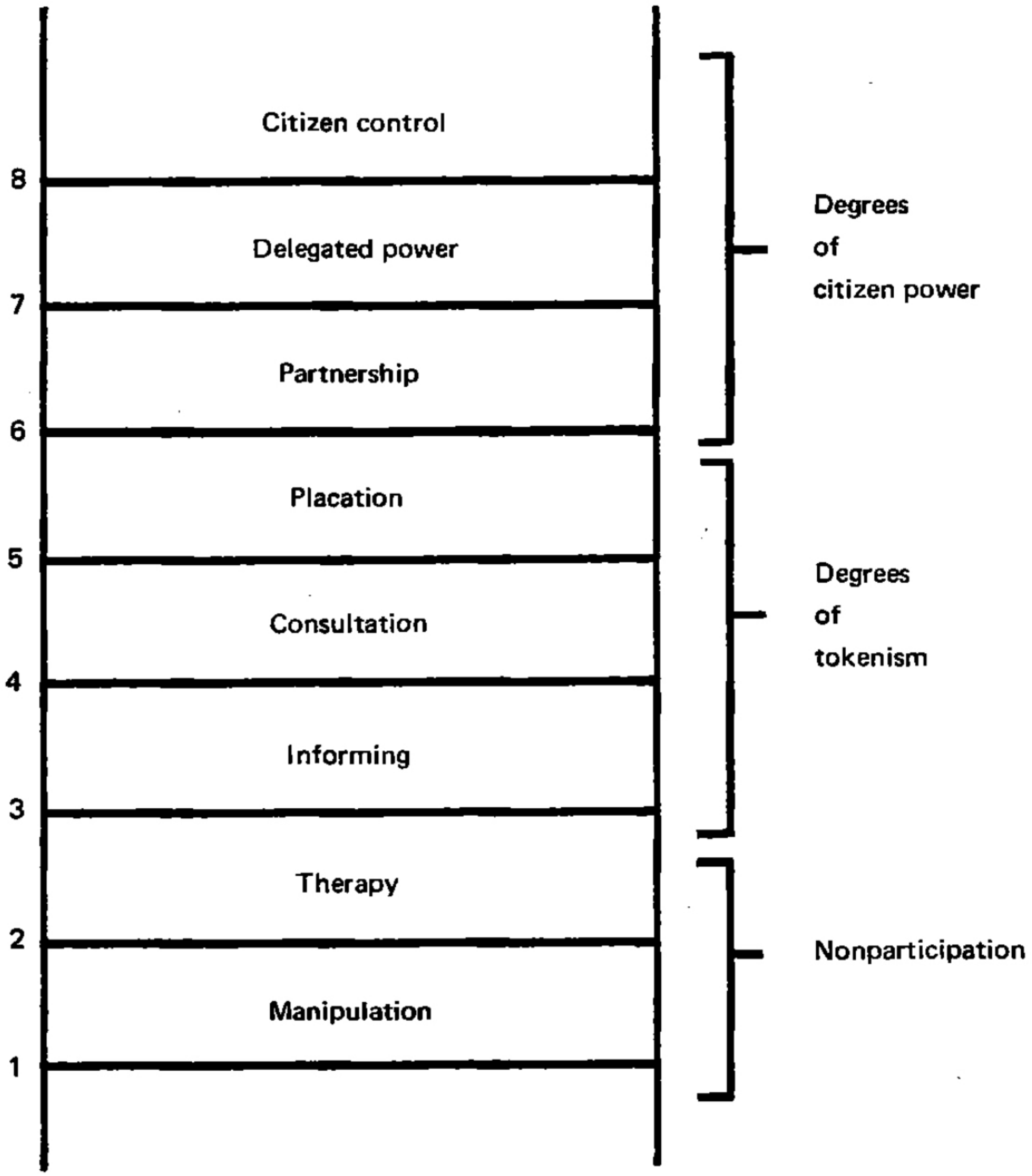


Figure 1.4: *Arnstein's ladder of citizen participation [17], which depicts a typology of eight levels of participation. The levels are arranged to convey an increasing extent of citizen power from bottom to top.*

relying on sighted assistance.

In ethical data collection, I observe that people are often in a state of disempowered participation for reasons relevant to design — they are included in large-scale data collection processes in which software tools do not respect their individual autonomy, and are unable to

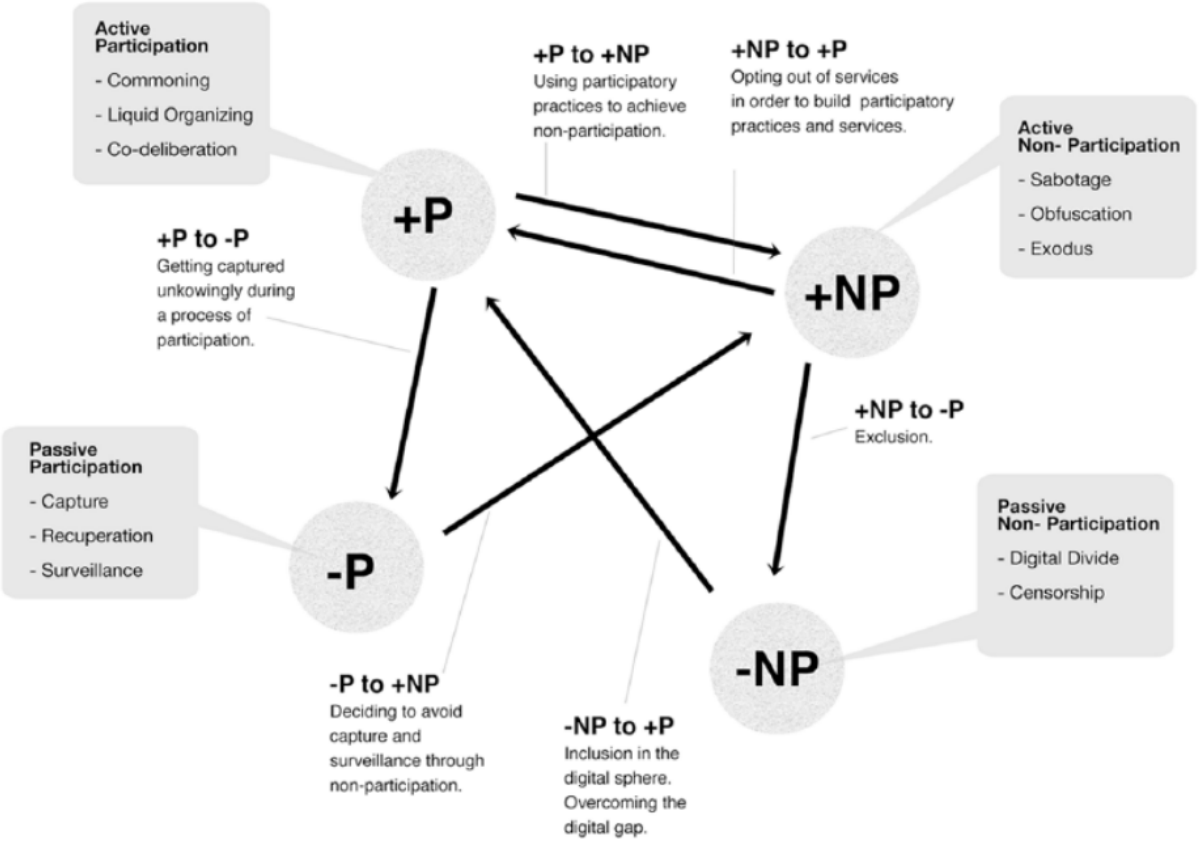


Figure 1.5: *Casemajor et al.'s framework of participation and non-participation in digital media [56], which depicts states of intersecting active/passive intentionality and participation/non-participation. Arrows indicate possible transitions between states.*

exert structural influence that would shift the goals of the data collectors toward alternatives they find more agreeable. In this domain, my work aims to help move people toward a state of empowered non-participation (Figure 1.6B). I introduce a novel software system for managing opt-out in large-scale social media research, apply interdiscplinary theory to help researchers interpret the success of such systems, and introduce a conceptual framework for understanding refusal in socio-technical systems as a form of design.

1.3 Prior Publications and Authorship

The work presented in this dissertation includes research and thinking that developed through collaborations with my advisor Arvind Satyanarayan and colleagues in the MIT Visualization Group and Cornell Citizens and Technology Lab. It also includes work that has been previously published in peer reviewed journals and conference proceedings. Below, I have included a list of previous publications that the material of this dissertation draws from.

- Jonathan Zong, Crystal Lee, Alan Lundgard, JiWoong Jang, Daniel Hajas, and Arvind Satyanarayan. Rich Screen Reader Experiences for Accessible Data Visualization.

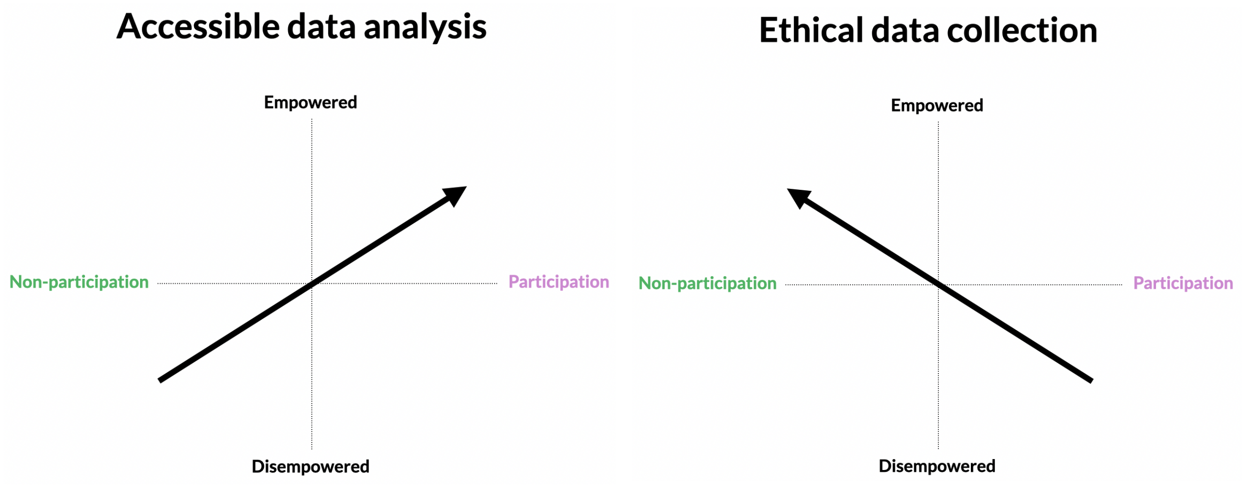


Figure 1.6: *A figure with two subfigures depicting the two problem domains examined in this dissertation: (a) Accessible data analysis, which involves movement from disempowered non-participation to empowered participation; (b) Ethical data collection, which involves movement from disempowered participation to empowered non-participation.*

Computer Graphics Forum, 2022. ISSN 1467-8659. doi:10.1111/cgf.14519. URL https://vis.csail.mit.edu/pubs/rich-screen-reader-vis-experiences/

- Jonathan Zong, Isabella Pedraza Pineros, Mengzhu Katie Chen, Daniel Hajas, and Arvind Satyanarayan. Umwelt: Accessible Structured Editing of Multimodal Data Representations, March 2024. URL http://arxiv.org/abs/2403.00106. Forthcoming in ACM CHI 2024
- Jonathan Zong and J. Nathan Matias. Bartleby: Procedural and Substantive Ethics in the Design of Research Ethics Systems. *Social Media + Society*, 8(1):205630512210770, January 2022. ISSN 2056-3051, 2056-3051. doi:10.1177/20563051221077021. URL http://journals.sagepub.com/doi/10.1177/20563051221077021
- Jonathan Zong and J. Nathan Matias. Data Refusal from Below: A Framework for Understanding, Evaluating, and Envisioning Refusal as Design. *ACM Journal on Responsible Computing*, 1(1):10:1–10:23, March 2024. doi:10.1145/3630107. URL https://dl.acm.org/doi/10.1145/3630107

Chapter 2

Related Work

2.1 Participation and Power in Design

Conversations about the practice of design inevitably reflect questions about participation and power. In this section, I summarize two lines of inquiry within these conversations about participation in design, and design as an exercise of power.

How does power shape who participates in design? Design is a creative act that takes on many forms. However, design as an institutionally recognized profession has typically excluded groups of people with less power in society. Sasha Costanza-Chock writes that "intersectional inequality systematically structures paid professional design work" [71]. Consequently, the outputs of design do not always reflect the values and priorities of the people affected by them.

Designers have sought to address this problem by broadening participation in design to include ideas of community power and accountability. Traditions like participatory design and co-design have sought to increase the amount of influence affected parties have over the design process [69]. Many of these approaches can be encapsulated in the disability activism slogan, "Nothing about us, without us" [71], which calls for equitable participation in design.

How do people use design to exercise power and reshape society? The idea of imagining speculative futures is an important frame for understanding the role of design in society. Dunne and Raby argue that design can be understood as a process of imagining and exploring alternate possibilities [86]. The process of imagining possibilities is not a straightforward one; it is shaped by dynamics of participation. Ruha Benjamin writes that "imagination is a field of struggle," arguing that power shapes what ideas become dominant in society, which in turns limits our perception of what is possible. When designers create artifacts that reflect dominant understandings of possible futures, they are potentially reinforcing structures of power. To reflect on the act of design as an exercise of power, designers have suggested approaches like "critical making" [80] and "critical technical practice" [11] to foreground the designer's role in exerting social influence through technical artifacts.

2.2 Accessibility as Empowered Participation

2.2.1 Screen Reader Accessibility for Blind and Low-Vision Users

Screen Reader Assistive Technology. A screen reader is an assistive technology that conveys digital text or images as synthesized speech or braille output. Screen readers are available as standalone third-party software or can be built-in features of desktop and mobile operating systems. A screen reader allows a user to navigate content linearly with input methods native to a given platform (e.g., touch on smartphones, mouse/keyboard input on desktop). Content authors must generate and attach alt text to their visual content like images or charts in order for them to be accessible to screen reader users. Functionality and user experience differs across platforms and screen readers. In this dissertation, however, I focus on interacting with web-based visualizations with the most widely used desktop screen readers (JAWS/NVDA for Windows, VoiceOver for Mac).

Web Accessibility Standards. In 2014, the World Wide Web Consortium (W3C) adopted the Web Accessibility Initiative's Accessible Rich Internet Applications protocol (WAI-ARIA) which introduced a range of semantically-meaningful HTML attributes to allow screen readers to better parse HTML elements [193]. In particular, these attributes allow a screen reader to convey the state of dynamic widgets (e.g., autocomplete is available for text entry), alert users to live content updates, and identify common sections of a web page for rapid navigation (e.g., banners or the main content). In 2018, the W3C published the WAI-ARIA Graphics Module [277] with additional attributes to support marking up structured graphics such as charts, maps, and diagrams. These attributes allow designers to annotate individual and groups of graphical elements as well as surface data values and labels for a screen reader to read aloud.

2.2.2 Accessible Data Representations

Accessible Visualization Design. In a recent survey, Kim et al. [167] describe the rich body of work that has explored multi-sensory approaches to visualization for multiple disabilities [31, 126, 162, 179, 289, 293]. Here, I focus on screen reader output native to web-based interfaces for blind users (namely via speech). Sharif et al. [239] find that many web-based charts are intentionally designed to cause screen readers to skip over them. For charts that a screen reader does detect, blind or low vision users nevertheless experience significant difficulties: these users spend 211% more time interacting with the charts and are 61% less accurate in extracting information compared to non-screen-reader users [239]. Despite the availability of ARIA, alt text and data tables remain the most commonly used and recommended methods for making web-based charts accessible to screen readers [60, 114, 279]. However, each of these three approaches comes with its own limitations. Static alt text requires blind readers to accept the author's interpretation of the data; by not affording exploratory and interactive modes, alt text robs readers of the necessary time and space to interpret the numbers for themselves [178]. Recent research also suggests that blind people have nuanced preferences for the kinds of visual semantic content conveyed via text [178, 213], and desire more interactive and exploratory representations of pictorial images [196]. Data tables, on the other hand, undo the benefits of abstraction that visualizations enable — they force readers to step sequentially through data values making it difficult to identify larger-scale patterns or trends, and do not leverage the structure inherent to web-based grammars of graphics [43, 232]. Finally, ARIA labels are not a panacea; even when they are used judiciously — a non-trivial task which often results in careless designs that cause screen readers to simply read out long sequences of numbers without any other identifiable information [227] — they present a fairly low expressive ceiling. The current ARIA specification does not afford rich and nuanced information-seeking opportunities equivalent to those available to sighted readers.

There has been some promising progress for improving support for accessibility within visualization toolkits, and vice-versa for improving native support for charts in screen reader technologies. For instance, Vega-Lite [232] and Highcharts [132] are beginning to provide ARIA support out-of-the-box. Apple's VoiceOver Data Comprehension feature [74] affords more granular screen reader navigation within the chart, beyond textual summaries and data tables, via four categories of selectable interactions for charts appearing in Apple's Stocks or Health apps. These interactions include *Describe Chart*, which describes properties of the chart's construction, such as its encodings, axis labels, and ranges; *Summarize Numerical Data*, which reports min and max data values, and summary statistics like mean and standard deviation; *Describe Data Series*, which reports the rate-of-change/growth of a curve, trends, and outliers; and *Play Audiograph*, which plays a tonal representation of the graph's ascending/descending trend over time [74]. While Apple's features are presently limited to single-line charts, SAS' Graphics Accelerator [10] supports a similar featureset (including sonification, textual descriptions, and data tables) but for a broader range of statistical charts including bar charts, box plots, contour plots, and scatter plot matrices. My work follows in the spirit of these tools but focuses on web-based visualizations rather than standalone- or platform-integrated software. I go beyond what ARIA supports today to enable high-level and fine-grained screen reader interactions, and hope that my work will help inform ongoing discussions on improving web accessibility standards (e.g., via an Accessibility Object Model [44]).

Multimodal Data Representations. Researchers and practitioners have explored a variety of approaches to data representations beyond visualization. A large body of systems has explored how multiple non-visual modalities can be used in concert. For instance, Apple's VoiceOver Data Comprehension feature on iOS [74] offers out-of-the-box support for making data accessible through verbal descriptions and sonification (or non-speech audio). Similarly, research systems have explored methods for combining tactile graphics with voice [22, 23], sonification with voice [140], haptics and sonification [94], and sonification and interactive question-answering. Among such multimodal systems, Chart Reader [262] is a particularly apt point of comparison to my work because it incorporates best practices in visualization, structured textual description, and sonification into a single analysis interface.

While these systems make important and necessary contributions to accessible visualization, they share a common assumption: they begin with a visual artifact and attempt to retarget visual affordances to non-visual modalities. For example, VoxLens and Chart Reader require a visualization specification as input to generate their non-visual representations. As a consequence, these systems are often unable to express data representations that do not have an analogous visualization. Chart Reader, for instance, can only express sonifications that directly correspond to the specific typology of chart types it supports. In contrast, I present work in chapter 5 that does not derive its non-visual representations from the visual specification. Instead, its three modalities are treated as equal outputs, all derived from an abstract data model that is shared across modalities.

2.2.3 Accessible Authoring Tools for Non-Visual Representations

In contrast to tools that convert an existing artifact into another representation, researchers have also explored authoring toolkits for multimodal representations. However, most existing toolkits correspond to a single non-visual modality. For example, Highcharts Sonification Studio [55] is an authoring tool for producing charts with sonification, and SVGPlott [89] is an authoring tool for tactile charts. Though these tools are designed to author non-visual data representations, they require a user to specify a visualization to convert into a non-visual form. Consequently, they suffer from expressiveness limitations. For example, SAS Graphics Accelerator [4] includes an authoring workflow that makes charts accessible via sonification and textual description, yet does not support sonification for many chart types.

Because these authoring environments require users to specify visualizations, they impose additional demands on BLV users. For instance, users must have a visual form in mind before creating a non-visual representation, and need an accessible way to verify the accuracy of their visual specification. In chapter 5, I address these concerns by allowing users to create representations in any order, and without requiring users to specify all three modalities. Instead, representations are authored independently, reducing the need to conceptualize all outputs in terms of the visual modality. For example, when a user loads a dataset, a textual structure will be generated describing the data in terms of its fields. They can directly specify a sonification by assigning audio encodings, without needing to specify them in terms of visual concepts like the x- or y-axis.

Of existing authoring toolkits, the closest point of comparison is PSST [214], which enables BLV users to create multimodal representations of streaming data that include sonification, spoken description, and physical laser-cut artifacts. Just as I propose in chapter 5, PSST does not require a visual specification. However, PSST differs from my work in terms of its level of abstraction; where I explore a higher-level workflow and abstractions, PSST exposes low-level abstractions such as event streams, handlers, and a dataflow graph.

2.3 Refusal as Empowered Non-Participation

2.3.1 Data Refusal

Scholars across anthropology, feminist theory, and bioethics have developed refusal as a concept for understanding how people reshape social and political relations. Anthropologist Carole McGranahan argues for the need to "recognize and theorize refusal" as a way that people "[stake] claims to the sociality that underlies all relationships, including political ones" [192]. Important conversations about refusal have been situated in conversations about citizenship and indigeneity [191, 247, 267]. One particular source of inspiration for my work in Part III is the work of sociologist Ruha Benjamin, who developed the idea of informed refusal within a family of conversations about the "relationships between biological and political experiments" [35]. For example, knowledge-making endeavors around stem cells and

DNA shape how institutions see people and how power is structured in society. Benjamin describes how "biocitizens" such as stem cell advocates contend over biological research by organizing to draw attention to inequalities and injustices in those systems of knowledge. Citing Rapp [217], Benjamin describes "moral pioneers" and "biodefectors" who work individually or collectively to opt out of studies, prevent labs from being built, or repatriate blood samples that geneticists used beyond agreed limits [36].

Drawing from these prior conversations, the idea of data refusal sheds light on ways that people contest the power of data in their lives, while also making the power of data collectors more legible [247]. Scholars have argued that multinational corporations' domination over digital ecosystems grants firms "direct power over political, economic and cultural domains of life," an arrangement scholars have described as digital colonialism [73, 173]. Like the historical subjects of colonialism, data subjects continue to design innovative ways to reshape their social relations with data collectors, even when power imbalances are great.

People often take up refusal in opposition to a variety of harms caused by data systems. Critical data scholar Anna Lauren Hoffmann has drawn a distinction between analyses of harm that focus on access to "rights, opportunities, and resources," and those that attend to social, discursive, and representational dimensions of data [136]. Privacy scholars Citron and Solove have proposed a legal typology of privacy harms that cover a broad range of harms to do with rights, opportunities, and resources [65]. Their framework includes types of harm that can easily be observed because they affect an individual's physical and psychological safety, reputation, or economic status. Crucially, their framework also includes collective harms including discrimination, risks to relationships, and constraints on collective autonomy. Scholars from a variety of disciplines also have written about important cases of symbolic and representational violence that are not well-captured by privacy frameworks. For example, Safiya Noble has written about the dignitary harm of Google search's stereotype-reinforcing results [203]. Sasha Costanza-Chock notes the cisnormativity of by airport scanning systems that inevitably subject trans, intersex, and gender non-conforming people to additional scrutiny [70]. Hoffmann explains that symbolic harms are also upheld by shaping discourse to normalize oppressive conditions and diffuse resistance [138].

When people refuse a data system, they are taking pragmatic actions toward changing those systems. For Ruha Benjamin, refusal is a form of agency that involves rejecting "the terms set by those who exercise authority in a given context" and which "may also extend beyond individual modes of opting out to collective forms of conscientious objection" [36]. While computer science and communication scholars often imagine lack of technology access or use as a form of deprivation [78, 223], refusal is "a politically significant action: one which opens up possibilities for power shifts, resistance to dominant political structures, and emancipation" [56].

2.3.2 Refusal and Standpoint

Refusal is a situated action that is shaped by the standpoint and imagination of the people who do the refusing [62]. Feminist standpoint theorists have argued that people's ways of understanding the world depend on their position within social structures [123]. The language of "above" and "below" are sometimes used to "emphasize and articulate the tensions between the wants, needs, and knowledge of unequally empowered stakeholders" [282]. Feminist scholar Donna Haraway, for example, argues that scientists should value knowledge "from below," which begins from the experiences of those "subjugated" under hierarchies of power [122]. This language of power clarifies an important distinction: because refusal happens within these systems of power, people's options for refusal will depend on their relationship to power.

A growing body of research considers actions of refusal by academics, designers, and industry professionals who are positioned to influence the creation of technology. For example, the authors of the Feminist Data Manifest-No commit in their capacity as academics to "entering ethically compromised spaces like the academy and industry not to imbricate ourselves into the hierarchies of power but to subvert, undermine, open, make possible" [62]. Other work has advanced efforts by tech designers to re-shape the values and practices of technology production [16], including by refusing to design [25, 32, 115, 287] or participate in industry events [112], or by engaging in labor organizing [72] and sabotage [197].

Another subset of work on data refusal focuses on the practices of those who are less proximal to the creation of new technology, yet are affected by its design and deployment. Recently, feminist scholars who also contributed to the Feminist Data Manifest-No have articulated the concept of *critical refusal* through case studies illustrating "how data practices can be at the center of issues that impact vulnerable communities" [107]. Researchers in computer science have studied practices including crowd-worker mutual aid [151], collective actions by platform moderators [187], and political organizing around smart city implementation [282], although this work is not typically examined explicitly through the lens of refusal.

My purpose in advancing the idea of "data refusal from below" in chapter 9 is to emphasize the value of refusal by people harmed by data systems who have low direct power over technology creation. In computer science, designers do not always see actions by people who resist or reject technology as legitimate or worthy of study [223]. Yet as Communication scholar Seeta Peña Gangadharan argues, "we cannot just rely upon the enlightened goodwill of privileged elites to recognize and rectify [injustices]," because refusal initiated by the marginalized "matters just as much as recognition of the marginalized by privileged people or institutions" [106].

Even though technology designers can refuse in ways that others can't, their refusals can be consistent with the values of other refusers working "from below." Anthropologist and Kahnawà:ke citizen Audra Simpson has described how academics can respect the limits established by others with less institutional power. She writes about refusals taken up by her Kahnawà:ke community to reject settler-colonial ideas about borders, citizenship, and membership [248]. One such everyday refusal involved feigning ignorance about certain information during ethnographic interviews [247]. Simpson, in her capacity as an academic, chose to develop her own practice of *ethnographic refusal* by articulating limits to what she would report. By carefully considering "what you need to know and what I refuse to write in," [247] Simpson's ethnographic refusal has upheld the sovereignty of the Kahnawà:ke by respecting the limits the community places on what can be known by outsiders.

My encouragment that scholars and designers take a view from below is not meant to exclude or dismiss refusals by tech workers and academics. As other HCI scholars of tactics from below have noted, "above" and "below" are oversimplifications — researchers are often simultaneously situated "above" and "below" in different contexts [282]. Similarly, Ganesh and Moss note that "'inside' and 'outside' Big Tech [are not] watertight categories," even as they acknowledge the different discursive implications of different kinds of responses to harm [105]. Yet standpoint is a generative idea for thinking about acts of refusal. For example, organizers of the #NoTechForIce campaign name "educating communities about how to protect themselves against new forms of criminalization, taking direct action to confront corporate actors, [and] organizing with tech workers and students to leverage their influence over Silicon Valley" as some of several ways they are resisting the relationship between tech companies and U.S. immigration enforcement. These different approaches all contribute to a wider movement by explicitly incorporating an analysis of power and standpoint. Similarly, Indigenous studies scholars Tuck and Yang point out the need for analytic practices of refusal that "negotiate how we as social science researchers can learn from experiences of dispossessed peoples — often painful, but also wise, full of desire and dissent — without serving up pain stories on a silver platter for the settler colonial academy, which hungers so ravenously for them" [267]. I hope that this framework helps scholars see refusal practices from above and below as parallel and complementary.

2.3.3 Design for Consent and Refusal

Research is increasingly delivered digitally, and researchers and designers have investigated how to design software and procedures relevant to research ethics. The field of HCI can offer valuable contributions to scholarly conversations about research ethics, since so many matters in research ethics hinge on the design and user experience of ethics procedures.

Many designers work to make research ethics processes cost-effective for university bureaucracies, maximizing research output, and minimizing compliance risks. Many companies offer white-label IRB systems, which universities buy and apply their own branding to. Examples of commercial products include Cayuse IRB, Kuali Protocols, iRIS, and others [109]. These systems provide user interfaces for researchers to submit protocols for IRB review, and for IRB staff to quickly review large numbers of protocols. Case management systems like EthicsPoint [149] also provide a process for participants to anonymously report researcher misconduct.

One way to scale research is to design standardized consent forms for participants, which can be evaluated for readability using graphic design principles and cognitive measures [18, 39, 135, 259]. Researchers have also empirically evaluated the design of consent forms for participant comprehension and awareness of legal implications [12, 257, 288]. For example, studies have found that shorter consent forms are better understood by participants [84, 91]. However, critics note that efficient forms may not always lead to outcomes that protect participants. For example, researchers have found that people sign consent forms even when they are designed illegibly, concluding that consent forms do more to facilitate submission to authority than protect participant autonomy [154].

Designers have also worked to help researchers deliver digital equivalents of paper-based research ethics procedures. In psychology, designers of survey software have used pop-up windows to deliver debriefing information for consented participants who exit surveys before completion [171]. Researchers have also considered the trade-offs of using digital signatures to legally document informed consent [27]. More recent work has explored eConsent systems that fully replace paper documentation [67, 166]. Researchers have evaluated eConsent's effectiveness in terms of factors such as trust, scalability, and user experience [58]. As scholars of Feminist HCI have observed, affirmative consent involves more than simply providing users with a checkbox [150]. Wilbanks notes that eConsent procedures are usually implemented as single-point transactions, and proposes to "transform informed consent into an ongoing relationship of trust-based permission" in a digital context [284]. This involves not only delivering consent procedures digitally, but also designing interactive experiences to ensure participant comprehension and ongoing engagement in consent procedures.

As part of Wilbanks' work, Sage Bionetworks released an open source toolkit of interface components for designers to adapt into informed consent user experiences. This work has also been adapted into Apple's ResearchKit framework for iOS developers [285]. Typical user experiences with these tools involve a series of interactive concept assessments before subjects are presented with a consent form to sign.

Novel ethics procedures are not guaranteed to increase protections for participant autonomy. As Wilbanks points out, "it is just as possible to use the visual interface to obscure [concepts] as it is to ... reveal them" [284]. Outside of research ethics, consent management platforms have been widely adopted by tech companies in response to the European Union's General Data Protection Regulation (GDPR). Much design effort in consent management has gone toward nudging people to consent through "dark patterns" or misleading user experience designs that undermine autonomy [204]. Even without misleading designs, procedures that introduce incentives, barriers, or irrelevant information into a consent process can easily influence people to give up personal data [19].

Any research ethics system that relies on individual choice also struggles with a "consent dilemma" [252]. Scholars have argued that this model of "privacy self-management" overburdens individuals with an impossible task of never-ending decisions within a rapidly-changing, complex information landscape [252]. If people check a box out of resignation at the impossibility of privacy management, their recorded privacy preferences could be inconsistent with their actual preferences or behavior.

To overcome the limitations of systems based on individual autonomy, researchers have explored collective governance schemes for research ethics. For example, researchers have convened a representative group of citizens to discuss the details of genetic testing. If the representative body approves the research on the behalf of the group, individuals are offered a choice to consent to be governed by their deliberations [77, 169]. Similarly, community IRBs are formed by participants and work in partnership with institutional IRBs to review and negotiate over potential research [2, 3, 176, 215].

Another thread of work in HCI seeks to address problems of autonomy by restructuring the relationship between participants and researchers. The CivilServant system supports moderators of online communities in designing studies with the help of researchers, and provides processes for "community debriefings" involving public discussions of research results [188]. Research that is co-designed with participants, who are directly affected by and exert agency over how research is conducted, falls under the broader category of participatory research [69]. However, communities are often heterogenous and can contain multiple conflicting parties. For example, some community members may oppose the power held by a dominant group that may be working with researchers. Researchers working with online communities must navigate how they are positioned in relation to multiple conflicting social actors [164].

By designing and evaluating a research ethics system in chapter 8, I am advancing a body of design research akin to ResearchKit and eConsent that develops scalable user interfaces for common research ethics procedures. By automating and scaling the debriefing procedure, I extend rights-based autonomy protections to a large number of online research participants.

Part II

Empowered Participation in Data Analysis

Chapter 3

Introduction

Multi-modal interfaces that incorporate visualization, textual description, and sonification have become an important means of making data analysis accessible, but existing approaches often prioritize the visual modality — a critical limitation when it comes to BLV people's equal participation in data analysis. Tools often require an existing visualization to convert into an accessible format, making it difficult for a BLV user to independently create data representations or engage in rapid prototyping. Further, designers often approach non-visual representations using direct sensory re-mappings of existing visual chart forms, but this approach fails to leverage modality-specific, non-visual affordances.

Current web accessibility guidelines ask visualization designers to support screen readers via basic non-visual alternatives like textual descriptions and access to raw data tables. But charts do more than summarize data or reproduce tables; they afford interactive data exploration at varying levels of granularity — from fine-grained datum-by-datum reading to skimming and surfacing high-level trends. In response to the lack of comparable non-visual affordances, I present a set of rich screen reader experiences for accessible data visualization and exploration (Chapter 4). Through an iterative co-design process, I identify three key design dimensions for expressive screen reader accessibility: *structure*, or how chart entities should be organized for a screen reader to traverse; *navigation*, or the structural, spatial, and targeted operations a user might perform to step through the structure; and, *description*, or the semantic content, composition, and verbosity of the screen reader's narration. I operationalize these dimensions to prototype screen-reader-accessible visualizations that cover a diverse range of chart types and combinations of my design dimensions. I evaluate a subset of these prototypes in a mixed-methods study with 13 blind and low vision readers. My findings demonstrate that these designs help users conceptualize data spatially, selectively attend to data of interest at different levels of granularity, and experience control and agency over their data analysis process.

To de-center the visual modality in tools for authoring data representations, I present Umwelt: an authoring environment for interactive multimodal data representations (Chapter 5). In contrast to prior approaches, which center the visual modality, Umwelt treats visualization, sonification, and textual description as coequal representations: they are all derived from a shared abstract data model, such that no modality is prioritized over the others. To simplify specification, Umwelt evaluates a set of heuristics to generate default multimodal representations that express a dataset's functional relationships. To support smoothly moving between representations, Umwelt maintains a shared query predicate that is reified across all modalities — for instance, navigating the textual description also highlights the visualization and filters the sonification. In a study with 5 blind / low-vision expert users, I found that Umwelt's multimodal representations afforded complementary overview and detailed perspectives on a dataset, allowing participants to fluidly shift between task- and representation-oriented ways of thinking.

Chapter 4

Rich Screen Reader Experiences for Accessible Data Visualization

4.1 Introduction

Despite decades of visualization research and recent legal requirements to make web-based content accessible [131, 278], web-based visualizations remain largely inaccessible to people with visual disabilities. Charts on mainstream publications are often completely invisible to screen readers (an assistive technology that transforms text and visual media into speech) or are rendered as incomprehensible strings of *"graphic graphic graphic"* [227, 239]. Current accessibility guidelines ask visualization designers to provide textual descriptions of their graphics via alt text (short for alternative text) and link to underlying data tables [114, 279]. However, these recommendations do not provide modes of information-seeking comparable to what sighted readers enjoy with interactive visualizations. For instance, well-written alt text can provide a high-level takeaway of what the visualization shows, but it does not allow readers to drill down into the data to explore specific sections. While tables provide readers with the ability to hone in on specific data points, reading data line-by-line quickly becomes tedious and makes it difficult to identify overall trends.

Developing rich non-visual screen reader experiences for data visualizations poses several unique challenges. Although visuomotor interactions (like hovering, pointing, clicking, and dragging) have been core to visualization research [79], screen readers redefine what interaction is for visualization. Rather than primarily *manipulating* aspects of the visualization or its backing data pipeline [79, 128, 291], screen readers make *reading* a visualization an interactive operation as well — users must intentionally perform actions with their input devices in order to cognize visualized elements. Moreover, as screen readers narrate elements one-at-a-time, they explicitly linearize reading a visualization. As a result, in contrast to sighted readers who can choose to selectively attend to specific elements and have access to the entire visualization during the reading process, screen reader users are limited to the linear steps made available by the visualization author and must remember (or note down) prior output conveyed by the screen reader. Despite these modality differences, studies have found that screen reader users share the same information-seeking goals as sighted readers: an initial holistic overview followed by comparing data points [239], akin to the information-seeking mantra of “overview first, zoom and filter, and details on demand” [244].

In this chapter, I begin to bridge this divide by conducting an iterative co-design process (collaborator Daniel Hajas is a blind researcher with relevant experience) prototyping rich and usable screen reader experiences for web-based visualizations. I identify three design dimensions for enabling an expressive space of experiences: *structure*, or how the different elements of a chart should be organized for a screen reader to traverse; *navigation*, which describes the operations a user may perform to move through this structure; and, *description*, which specifies the semantic content, composition, and verbosity of text conveyed at each step. I demonstrate how to operationalize these design dimensions through diverse accessible reading experiences across a variety of chart types.

To evaluate the contribution, I conduct an exploratory mixed-methods study with a subset of my prototypes and 13 blind or low vision screen reader users. I identify specific features that make visualizations more useful for screen reader users (e.g., hierarchical and segmented approaches to presenting data, cursors and roadmaps for spatial navigation) and identify behavior patterns that screen reader users follow as they read a visualization (e.g., constant hypothesis testing and validating their mental models).

4.2 Design Dimensions for Rich Screen Reader Experiences

Currently, the most common ways of making a visualization accessible to screen readers include adding a single high-level textual description (via alt text), providing access to low-level data via a table, or tagging visualization elements with ARIA labels to allow screen readers to step through them linearly (e.g., as with Highcharts [132]). While promising, these approaches do not afford rich information-seeking behaviors akin to what sighted readers enjoy with interactive visualizations. To support systematic thinking about accessible visualization design, I introduce three design dimensions that support rich, accessible reading experiences: *structure*, or how elements of the visualization should be organized for a screen reader to traverse; *navigation*, or the mechanisms by which a screen reader user can move from one element to another; and *description*, or what semantic content the screen reader conveys.

Methods. I began by studying the development of multi-sensory graphical systems, covering work in critical cartography [168, 283], blind education [13, 111], tactile graphics [15, 52, 75, 103, 126], and multi-sensory visualization [23, 47, 61, 184]. Drawing on conventions and literature on crip, reflective, and participatory design [71, 121, 238], I began an iterative co-design process with collaborator Daniel Hajas, who is a blind researcher with relevant expertise. Hajas is a screen reader user with a PhD in HCI and accessible science communication, but he is not an expert in visualization research. Co-design — particularly as encapsulated in the disability activism slogan, *“Nothing about us, without us”* [71] — is important because it can eliminate prototypes that replicate existing tools, solve imaginary problems (i.e., by creating disability dongles [153]) or unintentionally produce harmful technology [242]. To balance engaging disabled users while acknowledging academia’s traditionally extractive relationship with marginalized populations [69], I intentionally acknowledge

Hajas as both co-designer and co-researcher. We believe that the distinction between co-designer — a phrase that often discounts lived experience as insufficiently academic — and researcher is minimal; technical, qualitative, and experiential expertise are all important components of this research. Hajas' profile is a perfect example of the intersection between lived experience of existing challenges and solutions, academic experience of research procedures, and an interest in the science of visualization. While he does not represent all screen reader users, his academic expertise and lived experience uniquely qualify him to be both researcher and co-designer. Nevertheless, to incorporate a diverse range of perspectives, I recruited additional participants as part of an evaluative study (§ 4.4).

Our work unfolded over 6 months and yielded 15 prototypes. Collaborators met weekly for hour-long video conferences. In each session, we would discuss the structure and affordances of the prototypes, often by observing and recording Hajas' screen as he worked through them. We would also use these meetings to reflect on how the prototypes have evolved, compare their similarities and differences, and whiteboard potential design dimensions to capture these insights. Following these meetings, Hajas wrote memos detailing the motivations for each prototype, tagging its most salient features, summarizing the types of interactions that were available, enumerating questions that the prototype raises, and finally providing high-level feedback about its usefulness and usability. In the following section, I liberally quote these memos to provide evidence and additional context for the design dimensions.

4.2.1 Structure

I define *structure* to mean an underlying representation of a visualization that organizes its data and visual elements into a format that can be traversed by a screen reader. Through the co-design process, I identified two components important to analyzing accessible structures: their *form*, or the shape they organize information into; and *entities*, or which parts of the visualization specification are used to translate a chart into a non-visual structure. Design decisions about form and entities are guided by considerations of *information granularity*, or how many levels comprise the range between a high-level overview and individual data values.

Form. Accessible structures organize information about the visualization into different forms, including lists, tables, and trees. Consider existing best practices and common approaches. A rasterized chart with alt text is represented to a screen reader as a single node. SVG-based visualizations can additionally be tagged with ARIA labels to describe the axes, legends, and individual data points. Despite SVG's nesting, screen readers linearize these ARIA labels into a list structure so that the user can step through them sequentially. Data tables, on the other hand, provide a grid structure for screen readers to traverse. At each cell of the grid, the screen reader reads out a different textual description, allowing the user to explore a space by traversing the grid spatially (up, down, left, and right) instead of merely linearly. Accessible visualization research has begun to explore the use of tree structures for storing chart metadata [281], but they remain relatively rare in practice. My prototypes primarily use trees as their branching and hierarchical organization allows users to browse different components of a visualization and traverse them at different levels of detail.

Entities. Where form refers to how nodes in a structure are arranged, entities instead refers to what aspects of the visualization the nodes represent. These aspects can include:

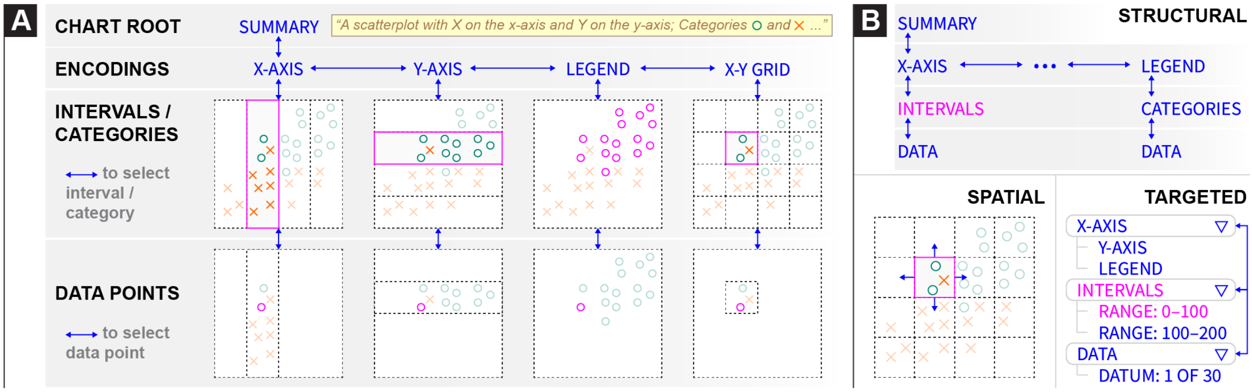


Figure 4.1: *(a) An accessible visualization structure in the form of a tree and comprised of encoding entities. Solid magenta outlines indicate the location of the screen reader cursor. Solid blue arrows between labels indicate available next steps via keyboard navigability (up, down, left, right). (b) Three ways of navigating accessible visualization structures: structural, spatial, and targeted.*

- *Data*, where nodes in the structure represent individual data values or different slices of the data cube (e.g., by field, bins, categories, or interval ranges). For example, in a data table, every node (i.e. cell) represents a data value designated by the row and column coordinates. Depending on the form, data entities can be presented at different levels of detail. For example, one prototype I explored represents a line chart as a binary tree structure (Fig. 4.2e): the root node represents the entire x-axis domain, and each left and right child node recursively splits the domain in half. Users can traverse the tree downward to binary search for specific values or understand the data distribution.

- *Encodings,* where nodes in the structure correspond to visual channels (e.g., position, color, size) that data fields map to. For instance, consider Figure 4.1a which depicts the encoding structure of a Vega-Lite scatterplot. The visualization is specified as mappings from data fields to three visual encoding channels: `x`, `y`, and `color`. Thus, the encoding structure, which here takes the form of a tree, comprises a root node that represents the entire visualization and then branches for each encoding channel as well as the data rectangle (x-y grid). Descending into these branches yields nodes that select different categories or interval regions, determined by the visual affordances of the channel. For instance, descending into axis branches yields nodes for each interval between major ticks; x-y grid nodes represent cells in the data rectangle as determined by intersections of the axes gridlines; and legend nodes reflect the categories or intervals of the encoding channel (i.e., for nominal or quantitative data respectively). Finally, the leaves of these branches represent individual data values that fall within the selected interval or category.

- *Annotations,* where nodes in the structure represent the rhetorical devices a visualization author may use to to shape a visual narrative or guide reader interpretation of data (e.g., by drawing attention to specific data points or visual regions). Surfacing annotations in the visualization structure allows screen reader users to also benefit from and be guided by the author’s narrative intent. For example, Figure 4.2d illustrates an annotation tree structure derived from an example line chart with two annotations highlighting intervals in the temporal x-axis. The root of the tree has two children representing the two annotated regions. The these two annotation nodes have a child node for each data point that is highlighted within the region of interest.

Considerations: Information Granularity. When might users prefer nested structures (i.e. trees) over flat structures (i.e., lists and tables)? Like sighted users, screen reader users seek information by looking for an overview before identifying subsets to view in more detail [239]. Trees allow users to read summary information at the top of the structure, and traverse deeper into branches to acquire details-on-demand. Kim et al. use the term *information granularity* to refer to the different levels of detail at which an accessible visualization might reveal information [167]. They organize granularity into three levels: existence, overview, and detail. *Existence* includes information that a chart is present, but no information about underlying data. *Overview* includes summary information about data — e.g. axes, legends, and summary statistics like min, max, or mean — but not individual data points. *Detail* includes information about precise data values.

I use the root node to signal the existence of the tree, and deeper nodes in the tree reflect finer levels of granularity. Branch nodes give an overview summary about the data underneath, providing information scent [212], while leaf nodes map to individual data points. In his feedback about the prototype shown in Figure 4.1, Hajas wrote *"considering how difficult reading a scatterplot with a screen reader is due to its sequential reading nature, the tree structure makes the huge number of data points fairly readable"*.

Entities are not mutually exclusive, and a structure might opt to surface different entities in parallel branches. I prototyped a version of Figure 4.2d which placed an encoding tree and annotation tree as sibling branches under the root node. Users could descend down a given branch, and switch to the equivalent location in the other branch at will. These design decisions are motivated by findings in prior work: by placing encodings and annotations as co-equal branches, I produce a structure that preserves the agency of screen reader users either to start with the narrative arc of annotations, or follow it after having the chance to interpret the data for themselves [178]. As Hajas confirms *"Depending on my task, either the encoding or annotation tree could be more important. If my task involved checking population growth in the last 100 years, I would start with the encodings. If I were to look for sudden changes in population numbers, such war-time mortality effects, I would start exploring the annotations, then tunnel back to the other tree."*

4.2.2 Navigation

Screen reader users need ways to traverse accessible structures to explore data or locate specific points. When browsing a webpage, screen readers provide a cursor that represents the current location in the page. Users use keyboard commands to step the cursor backward and forward in a sequential list of selectable items on the page, or jump to important locations such as headers and links. Through the prototyping process, I developed three ways of navigating through an accessible structure: *structural navigation*, *spatial navigation*, and *targeted navigation* (Fig. 4.1b). A key concern across these navigation schemes is reducing a user's cognitive load by affording a sense of the *boundaries* of the structure.

Structural Navigation. Structural navigation refers to ways users move within the accessible structure. I identify two types of structural navigation. *Local navigation* refers to step-by-step movements between adjacent nodes in the structure. This includes moving up and down levels of a hierarchy, or moving side to side between sibling elements. *Lateral navigation* refers to movement between equivalent nodes in adjacent sub-structures. For example, Fig. 4.2a depicts a multi-view visualization with six facets. When the cursor is on a Y-axis interval for the first facet, directly moving to the same Y-axis interval on the second facet is a lateral move.

Spatial Navigation. Sometimes users want to traverse the visualization according to directions in the screen coordinate system. I refer to this as spatial navigation. For example, when traversing part of an encoding structure that represents the visualization's X-Y grid, a downward structural navigation would go down a level into the currently selected cell of the grid, showing the data points inside the cell. A downward spatial navigation, in contrast, would move to the grid cell below the current one — i.e. towards the bottom of the Y-axis. Spatial navigation is also useful when navigating lists of data points, which may not be sorted by X or Y value in the encoding structure. Where a leftward structural navigation would move to the previous data point in the structure, a leftward spatial navigation would move to the point with the next lowest X value.

Targeted Navigation. Navigating structurally and spatially requires a user to maintain a mental map of where their cursor is relative to where they want to go. If the user has a specific target location in mind, maintaining this mental map in order to find the correct path in the structure to their target can create unnecessary cognitive load. I use targeted navigation to refer to methods that only require the user to specify a target location, without needing to specify a path to get there. For example, the user might open a list of locations in the structure and select one to jump directly there. Screen readers including JAWS and VoiceOver implement an analogous form of navigation within webpages. Instead of manually stepping through the page to find a specific piece of content, users can open a menu with a list of locations in the page. These locations are defined in HTML using ARIA landmark roles, which can designate parts of the DOM as distinct sections when read by a screen reader. When a screen reader user open the list of landmarks and selects a landmark, their cursor moves directly to that element.

Considerations: Boundaries & Cognitive Load. Screen reader users only read part of the visualization at a time, akin to a sighted user reading a map through a small tube [126]. How do they keep track of where they are? In our co-design process, we found it easiest for a user to remember their location relative to a known starting point, which is corroborated by literature on developing spatial awareness for blind people [61, 174, 283]. Hajas noted the prevalence of the `Home` and `End` shortcuts across applications for returning to a known position in a bounded space (e.g. the start/end of a line in a text editor). We also found that grouping data by category or interval was helpful for maintaining position. Hajas noted that exploring data within a bounded region was like entering a room in a house. In his analogy, a house with many smaller rooms with doors is better than a house with one big room and no doors. Bounded spaces alleviate cognitive load by allowing a user to maintain their position relative to entry points.

Comparing navigation techniques, Hajas noted that spatial felt *"shallow but broad"* while targeted felt *"deep but narrow."* While he expressed a personal preference for deep-narrow structures, he nevertheless *"would not give up [spatial navigation] because it makes me believe I'm actually interacting with a visualization."* This insight demonstrates the value of offering multiple complementary navigation techniques. Moreover, while targeted navigation facilitates quick searching and doesn't require the user to maintain a mental map to find specific data points, structural and spatial exploration enable more open-ended data exploration. It also provides a mechanism for establishing common ground with sighted readers (e.g., allowing both blind and sighted readers to understand a line segment as being "above" or "higher" than another).

4.2.3 Description

When a user navigates to a node in a structure, the screen reader narrates a description associated with that node. For example, when navigating to the chart's legend, the screen reader output might articulate visual properties of the chart's encoding: *"Category O has color encoding green; X has color encoding orange"* (Figure 4.1). Or, if that visual semantic content isn't relevant to understanding the data, it might ignore the color: *"each datum belongs to either Category O or X."* The *content*, *composition*, and *verbosity* of the description can affect a user's comprehension of the data. Designers must consider *context & customization* when describing charts.

Content. Semantic *content* is the meaningful information conveyed not only through natural language utterances, but also through the visualization (a graphical language [38]). Because graphics convey myriad different kinds of content, the challenge of natural language description is to convey information that is not only commensurate with what the chart expresses via graphical language, but also useful to its readers. Accessible chart description guidelines from WGBH [114], W3C [279], and others [159] offer prescriptions for conveying specific content for blind readers (such as the chart's title, axis encodings, and noteworthy trends). Lundgard and Satyanarayan expand the scope of these guidelines with a more general conceptual model of four levels of semantic content: *chart construction properties* (e.g., axes, encodings, marks, title); *statistical concepts and relations* (e.g., outliers, correlations, descriptive statistics); *perceptual and cognitive phenomena* (e.g., complex trends, patterns); and *domain-specific insights* (e.g., socio-political context relevant to the data) [178].

Decoupling a chart's semantic content from its visual representation helps designers better understand what data representations afford for different readers. For instance, Lundgard and Satyanarayan find that what blind readers report as most useful in a chart description is not a straightforward translation of the visual data representation. Specifically, simply listing the chart's encodings is much less useful to blind readers than conveying summary statistics and overall trends in the data [178]. As Hajas noted, *"I want to see the global trend, which is why sighted people rely on visualization."* For instance, for a stock market chart the reader *"might see the overview from first to last data points, and then zoom into an outlier in the middle."* These findings suggest opportunities interleave different kinds of content at different levels of a hierarchical structure to yield richer, more useful screen reader navigation. For example, injecting summary statistics (say, the existence of outliers within a particular subcategory of the data) higher up in the chart's tree structure (e.g., at the `legend` encoding node) might afford "scent" for "information foraging" [212], or further exploration down a particular branch (data subcategory) of the tree. Or, if navigating in a targeted fashion, the user might be afforded the option to directly navigate to outliers without traversing the tree.

Composition. The usefulness of a description depends not only on the content conveyed by its constituent sentences, but also on its *composition*: how those sentences are ordered in relation to each other. For example, during our co-design process, Hajas found that when navigating a chart's tree structure, the screen reader output could quickly become redundant, affecting how quickly and efficiently he could pick out the meaningful information at each node. For instance, the utterance *"Category: O, Point 3 of 15, x = 5, y = 12"* and the utterance *"x = 5, y = 12, Category: O, Point 3 of 15"* afford significantly different experiences for a user who wishes to quickly scan through individual data points. In the first utterance, the reader immediately receives content that helps to situate them in a broader data context, namely data labeled as "Category: O" at the `legend` node. In the second utterance, the reader immediately receives datum-specific content that helps to rapidly explore the fine-grained details within that data context. Whether a reader prefers one compositional ordering to another will depend on the task they are attempting to accomplish. As Hajas noted *"I like the label at the beginning of the information, saying at which level of the tree I am at. It is important for knowing where I am. It is also great that this information is only spoken out when I change level, but not when I navigate laterally."* These compositional choices are highly consequential for readers' experience, when they must repeatedly read nearly-identical utterances while navigating a structure.

Verbosity. Whereas composition refers to the ordering of content, *verbosity* refers to how much content the screen reader conveys. More content is not always better. As Hajas noted of Apple's Data Comprehension feature [74]: *"It can sometimes be too much information all at once, if it starts reading out all of the data. This is very difficult if you're interested in some data points that are in the middle. It is very play-or-stop."* Depending on the screen reader software, a user may be afforded control over how much content is conveyed. For instance, JAWS offers high, medium, and low verbosity levels [235]. At higher verbosity the screen reader announces more structural, wayfinding content (e.g. the start and end of regions). For data tables, verbosity configurations can affect whether the table size is read as part of the description, and whether row and column labels are repeated for every cell. Descriptions of nodes in an encoding structure might analogously include information about the path from the root — for example, by reminding the user that they are reading Y-axis intervals. These repetitions can help users remember their location within a structure, but additional verbosity is less efficient for comprehending the data quickly.

Considerations: Context & Customization. Apart from its constituent parts (content, composition, verbosity), a description's usefulness also depends on the *context* in which it is read: namely, the reader's task or intent, and familiarity with the data interface. The same description might be useful in some situations, but relatively useless in others. A reader's information needs are fundamentally context-sensitive. For example, as Hajas noted, when reading a news article, it may be satisfactory to accept a journalist's description of the data on good faith. But, when reviewing scientific research, *"I don't necessarily want to just believe what is said in the text, I want to check and double-check the authors' claims. Go down to the smallest numbers in the analysis. I want to be able to look at the confusion matrix and see if they made a mistake or not."* This targeted verification requires a description to afford users with precise look-up capabilities, in contrast to descriptions that may be generated when browsing or exploring the data.

This context-sensitivity reveals an important aspect of usability: a user's familiarity (or lack thereof) with the data interface. Wayfinding content (e.g., *"Legend. Category: O."*) can help a user remember their location in a structure, and may be useful while they assemble a mental map of the visualization. But, as they become accustomed to the interface and visualization, such descriptions may prove cumbersome. Because user needs depend on their task, preferences, and familiarity, interfaces might afford personalization and customization to facilitate context-sensitive description.

4.3 Example Gallery

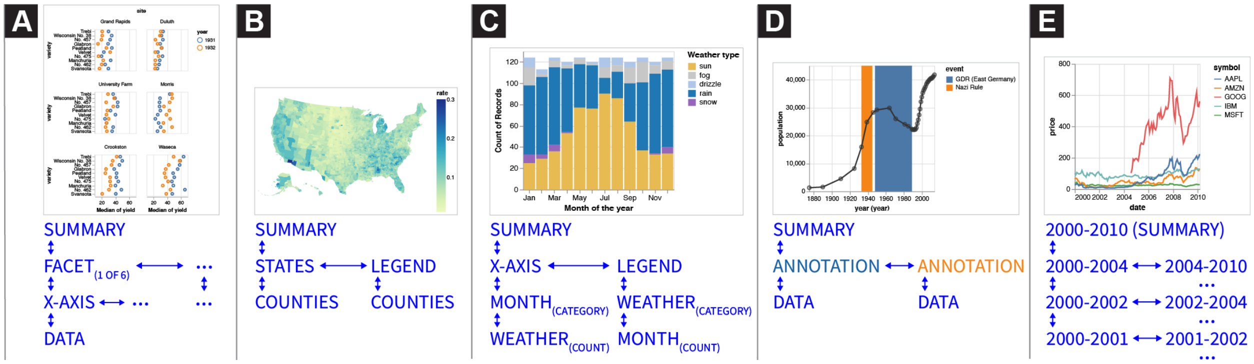


Figure 4.2: *Example structural and navigational schemes generated as part of the co-design process, and applied to diverse chart types.*

The co-design process yielded prototypes that demonstrate a breadth of ways to operationalize the design dimensions. Figure 4.2 excerpts some of the highest-fidelity prototypes, implemented on top of Vega-Lite [232]. As deeply nested structures and dynamic content are not well-supported by ARIA, I implemented my designs as in-memory data structures. Event listeners update the user's position in the structure on keypress, and write text descriptions to an ARIA-Live region (an ARIA role typically used for temporary notifications). To establish common ground with sighted users, I also render the visualization graphically. The user's position in the tree drives a Vega-Lite selection that highlights points when the screen reader user is attending to them.

For every prototype, the `up`, `down`, `left`, and `right` arrow keys enable structural navigation (moving up or down a level, or stepping through siblings respectively). For example, within the facet level of Fig. 4.2(a), the user can press `left` or `right` keys to move between the six subplots of the multiview chart. On charts that contain a node representing the x-y grid, users can also use the `WASD` keys to spatially navigate the grid and data points within that branch (mimicking an interaction found in video games).

These prototypes highlight different compositions of structures and navigation schemes. Fig. 4.2(a) includes `shift+left` and `shift+right` for lateral navigation across facets: pressing these keys at any node within a facet branch will navigate to the same location under an adjacent branch (subplot). With the chloropleth (Fig. 4.2(b)), I group data in the encoding structure by U.S. state; users can then drill down into counties across either this branch or the legend one. Fig. 4.2(c) offers two different paths for drilling down: month first, or weather first. Fig. 4.2(d) structures the tree by annotations rather than encoding: users can descend into the time intervals designated by the orange and blue rectangles, and view points within those intervals. Finally, Fig. 4.2(e) organizes its tree in terms of data, offering a binary search structure through the years.

4.4 Evaluation

To evaluate my contribution, I conducted 90-minute Zoom studies with 13 blind and low vision participants. Participants were asked to explore three prototype accessible screen reader experiences, shown one after another each with a different dataset. The goal of the evaluation was not to determine which particular combination of design elements was "best," but rather to be exploratory — to compare the relative strengths and advantages of instantiations of the design dimensions, and understand how they afford different modes of accessible data exploration.

4.4.1 Evaluation Setup & Design

Following Frøkjær and Hornbæk's Cooperative Usability Testing (CUT) method [102], I conducted each session by alternating between the role of guide (i.e., talking to the user and explaining the prototype) and logger (i.e., keeping track of potential usability problems, interpreting the data to prepare for becoming the guide). I began each session with a semi-structured interview to understand participants' current experiences with data and the methods they use to make inaccessible forms of data representation usable. The rest of the session focused on each of the three prototypes in turn, with each condition split into two phases: interaction and interpretation. In the interaction phase, I guided participants through the prototypes and asked participants to use them and comment on their process, in the style of Hutchinson et al.'s technology probes [147]. Then, I began a cooperative interpretation phase, where researcher and participants engaged in a constructive dialogue to jointly interpret the usability problems and brainstorm possible alternatives to the current prototype. In this method, participants influence data interpretation, allowing for more rapid analysis than traditional think-aloud studies as some analysis is built into each evaluation session with instant feedback or explanation from participants [102].

Prototypes. The in-depth nature of the cooperative interpretation sessions required me to balance the total number of prototypes evaluated (so that participants would have time to thoroughly learn and interact with each one) with a time duration appropriate for a Zoom session (limited to 90 minutes to avoid exhausting participants). Accordingly, I selected the following three prototypes, each representing a different aspect of the design dimensions:

- TABLE: An accessible HTML data table with all rows and three columns from the classic Cars dataset, in order to compare the prototypes with existing accessibility best practice.

- MULTI-VIEW: Becker's barley yield trellis display [34] as shown in Fig. 4.2a. This prototype features local and lateral structural navigation via the arrow keys and with the shift modifier respectively, as well as spatial navigation via `WASD`.

- TARGET: A single-view scatterplot, illustrated in Fig. 4.1, depicting the Palmer Penguins dataset [141]. In addition to structural and spatial navigation, targeted navigation is available via three dropdown menus corresponding to the structural levels.

TABLE is the control condition, as it follows existing best practice for making data accessible to screen readers. MULTI-VIEW enables me to study how users move between levels of detail, and whether they could navigate and compare small multiple charts. Finally, TARGET allows me to compare how and when the participants use the three different styles of navigation (structural, spatial, and targeted). I presented the prototypes in this sequence to all participants to introduce new features incrementally.

Participants. I recruited 13 blind and low vision participants through collaborators in the blind community and through a public call on Twitter. Each participant received $50 for a 90-minute Zoom session. I provide aggregate participant data following ethnographic practice to protect privacy and not reduce participants to their demographics [233]. Half of the participants were totally blind (n=7), while others were almost totally blind with some light perception (n=4) or low vision (n=2). Half of them have been blind since birth (n=7). Participants were split evenly between Windows/Chrome (n=7) and Mac/Safari (n=6). Windows users were also split evenly between the two major screen readers (JAWS, n=3; NVDA, n=4), while all Mac participants used Apple VoiceOver. These figures are consistent with recent surveys conducted by WebAIM which indicate that JAWS, NVDA, and VoiceOver are the three most commonly used screen readers [280]. Demographically, 70% of the participants use he/him pronouns (n=9) and the rest use she/her pronouns (n=4). One participant was based in the UK while the rest were spread across eight US states. Participants self-reported their ethnicities (Caucasian/white, Asian, and Black/African, Hispanic/Latinx), represented a diverse range of ages (20–50+) and had a variety of educational backgrounds (high school through to undergraduate, graduate, and trade school). Nine participants self-reported as slightly or moderately familiar with statistical concepts and data visualization methods, two as expertly familiar, and one as not at all familiar. Five participants described data analysis and visualization tools as an important component in their professional workflows, and 8 interacted with data or visualizations more than 1–2 times/week.

4.4.2 Quantitative Results

To supplement the cooperative interpretation sessions, participants rated each prototype using a series of Likert questions. I designed a questionnaire with six prompts measuring a subset of Brehmer and Munzner's multi-level typology of abstract visualization tasks [46]. This framework, however, required some adaptation for non-visual modes of search. In particular, searching with a screen reader requires a sequential approach to data that is at odds with the "at-a-glance" approach sighted readers take to browsing and exploring data. As my prototypes focus on navigation through charts, I collapsed the *location* dimension of Brehmer and Munzner's search decomposition resulting in two prompts that jointly measure `lookup-locate` and `browse-explore`. I formulated additional questions to measure Brehmer and Munzner's `discover` and `enjoy` tasks as well as more traditional aspects of technology acceptance including *ease-of-use* and *perceived usefulness* [163]. Participants responded on a five point scale where 1 = Very Difficult/Unenjoyable and 5 = Very Easy/Enjoyable.

Table 4.1: *Rating scores for each prototype (Table, Multi-view, Targeted) on a five point Likert scale where 1 = Very Difficult (Very Unenjoyable) and 5 = Very Easy (Very Enjoyable). Median scores are shown in boldface, averages in brackets, standard deviations in parentheses.*

PROMPT : *When using this prototype ...*	TASK [46]	TABLE	MULTI-VIEW	TARGETED
How enjoyable was it to interact with the data?	`enjoy`	**3** [3.31] (0.95)	**4** [3.77] (1.01)	**4** [3.54] (0.97)
How easy was it to generate and answer questions?	`discover`	**4** [3.15] (1.34)	**3** [3.00] (1.08)	**3** [3.23] (1.17)
If you already knew what information you were trying to find, how easy would it be to look up or locate those data?	`lookup-locate`	**3** [3.31] (1.32)	**4** [3.77] (1.17)	**4** [3.38] (1.19)
If you didn't already know which information you were trying to find, how easy would it be to browse or explore the data?	`browse-explore`	**2** [3.00] (1.68)	**2** [2.69] (1.11)	**3** [3.00] (1.29)
PROMPT : *When using this prototype ...*	USE [163]	TABLE	MULTI-VIEW	TARGETED
How easy was it to learn to use?	ease-of-use	**4** [4.15] (0.99)	**3** [2.69] (0.75)	**3** [3.15] (1.34)
How useful would it be to have access to this interaction style for engaging with data?	perceived usefulness	**4** [4.15] (0.80)	**4** [4.00] (0.82)	**4** [4.15] (1.07)

Table 4.1 displays the questionnaire prompts, their corresponding tasks, and statistics summarizing the participants' ratings. A Friedman test found a significant rating difference for the ease-of-use of the prototypes $\chi^2(2, N = 13) = 15.05, p < 0.01$, with a large effect size (Kendall's $W = 0.58$). Follow-up Nemenyi tests revealed that MULTI-VIEW was more difficult to use than TABLE with statistical significance ($p < 0.01$), but TARGET was not. Additional tests for the other prompts found neither statistically significant differences, nor large effect sizes, between the prototypes. However, median scores (which are more robust to outliers than means [196]) suggest that participants generally `enjoy` interacting with MULTI-VIEW and TARGET more, and found them easier to `lookup` or `locate` data with. Moreover, TARGET had the highest median score for affording `browse` or `explore` capabilities. Conversely TABLE was easiest to learn to use, and generally made it easy to `discover`, or ask and answer questions about the data. Notably, in response to the question "*How useful would it be to have access to this interaction style for engaging with data?*" participants on average ranked all prototypes as more-than-useful ($med = 4, \mu \geq 4$). These statistics provide only a partial picture of participants' experiences with the prototypes [21]. Thus, I elucidate and contextualize reasons behind their scores through qualitative analysis.

4.4.3 Qualitative Results

After the interviews, I qualitatively coded the notes taken by the logger with a grounded theory approach [57]. I performed open coding in parallel with the interviews (i.e., coding Monday's interviews after finishing Tuesday's interviews). I then synthesized the codes into memos, from which I derived these themes.

Tables are familiar, tedious, but necessary. Every participant noted that tables were their primary way of accessing data and visualizations. While tables are an important accessible option, participants overwhelmingly reported the same problems: they are ill-suited for processing large amounts of data and impose high cognitive load as users must remember previous lines of the table in order to contextualize subsequent values. As P2 reported, *"if I'm trying to get a general sense of the table, I'll just scroll through and see what values there are. But there's 393 rows, so I'll never scroll through all of it...I can't really get a snapshot."* P11 said that *"Finding relationships can be tricky if you're in a table, because you've got to either have a really good memory or just get really lucky. [...] If I didn't know what I was looking for, forget it."* At most, participants tabbed through 20–30 rows during their sessions, but did so only because of the questions I posed (e.g., "is there a relationship between horsepower and mileage?") and noted that if they encountered this table outside of the study, they would tab past a few rows to check for summary statistics and then move on.

While it is not enjoyable to explore or build a mental model of data with static tables, participants still emphasized their necessity because of the format's familiarity: *"in terms of accessibility, tables are infinitely more useful because there is a standard way of navigating them in whatever your preferred screen reader is. With different representations, a blind person may not be trained to interpret it"* (P2). This builds on prior literature [239] and echoes testimony from participants who had some difficulty with the new prototypes; they reported that they lacked expertise and therefore found it difficult to work with non-tabular data (P8, 10). In other words, to maximize accessibility, it is crucial to include a table view of the data *in addition to* other forms of novel interaction.

Prior exposure to data analysis and representations increases the efficacy of spatial representations. Participants who had experience conducting data analysis or reading tactile graphs/maps were able to easily develop a spatial understanding of how each prototype worked. Five participants (P2–4, 11, 13) made direct connections between the MULTI-VIEW and TARGET prototypes, and the tactile graphs they encountered in school. Three participants (P2, 11, 12) found their software engineering experience made it easier to understand and navigate the prototypes' hierarchical structure. Previous literature on tactile mapping has also shown how developing tactile graphical literacy is crucial for building spatial knowledge, but they emphasize that it is not a sufficient for being able to conduct and understand data analysis. [111, 126] Since the participants already had an existing spatial framework, it became easier to explain how a prototype might work using their prior experience as a benchmark, which has been corroborated by similar studies in tactile cartography. [13, 241, 283] Importantly, the participants were able to find specific origin points that they could return to in order to navigate the different branches of the tree, which would be further aided with help menus and mini-tutorials to understand the keyboard shortcuts (P2). Being able to shift between origin points is especially important for switching between graphs or between variables. By contrast, participants who had more difficulty with the prototypes (P8, 10) pointed to their lack of experience working with non-tabular data. P10 reported that being able to mentally visualize data points within a grid was a specific challenge. *"I suspect that this might be understandable to someone who's done this before,"* he said, *"I don't do well with these charts unless they're converted back into tables."*

***Structure*: Hierarchical representations make it possible to effectively convey insights with minimal cognitive load.** While static tables are the most common accessible option to interactive visualizations, eight of the participants (P2–5, 7, 10, 11, 13) expressed a desire to filter and sort the data so that they could begin to explore possible trends without wading line by line. Sorting and filtering a table is one way to look for trends but, to get a summary view of the data quickly, a system must provide snapshots in smaller intervals so that users can easily construct a larger picture or choose specific slices of the data to explore further (i.e., "details on demand"). [167, 244] With MULTI-VIEW and TARGET P4 said, *"I always want more layers and details, but some charts had too much…This was a happy medium between having the information I wanted and presenting it in a way that I can keep up with."* P5 also noted that he liked *"having the ability to scroll through at a higher level and then drill down deeper if that's of interest."* By giving users a way to quickly skip through the data across specific axes, they are able to rapidly generate a broader mental image of each graph and drill down further to collect more details. *"When I was working with the table, I [started building] a table in my head,"* P2 shared. *"I had a rough representation of it as a scatter plot. But here, I know how to drill down and up between different layers of data grids, so that I can get the overall picture… [It gives me] different ways of thinking."* Being able to control the parts of the data that were most important to them was also an issue of trust, as it also provided a way for users to reach conclusions for themselves rather than rely on the interpretation of others: *"It's hard to mix…doing your own analysis and be given a text description that you have to just trust"* (P12). In their own workflows, these participants reported downloading static tables to further examine and manipulate with Excel, which they would use to create summary statistics or intervals to move more quickly through the data.

***Navigation*: Reading a visualization with a screen reader entails constant hypothesis testing and pattern-making.** Since screen reader users parse data iteratively, nine of the participants (P1–5, 7, 8, 11, 13) described reading a visualization as a process of slowly building up a mental model and constantly testing it to see where the patterns may no longer hold. *"I'm going row by row, not memorizing exact numbers but building a pattern in my head, and looking at the other rows to test my theory,"* reported P3. In other words, the participants engaged in a continuous state of updating and validating [199] their mental images as new data challenged the existing patterns they have pieced together. MULTI-VIEW and TARGET accelerated this process, as participants were able to more rapidly identify specific components that they wanted to test. For example, P2 intentionally moved quickly across each level of the structure hoping to find its "edges," or the minimum and maximum limits of each axis and grid. *"Visually, it might look like I'm doing a lot of jumping around,"* he said, *"[but it's] because I'm trying to build the picture in a way that makes sense for me."* Similarly, P5 started building his mental model of the visualization by drilling up and down the grid to create a spatial image of the data: *"I'm thinking more in spatial terms just because [this] is a new method of navigating to me. [...] I'm moving through the grid...I'm thinking of drilling down into that square to get more information."*

TARGET made it especially easy for participants to test their hypotheses by giving them direct access to components that might break their hypotheses. P5 reported that it allowed him to *"navigate to areas...that I'm interested in, skipping over stuff that's not of interest,"* and P4 likened it to *"[being] able to go directly to what you want in a grocery inventory rather than going through each item one by one."* The ability to use structural, spatial, and target navigation in both MULTI-VIEW and TARGET respectively facilitated the hypothesis-testing and pattern-making behaviors that the participants were accustomed to with static tables, and gave them an additional mental model for working with the data. As P1 noted, these prototypes gave her a richer understanding of the data by helping her piece together *"both the picture and the mathematical pattern,"* whereas TABLE afforded only the latter.

***Description*: Cursors and roadmaps are important for understanding where you are.** Being able to capture both a high-level overview of the information while preserving the ability to drill down into the data is a crucial component to accessing interactive visualizations [239]. To navigate between these two levels, however, the participants emphasized the importance of markers to help them understand where they could move. TARGET addressed this with dropdown menus that allowed participants to navigate to any part of the visualization, explore, and then return to where they had started. In the words of P4, *"[This] mode is freedom for the user. Being able to jump around and move in real time as you would with your hand gives you a new way of exploring the information."* MULTI-VIEW approached this issue by allowing participants to move throughout the grid. *"With the table, I was trying to hold the numbers in my head and I wasn't trying to visualize it or anything,"* said P3. *"With [MULTI-VIEW], I can sort of think about it more like a visualization since I can move up and down, left and right. Even though I can use the arrows in the table, it just doesn't feel the same. I'm still feeling around and seeing what I can find."* Without these navigation tools, P7 noted that *"It's too easy to get lost ...I don't know how to backtrack."* To orient herself, P13 would first test to see if she was at the corner cells in the visualizations (e.g., *"Am I in the upper left or the bottom right cell here?"*) so that she could contextualize her position within the visualization and return to a point of origin. *"I know that I must be at the bottom left cell here because I can't go to the left,"* P13 said, *"but being able to know where that is beforehand would be very helpful."*

4.5 Discussion and Future Work

In this chapter, I explore how structure, navigation, and description compose together to yield richer screen reader experiences for data visualizations than are possible via alt text, data tables, or the current ARIA specification. My results suggest promising next steps about accessible interaction and representation for visualizations.

4.5.1 Enabling Richer Screen Reader Experiences Today

Although the design dimensions highlight a diverse landscape of screen reader experiences for data visualizations, the study participants attested to the value of following existing best practices. Namely, alt text and data tables provide a good baseline for making visualizations accessible. Thus, visualization authors should consider adopting the design dimensions to enable more granular information access patterns only after these initial pieces are in place.

Existing visualization authoring methods, however, are likely insufficient for instantiating the design dimensions or producing usable experiences for screen reader users. In particular, it currently falls entirely on visualization authors to handcraft appropriate structures, navigational techniques, and description schemes on a per-visualization basis. As a result, besides being a time-consuming endeavor, idiosyncratic implementations can introduce friction to the reading process. For instance, per-visualization approaches might not account for an individual user's preferences in terms of verbosity, speed, or order of narrated output — three properties which varied widely among the study participant in ways that did not correlate with education level or experience with data. Thus, to scale and standardize this process, some responsibility for making visualizations screen reader accessible must be shared by *toolkits* as well. For example, the prototypes suggest a strategy for translating visualization specifications into hierarchical encoding structures (i.e., encoding channels as individual branches, and using visual affordances such as axis ticks and grid lines to populate the hierarchy levels). If toolkits provide default experiences out-of-the-box, visualization authors can instead focus on customizing them to be more meaningful for their specific visualization, and screen reader users have a stronger guarantee that the resultant experiences will be more usable and respectful of their individual preferences.

Current web accessibility standards also present limitations for realizing the design dimensions. For instance, there is no standard way to determine which element the screen reader cursor is selecting. Where ARIA has thus far focused on annotating documents with the semantics of a pre-defined palette of widgets, future web standards might instead express how elements respond to the *interaction affordances* of screen readers. For example, ARIA could offer explicit support for overview/detail hierarchies and different levels of description detail that can be progressively read according to user preferences.

4.5.2 Studying and Refining the Design Dimensions

My conversations with study participants also helped highlight that design considerations can differ substantially for users who are totally blind compared to those who have low-vision. For example, partially-sighted participants used screen magnifiers alongside screen readers. As a result, they preferred verbose written descriptions alongside more terse verbal narration. Magnifier users also wished for in situ tooltips, which would eliminate the need to scroll back and forth between points and axes to understand data values. However, promisingly, I found that using a screen reader and magnifier together affords unique benefits: *"I would have missed this point visually if I solely relied on the magnifier because the point is hidden behind another point"* (P12). Future work should more deeply explore how accommodations might complement and conflict when designing for different kinds of visual disability.

Similarly, in scoping my focus to screen readers and, thus, text-to-speech narration, I refrained from considering multi-sensory modalities in the design dimensions. Yet, I found that most participants had previous experience with multi-sensory visualization, including sonification (P5, 7, 9, 13), tactile statistical charts (P2–4, 10, 11, 13), and haptic graphics (P3, 4, 11, 13). Some participants reported that a *combination* of modalities would further enhance their experience — for example, getting a sonic overview of a line chart before reading more detailed text descriptions. Other participants, however, cautioned that adding multiple modalities can create additional confusion. For example, P7 noted that *"There's often a lack of explanation about how to map between sound and text."* Based on this testimony, it is unlikely that "sensory modalities" are merely an additional, independent dimension within the framework. Rather, future work must unpack the affordances of individual modalities, how they interact with one another, and how they impact the design of structure, navigation, and description.

4.5.3 What are Accessible Interactions for Data Visualizations?

Visualization researchers typically distinguish between static and interactive visualizations, where the latter allows readers to actively manipulate visualized elements or the backing data pipeline [128, 291]. Screen readers, however, complicate this model: reading is no longer a process that occurs purely "in the head" but rather becomes an embodied and interactive experience, as screen reader users must intentionally perform actions with their input devices in order to step through the visualization structure. While some aspects of this dichotomy may still hold, it is unclear how to cleanly separate *static reading* from *interactive manipulation* in the context of screen reader accessible visualizations, if these notions are conceptually separable at all. For instance, Hajas likened the navigation the prototypes afforded to *"shifting eye gaze, shifting focus of perceptual attention. When I navigate a visualization, naturally I would say 'I'm looking at this figure' and not that 'I'm interacting with this figure'."* Analogously, recent results in graphical perception find that sighted readers do not simply "see" visualizations in a single glance but rather perform active visual filtering operations [42]. However, when using the binary tree prototype (Fig. 4.2e), Hajas noted a more distinct shift from reading to interacting. He said, *"it gave me the impression that I'm not just looking selectively, but I focus and zoom into the data,"* analogous to zoom interactions that change the viewport for sighted readers. Better characterizing the shift that occurs with this prototype, and exploring accessible manipulations of visualizations that allow screen reader users to meaningfully conduct data analysis, are compelling opportunities for future work.

Chapter 5

Accessible Structured Editing of Multi-Modal Data Representations

5.1 Introduction

For blind and low-vision (BLV) people to be equal participants in interactive data analysis, they must be able to not only consume data representations created by others, but also create their own custom representations by rapidly prototyping and examining alternative designs [111, 145]. Critically, to have full agency over this process, BLV people must be able to independently author and understand data representations without relying on sighted assistance [111]. In pursuit of these goals, accessibility research has begun to investigate multimodal data representation — that is, not only visualization but also textual description, sonification, and other modalities — with initial research results suggesting that the complementary use of multiple modalities can effectively facilitate analysis. For instance, when both sonification and textual description are available, a screen reader user can get a high-level overview from sonification and use it to contextualize their detailed textual exploration [95], and structured textual description helps low-vision magnifier users understand a scatterplot even when data points are visually occluded [299]. Each modality structures information with different spatial and temporal trade-offs, so they often afford different tasks, comparisons, and navigation strategies. Because of these modality differences, a screen reader user might want to author multiple representations to accomplish different goals, and easily switch between representations to develop a more holistic understanding of the data.

Unfortunately, existing tools for creating multimodal data representations center the visual modality: they assume the existence of a visualization that can *then* be converted into an accessible representation. For instance, Chart Reader [262] and VoxLens [240] derive textual and sonified representations from an input specification of a visualization, while the SAS Graphics Accelerator [4] and Highcharts Sonification Studio [55] provide editors for non-visual representations that require users to first specify a visualization. This ordering imposes limitations on both the authoring process and the expressivity of the output representations. In particular, it is challenging for a BLV person to independently create and interpret non-visual data representations unless they can first generate a corresponding visual chart — a problem that pervades existing statistical software [111]. Moreover, a visualization-centric authoring process imposes an undue emphasis on replicating visual affordances non-visually by directly re-mapping encodings, instead of considering the distinct affordances of non-visual modalities. As a result, this approach constrains the set of output representations that systems consider — for instance, Chart Reader and VoxLens restrict their support to a limited subset of chart forms that are straightforwardly amenable to sonification (e.g. bar charts, line charts) while sonifications based on other chart forms (e.g. scatterplots), or that diverge from the original chart's visual encoding (e.g. because they involve data transformations or interactions not specified in the visualization) remain underexplored.

In this chapter, I present Umwelt[1]: an authoring environment for multimodal data representations designed to de-center the visual modality. A screen reader user can use Umwelt's *structured editor* to specify data representations that include visualization, structured textual description, and sonification. Instead of using a visual specification to generate non-visual representations, Umwelt derives each modality from a shared abstract data model. As a result, users can create these representations in any order and/or specify only a subset of the three modalities as desired. Moreover, via different sections in the editor, a user can switch between editing all modalities simultaneously, or making fine-grained edits to a particular modality. To help users manage the upfront complexity of authoring a multimodal representation, the editor evaluates set of heuristics to generate default representations that express the dataset's functional relationships, and that a user can freely modify. For example, a stocks dataset with a field `price` that depends on independent variables `date` and `symbol` will result in default representations that afford easily looking up the `price` for a given tuple of `date` and `symbol` — a multi-series line chart, a textual structure that can group by `symbol` and `date`, and a sonification that plays back the `price` for each `date`, by `symbol` (Figure 5.6A).

The editor's state is rendered in Umwelt's *viewer* as independent visual, textual, and sonification views that are interactively linked together. These interactions help maintain a shared context across modalities — for instance, navigating the text structure also highlights the corresponding data visually and filters the sonification domain to only play the selected values — and encourage users to think of the modalities as complementary views into the data. Keyboard shortcuts help a screen reader user quickly move back and forth between the editor and viewer, enabling a tight non-visual feedback loop for confirming the results of edits during prototyping. The editor state is backed by an internal declarative specification (Figure 5.3). This specification language describes an expressive space of multimodal representations in the viewer. For example, Umwelt extends Vega-Lite's concept of view composition to express multi-view textual structures and sonifications.

I designed Umwelt through an iterative co-design process involving collaborator Daniel Hajas, who is a blind researcher with relevant expertise. I evaluate my contribution with multiple evaluation methods, following best practices [219]. Through an example gallery, I demonstrate that Umwelt's abstractions can express multimodal representations that span

---

[1] The system is named after the concept of *umwelt* in Jakob von Uexküll's semiotic theories [268]. An organism's *umwelt* is the perceptual world produced by its subjective sense experience. Sense faculties vary across species; for example, bats hear ultrasound, and birds sense magnetic fields. As science journalist Ed Yong and disability activist Alice Wong note in an interview [14], the idea of *umwelt* does not support the notion that there is a normative sensory apparatus, either throughout nature or within the human species. Instead, it encourages us to equally value different subjective sense experiences.

a variety of dataset semantics, data types, and view compositions. I also conduct a study involving 5 expert BLV screen reader users to understand how the editor and viewer help users conceptualize, author, and explore multimodal data representations. My findings surface rich themes about how screen reader users approach multimodal data analysis. I found that participants relied on complementary representations to move between overview and detail, and to manage cognitive and sensory load. Interactive synchronization and runtime customizations enabled participants to access the data by reconfiguring and switching representations to use the one that best suited their immediate needs. Participants also envisioned multimodal representations playing a role in facilitating communication between people who rely on different senses. I also found that the editor reduced challenges associated with specifying representations, and surfaced different ways of thinking about the relationship between specifications and users' goals.

5.2 Motivation: De-Centering the Visual Modality

In this section, I discuss how the overarching motivation of de-centering the visual modality in data analysis translates to concrete design goals for multimodal authoring systems. I then walk through an example usage scenario to demonstrate how instantiating these design goals in Umwelt enables a user to conduct independent data exploration using multiple complementary modalities.

5.2.1 Design Goals

I designed Umwelt through an iterative co-design process led by myself and collaborator Daniel Hajas. Hajas is a blind researcher with relevant experience in designing accessible representations. Over the course of about a year, we developed multiple prototypes of various interactive sonification and textual description techniques, accessible editor interfaces, and syntax prototypes of Umwelt's abstract model. We discussed prototypes regularly over Zoom call and email, reflecting on their strengths and weaknesses and brainstorming directions for additional iteration.

Early in the design process, I identified several challenges in the design of an authoring environment for multimodal data representations that arose from the core motivating principle of de-centering the visual representation. I synthesized these challenges into a set of design goals (DGs) that guided our iterative process and influenced team discussions where we reflected on candidate designs. Sections 5.3 and 5.4 will elaborate on how these design goals are addressed in the editor and viewer, respectively.

1. **Deferred commitment to a modality.** Authors often do not begin a rapid prototyping process with a concrete idea of their desired end state. As such, it is important for an authoring tool to offer the flexibility to easily try many candidate representations in an exploratory manner. In a multimodal system, this might involve freely editing different modalities in any order, or easily changing a field's mappings from one modality to another as they explore possible designs. In existing systems, the non-visual modalities depend on a visual modality. This requires an author to *prematurely*

commit [40] to a visual representation before specifying other modalities. My goal is instead to encourage deferred commitment. For example, an author should be able to specify non-visual modalities independently without first needing to create a visual specification.

2. **Complementary use of modalities.** An advantage of a multimodal system is that users are not required to rely on a single representation to meet all of their needs. Due to differences in how each modality conveys information, it is difficult to expect any single representation to act as a standalone replacement for another. Instead, a goal of my system is to encourage users to choose the modality that best suits their task at any given time, or use multiple modalities together to gain a more complete understanding.

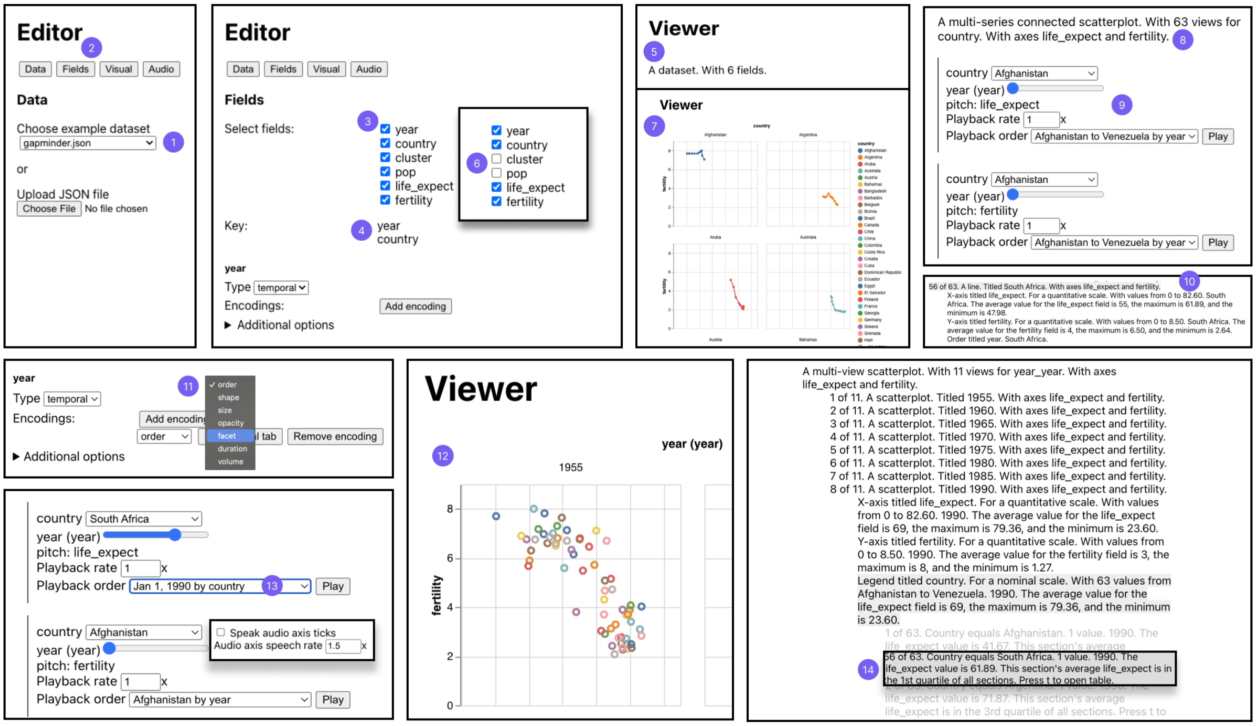
3. **Common ground between mixed-ability collaborators.** In a system designed to de-center the visual representation, the visualization still serves an important purpose for collaboration between people who primarily use different senses. For instance, screen reader users are not a monolith; people who use both a screen reader and a magnifier, like many low-vision users do, benefit from referencing a visual chart alongside other representations. Additionally, an important possible use of data analysis is to communicate about it with other people in personal or professional settings. BLV data analysts may need to communicate their findings to audiences that include people with different levels of vision, or participate in discussions where sighted colleagues are referencing visual concepts. Consequently, I find it important that representations establish common ground for diverse collaborators.

5.2.2 Authoring Multimodal Representations Co-Equally: An Example Usage Scenario

To demonstrate the process of authoring and analyzing multimodal representations in Umwelt, I walk through a scenario in which a screen reader user named Lula explores Hans Rosling's well-known Gapminder dataset [221].

Field-driven default specification. When Lula loads the dataset (Figure 5.1.1) into the editor in the DATA tab, the FIELDS tab (Figure 5.1.2) is populated with all of the dataset's fields. Lula inspects the set of checkboxes labeled "select fields" to ascertain that all six fields are initially checked (Figure 5.1.3) and thus are participating in the default multimodal representation. Reading the section below these checkboxes, Lula learns that Umwelt has inferred a composite key of `(year, country)` (Figure 5.1.4). Jumping over to the viewer by pressing the 'v' key on their keyboard, they observe that the initial representation is a textual structure that hierarchically groups and organizes the data for each field (Figure 5.1.5).

Lula decides that they want to analyze the life expectancy vs. fertility rate of countries over time, mirroring Rosling's original global health scatterplot [221]. They jump back to the editor with the 'e' key and tab through the checkboxes to keep only `year`, `country`, `life_expect`, and `fertility` checked (Figure 5.1.6). Umwelt's infers a new multimodal representation by reasoning about the dataset's keys and the measure types of the selected fields (i.e., nominal, quantitative, etc.). For the fields Lula has selected, as the data is keyed



A dataset. With 6 fields.

A multi-series connected scatterplot. With 63 views for country. With axes life_expect and fertility.

56 of 63. A line. Titled South Africa. With axes life_expect and fertility.
X-axis titled life_expect. For a quantitative scale. With values from 0 to 82.60. South Africa. The average value for the life_expect field is 55, the maximum is 61.89, and the minimum is 47.98.
Y-axis titled fertility. For a quantitative scale. With values from 0 to 8.50. South Africa. The average value for the fertility field is 4, the maximum is 6.50, and the minimum is 2.64.
Order titled year. South Africa.

A multi-view scatterplot. With 11 views for year_year. With axes life_expect and fertility.
1 of 11. A scatterplot. Titled 1955. With axes life_expect and fertility.
2 of 11. A scatterplot. Titled 1960. With axes life_expect and fertility.
3 of 11. A scatterplot. Titled 1965. With axes life_expect and fertility.
4 of 11. A scatterplot. Titled 1970. With axes life_expect and fertility.
5 of 11. A scatterplot. Titled 1975. With axes life_expect and fertility.
6 of 11. A scatterplot. Titled 1980. With axes life_expect and fertility.
7 of 11. A scatterplot. Titled 1985. With axes life_expect and fertility.
8 of 11. A scatterplot. Titled 1990. With axes life_expect and fertility.
X-axis titled life_expect. For a quantitative scale. With values from 0 to 82.60. 1990. The average value for the life_expect field is 69, the maximum is 79.36, and the minimum is 23.60.
Y-axis titled fertility. For a quantitative scale. With values from 0 to 8.50. 1990. The average value for the fertility field is 3, the maximum is 8, and the minimum is 1.27.
Legend titled country. For a nominal scale. With 63 values from Afghanistan to Venezuela. 1990. The average value for the life_expect field is 69, the maximum is 79.36, and the minimum is 23.60.
1 of 63. Country equals Afghanistan. 1 value. 1990. The life_expect value is 41.67. This section's average
56 of 63. Country equals South Africa. 1 value. 1990. The life_expect value is 61.89. This section's average life_expect is in the 1st quartile of all sections. Press t to open table.
life_expect value is 71.87. This section's average life_expect is in the 3rd quartile of all sections. Press t to

Figure 5.1: *An analyst's workflow in Umwelt: 1–6 illustrate an analyst's process of creating an initial multimodal data representation (shown in 7-9). 10 shows their initial exploration, before (11) making edits and then (12–14) continuing their analysis .*

by `(year, country)`, Umwelt assumes that a typical reader will use these fields to lookup the value fields `(life_expect, fertility)`.

Though there may be multiple ways to represent the same key-value semantics, the goal of Umwelt's heuristics is to provide an initial starting point rather than a single best representation. Thus, for the selected fields, Umwelt produces a multimodal representation that includes a small multiple of connected scatterplots (Figure 5.1.7), a textual structure (Figure 5.1.8), and two sets of audio controls (Figure 5.1.9), such that all modalities support this lookup operation via their modality-specific affordances. The visualization facilitates the lookup by faceting the data into multiple views by the `country` field, and using the `year` to order a connected scatterplot in each view. The sonification supports this same lookup by offering two sets of audio controls — each corresponding to an *audio unit*, or a single audio track that plays a continuous tone with pitch corresponding to `life_expect` or `fertility` respectively. Both audio units allow Lula to select `(year, country)` tuples, either via sequential playback by pressing the play button, or via interactive selection by manipulating a dropdown and slider. Finally, the textual structure facilitates the key-value lookup via its hierarchical structure. The first level below the root allows Lula to choose a `country`, and the next level allows Lula to drill further down into `year`, `life_expect`, or `fertility` (Figure 5.1.10). This hierarchy is generated based on the fields and key, and would exist even without a visual representation; however, because visual information is present, it's used to annotate the fields' descriptions in terms of x, y, and order encodings. At the lowest levels of the tree, Lula receives summary information about the selected data, and they can also press 't' to open a tabular view of the data.

Analyzing data using multiple modalities. Lula starts by pressing 'p' on their keyboard to play the sonification for `life_expect`. Umwelt orients them via *audio axis ticks*, which are spoken announcements of data values interleaved with the sonification to communicate playback progress. The system announces the first `country`, Afghanistan, and as the sonification progresses, the system speaks the `year` value prior to the sonified tone for each 5 year interval. After listening through a few `country` values, Lula understands how to interpret the sounds, and disables the audio axis ticks feature to more rapidly get a gist of the rest of the data. They observe that `life_expect` generally increases for most countries, but the min and max values can vary widely. For a few countries that sound different from the rest, Lula pauses the sonification by pressing 'p' on their keyboard again. To determine which country they were listening to, Lula tabs to the set of audio controls for `life_expect`, and inspects the `country` dropdown menu which reflects the current position in the paused playback.

Noticing that South Africa's `life_expect` sonification peaks in the middle before dropping again, Lula jumps to the textual structure with the 'o' key and navigates to the corresponding node by using the down arrow to move from the root level to the `country` level, and using the left and right arrow keys to find the node representing South Africa. Descending a level to the x-axis, they read the exact average, min, and max summary values of `life_expect` for South Africa — grounding the sonification they heard before in concrete numbers. By navigating to a sibling branch of the textual hierarchy, Lula is instead able to step through each year to read the exactly value of `life_expect` from `1955` to `2005`. To remind themselves of the overall trend, they press 'p'. Their position in the textual structure emits a query predicate that filters the sonification domain and highlights the corresponding data in the visualization. Because their cursor focus remains on the textual description for South Africa over a particular set of years, the sonification only plays through this data subset. Thus, they are able to identify that the peak they heard was for `1990`, when `life_expect` was `61.89`.

Editing the representation design. Umwelt's heuristics for these four fields prioritize using `country` as the "outermost" key, or the key that is used first in the composite key lookup operation. In other words, the multimodal representations afford looking up a specific `country` before exploring it by `year`. Now that Lula has done that, they may want to explore the data other way: by picking a `year` and exploring it by `country`.

To do so, Lula jumps back to the FIELDS tab of the editor, and removes the `facet` encoding from the definition for `country`. Umwelt updates the visualization to a single view containing a multi-line connected scatterplot (also known as a trace visualization [220]). The textual descriptions update to reflect this change, though the hierarchical structure remains unchanged; the sonification, similarly, does not change. Lula then updates the definition for `year`, switching the `order` encoding for a `facet` encoding instead (Figure 5.1.11). As a result, the visualization is faceted by year and each facet contains a scatterplot (Figure 5.1.12) with one point per country. The textual hierarchy now updates with `year` at the first level, and `country` nested underneath. The sonification still does not change as its traversal ordering is based on the key, which has remained the same throughout. Lula is able to verify these edits had the intended effect by quickly jumping back and forth between the viewer and editor with the 'e' and 'v' hotkeys.

Lula then repeats their preferred analysis process of sonification overview followed by detailed textual exploration. To listen to the `life_expect` values for each `country` in a given `year`, they first select a `year` using the slider — they start with `1990` to see what else was happening in the world during their previously observed notable year. Then, using the "playback order" dropdown, they select "1990 by country" (Figure 5.1.13). After turning "speak audio axis ticks" back on and pressing play, they hear the name of each `country` followed by a short tone corresponding to its `life_expect` in `1990`. This gives them a general sense of the variability of `life_expect` values in `1990`. Listening for South Africa, they have a sense of the relative position of that tone compared to the higher or lower tones that they've heard. Returning to the textual description, they navigate to `1990` and drill down into the `country` legend. They navigate to South Africa and are reminded of its average `life_expect`, and read that this value is in the 1st quartile of `life_expect` values — meaning that it is below the 25th percentile of values (Figure 5.1.14).

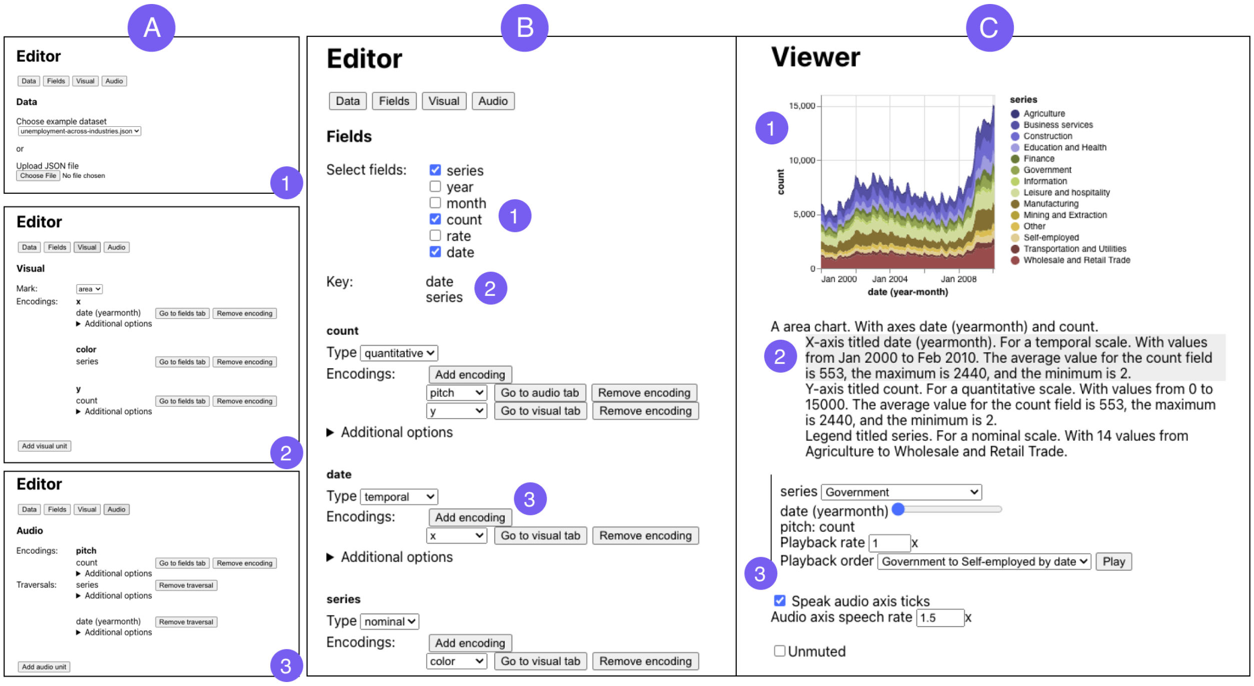
Summary. Using Umwelt, Lula was able to author multimodal data representations involving visualization, structured textual description, and sonification as part of a self-guided exploratory data analysis of the Gapminder data. Using heuristics that account for fields' measure type and the dataset's keys, Umwelt generated smart default specifications to help the analyst quickly get started without needing to think deeply about low-level specification across three modalities. Using an overview and detail strategy, Lula started by listening to the sonification, and contextualized what they heard with concrete data values by moving to the corresponding location in the textual hierarchy. This process of smoothly moving between modalities allowed them to leverage the distinct affordances of each modality in a complementary way. As they progressed, they recognized that their emergent goals during analysis would benefit from a change in the representations' affordances. By making a small number of atomic changes in the editor, Lula was able to generate a new textual hierarchy and adjust the sonification playback order to explore the data a different way.

5.3 The Umwelt Editor

With Umwelt's editor, users specify multimodal representations using an interface designed primarily for screen readers. The editor's internal state consists of a declarative JSON structure as shown in Figure 5.3. In this section, I first introduce key parts of the editor, including its main components and its default specification heuristics. Then, I discuss my design rationale and how it addresses the design goals.

5.3.1 Components of the Umwelt Editor

Umwelt's editor, as shown in Figure 5.2A and 5.2B, is split across four tabs. This organization is motivated by screen reader affordances, and navigation and wayfinding principles. In the co-design process, I found that when a screen reader user wanted to move back and forth between the editor and viewer with their screen reader, it was more difficult to maintain their position if the editor had too much extraneous content visible at once. Organizing the interface into tabs helps screen reader users think about what functionality they need at a given moment, and helps manage page length and the depth of the information hierarchy.



Editor

Data

Choose example dataset: unemployment-across-industries.json

or

Upload JSON file: No file chosen

Editor

Visual

Mark: area

Encodings:
- **x**: date (yearmonth) ▸ Additional options
- **color**: series
- **y**: count ▸ Additional options

Editor

Audio

Encodings:
- **pitch**: count ▸ Additional options

Traversals:
- series ▸ Additional options
- date (yearmonth) ▸ Additional options

Editor

Fields

Select fields:
- ☑ series
- ☐ year
- ☐ month
- ☑ count
- ☐ rate
- ☑ date

Key: date series

count

Type: quantitative

Encodings: pitch; y

▸ Additional options

date

Type: temporal

Encodings: x

▸ Additional options

series

Type: nominal

Encodings: color

Viewer

A area chart. With axes date (yearmonth) and count.

X-axis titled date (yearmonth). For a temporal scale. With values from Jan 2000 to Feb 2010. The average value for the count field is 553, the maximum is 2440, and the minimum is 2.

Y-axis titled count. For a quantitative scale. With values from 0 to 15000. The average value for the count field is 553, the maximum is 2440, and the minimum is 2.

Legend titled series. For a nominal scale. With 14 values from Agriculture to Wholesale and Retail Trade.

series: Government

date (yearmonth)

pitch: count

Playback rate: 1x

Playback order: Government to Self-employed by date

☑ Speak audio axis ticks

Audio axis speech rate: 1.5x

☐ Unmuted

Figure 5.2: *The Umwelt interface. A) The data, visual, and audio tabs of the editor. B) The editor's fields tab, where users specify field definitions and encodings. C) The viewer, where users analyze data with interactive multimodal data representations.*

Data Tab (Figure 5.2.A.1). A user begins by either loading a tabular dataset or choosing from a pre-populated list of example datasets. Umwelt then performs some simple type inference, and populates the other tabs with the dataset's fields.

Fields Tab (Figure 5.2.B). This tab lists all the fields in the dataset, with corresponding checkboxes to allow a user to pick which fields should participate in the multimodal representation. When a user checks or unchecks a field, the system evaluates a set of heuristics (described in 5.3.2) to produce a default multimodal representation. For each selected field, the editor provides a set of controls (Figure 5.2.B.3) to edit the field's inferred measure type, groupings, and transforms that may be calculated on the field (e.g. aggregation, binning). These definitions serve as a shared default across all modalities — defaults that can be overridden in modality-specific ways under the appropriate tab (described below). This tab also collates together the encodings a field is participating for both visual and audio modalities, offering user's a cross-modality perspective that can be important for generating cohesive and complementary experiences as I describe in § 5.3.3.

Visual Tab (Figure 5.2.A.2). This tab allows a user to make edits that apply only to the visual modality. A visual specification includes the visual-specific concept of a mark, and the encodings for that modality that were assigned in the Fields tab. Changes to a field definition (e.g. its transforms) apply only to the corresponding visual encoding. To allow users to be able to express multi-view displays (e.g., layered views or small multiples), Umwelt groups a mark and set of visual encodings into a *visual unit* — a concept Umwelt inherits from Vega-Lite [232]. Users can create multiple visual units, which can then be composed together as layers (where units are plotted one on top of the other) or as a concatenation (where units are laid out side-by-side).

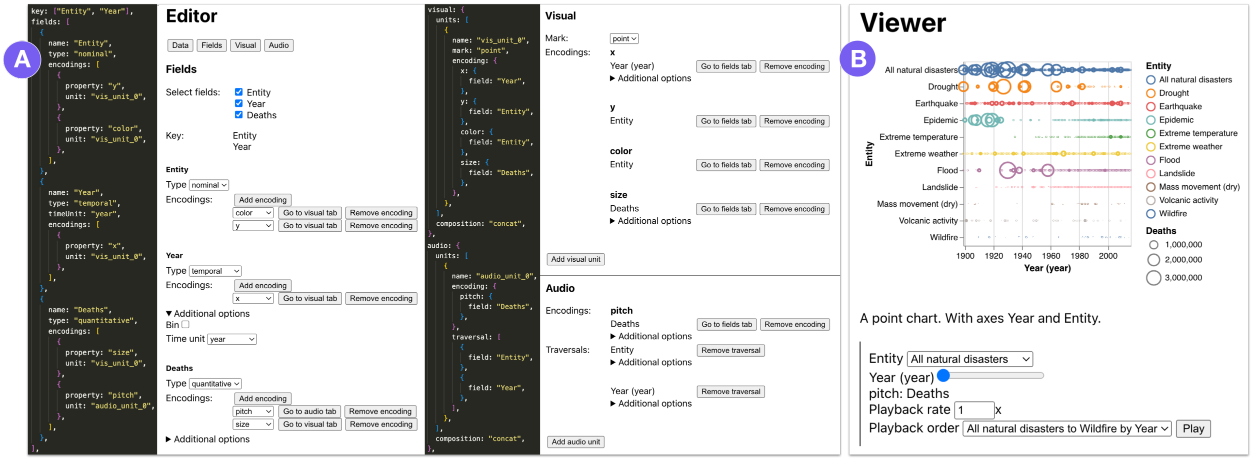
AUDIO Tab (Figure 5.2.A.3). This tab allows a user to make edits that apply only to the audio modality. An audio specification includes encodings from the FIELDS tab, which can be overridden, and *traversals*, an audio-specific abstraction I introduce to control the order in which data points are sonified.

While some visualization systems such as Vega-Lite offer an *order* visual encoding channel (and Tableau offers similar functionality via its *detail* shelf), this channel need only be used in special circumstances — for instance, to control the order that line segments are drawn as part of a connected scatterplot, or to determine z-axis and stack ordering. In contrast, ordering is much more central to the audio modality as data must be linearized into a fixed playback order, and different orderings afford different data lookups and comparisons. For instance, a reader of the stacked area chart in Figure 5.2.C may want to compare all values for `year`, one `series` at a time, to understand the trend of `count` within each `series`; or, they may want to traverse all `series` for a single `date` before moving onto the next `date`, to compare which series had the largest `count` at each `date`. A visualization reader could easily do both of these operations on the same chart. However, unlike the visual modality, a sonification can only afford one of these operations at a time — switching between the two requires re-ordering the data. Therefore an explicit specification of traversal is required.

A traversal specification is an ordered list of field definitions. The ordered list represents the precedence of groupby operations over the data, which are used to determine a linearized playback order. Consider the Figure 5.2 example again. The editor state in Figure 5.2A.3 defines a traversal `[series, date]`. This means that the data is first grouped by `series` before `date`. In the corresponding viewer state in Figure 5.2C.3, when a user presses play, the sonification will select the first value of `series` and iterate through all values for `date`. At each step, the current tuple of `(series, date)` values is used to query the value of `count` to encode it as pitch. After traversing all `date` values for the given `series`, the playback advances to the next value for `series` and iterates over all `date` values again. Consider the alternate traversal definition of `[date, series]`, which reverses the order in which the fields are grouped. In this case, the sonification would instead start with the first value for `date` and iterate through all values for `series` before proceeding to the next `date`. These two possible traversal specifications correspond to the two use cases described in the previous paragraph.

As this example demonstrates, it sometimes takes multiple sonification specifications to reproduce functionality afforded by a single visualization. To make it easier to provide multiple alternate sonifications, I also extend the concept of view composition to sonification. Like a unit visualization, a *unit sonification* contains a single set of encodings and traversals. Each audio unit corresponds to a single audio track that maps data to a tone, varying its properties (e.g. pitch, volume) according to the specified encodings. Concatenating two unit sonifications means providing two separate, independent audio playback controls side-by-side. A user can move between them to control which one they are listening to, and only one audio unit can be playing at once. Layering two unit sonifications means that they share a traversal and that their encodings are expressed through two audio tones playing simultaneously.

Future Work: Extending to Text and Other Modalities. Each modality is specified independently, yet each specification inherits from Umwelt's shared field definition. As



```
key: ["Entity", "Year"],
fields: [
  {
    name: "Entity",
    type: "nominal",
    encodings: [
      {
        property: "y",
        unit: "vis_unit_0",
      },
      {
        property: "color",
        unit: "vis_unit_0",
      },
    ],
  },
  {
    name: "Year",
    type: "temporal",
    timeUnit: "year",
    encodings: [
      {
        property: "x",
        unit: "vis_unit_0",
      },
    ],
  },
  {
    name: "Deaths",
    type: "quantitative",
    encodings: [
      {
        property: "size",
        unit: "vis_unit_0",
      },
      {
        property: "pitch",
        unit: "audio_unit_0",
      },
    ],
  },
],
```

Fields

Select fields: ☑ Entity ☑ Year ☑ Deaths

Key: Entity, Year

Entity
Type: nominal
Encodings: color, y

Year
Type: temporal
Encodings: x
▼ Additional options
Bin ☐
Time unit: year

Deaths
Type: quantitative
Encodings: pitch, size
▶ Additional options

```
visual: {
  units: [
    {
      name: "vis_unit_0",
      mark: "point",
      encoding: {
        x: {
          field: "Year",
        },
        y: {
          field: "Entity",
        },
        color: {
          field: "Entity",
        },
        size: {
          field: "Deaths",
        },
      },
    },
  ],
  composition: "concat",
},
audio: {
  units: [
    {
      name: "audio_unit_0",
      encoding: {
        pitch: {
          field: "Deaths",
        },
      },
      traversal: [
        {
          field: "Entity",
        },
        {
          field: "Year",
        },
      ],
    },
  ],
  composition: "concat",
},
```

Mark: point
Encodings:
x Year (year) ▶ Additional options
y Entity
color Entity
size Deaths ▶ Additional options

Encodings:
pitch Deaths ▶ Additional options
Traversals:
Entity ▶ Additional options
Year (year) ▶ Additional options

A point chart. With axes Year and Entity.

Entity: All natural disasters
Year (year)
pitch: Deaths
Playback rate 1 x
Playback order: All natural disasters to Wildfire by Year

Figure 5.3: *A) Fragments of an internal declarative specification shown next to their corresponding Umwelt editor states. B) The output multimodal representation for that specification.*

such, I expect it will be relatively straightforward to extend Umwelt's editor to support additional modalities (e.g., textual descriptions, tactile graphics, haptic feedback, etc.) provided there are well-defined abstractions and specification languages for these modalities. For now, although Umwelt currently supports textual output, I have chosen to not offer a TEXT tab as this remains a nascent research area without settled consensus on suitable abstractions. Instead, I have opted to preserve consistency between the textual structure and the visualization. I explain my rationale for this choice, and how it indicates a need for future research, in 5.3.3.

5.3.2 Default Specifications and Heuristics

Umwelt uses a set of simple heuristics to generate default multimodal representations based on a dataset's typings and key. In doing so, Umwelt seeks to avoid presenting a user with a blank slate whenever possible, and to accelerate a user in producing commonly used multimodal representations. Once the heuristics are evaluated, a user can modify the resultant defaults via the editor interface. These non-exhaustive heuristics are simple if-else statements that map combinations of field types and primary keys to specification fragments, which I document in Table 5.1. In the example gallery (Figure 5.6), examples A, B, C, G, and F were generated by Umwelt's heuristics while the rest required manual specification.

My heuristics are motivated by the idea of *functional dependence* between fields in a dataset. In database theory, a functional dependence is the relationship described by a dataset's primary key — a set of fields whose values uniquely index all rows of the dataset [66, 200]. Just as search algorithms over relational databases use the key to perform data lookups [66], an analyst using a data representation will often implicitly use the key to look up functionally dependent fields (also known as value fields). For example, a common way to read a single-series line chart is to choose a value for the x-axis field to look up a value for the y-axis field. Though some existing visualization systems, including Tableau, use key-value relationships to model visualizations [255], this concept is even more central to Umwelt because it provides a shared basis for expressing a dataset's semantics across multiple modalities.

I identified and validated the heuristics through the co-design process, manually authoring specifications for a diverse range of datasets with differently arrangements of typings and key. I worked to identify commonalities between the designs and try to articulate my intuition for why certain representations made more sense than others. For instance, I felt that the stacked area chart in Figure 5.2.C would be nonsensical if the `count` were encoded as `color`, and `series` as `y`. This can be explained by the functional dependence of `count` on `date` and `series`. In the sonification case, the key constrains which fields should be encoded at all; I was almost always interested in mapping a value field to an encoding property like pitch and using the key to determine the order of playback. For example, it does not make sense to sonify `date` or `series` in Figure 5.2.C — again, because `count` is functionally dependent `date` and `series`. Finally, in the case of text, the key imposes constraints on the hierarchical structure. In Figure 5.2.C, I was more likely to want to group by `date` and `series` to look up a `count` value than, for example, to look up a `date` by first finding its corresponding `count`.

Table 5.1: *Default specification heuristics based on a dataset's key and typings. T = temporal field, N = nominal field, Q = quantitative field. Each row represents a rule that matches a dataset's key and value tuples. The visualization, sonification, and textual description columns show default specifications for each rule, represented in YAML format for conciseness.*

Key	Value	Visualization	Sonification	Textual Structure	Example
T, N	Q				Figure 5.6A
T, N (>5 categories)	Q				Figure 5.6G

T, N

```
mark: "line"
encoding:
  x:
    field: t_key
  y:
    field:
        value[0]
  color:
    field: n_key
```

```
encoding:
  pitch:
    field:
        value[0]
traversal:
  - field: n_key
  - field: t_key
```

```
groupby: n_key
children:
  - groupby: t_key
  - groupby: value[0]
```

T, N (>5 categories)

```
mark: "point"
encoding:
  x:
    field: t_key
  y:
    field: n_key
  color:
    field: n_key
  size:
    field:
        value[0]
```

```
encoding:
  pitch:
    field:
        value[0]
traversal:
  - field: n_key
  - field: t_key
```

```
- groupby: t_key
- groupby: n_key
- groupby: value[0]
```

–	Q, Q, N	Figure 5.6B

```
mark: "point"
encoding:
  x:
    field:
        q_value[0]
  y:
    field:
        q_value[1]
  color:
    field: n_value
```

```
- encoding:
    pitch:
      field:
          value[0]
      aggregate:
          "mean"
  traversal:
    - field:
        value[1]
      bin: true
- encoding:
    pitch:
      field:
          value[1]
      aggregate:
          "mean"
  traversal:
    - field:
        value[0]
      bin: true
```

```
- groupby: q_value[0]
- groupby: q_value[1]
- groupby: n_value
```

T	Q, Q	Figure 5.6C

```
mark: "line"
encoding:
  x:
    field:
        value[0]
  y:
    field:
        value[1]
  order:
    field: key[0]
```

```
- encoding:
    pitch:
      field:
          value[0]
  traversal:
    - field:
        key[0]
- encoding:
    pitch:
      field:
          value[1]
  traversal:
    - field:
        key[0]
```

```
- groupby: key[0]
- groupby: value[0]
- groupby: value[1]
```

T, N, N

Q

```
mark: "point"
encoding:
  x:
    field:
        value[0]
  y:
    field:
        n_key[1]
  color:
    field: t_key
  facet:
    field:
        n_key[0]
```

```
encoding:
  pitch:
    field:
        value[0]
traversal:
  - field:
      n_key[0]
  - field:
      n_key[1]
  - field: t_key
```

```
groupby: n_key[0]
children:
  - groupby: value[0]
  - groupby: n_key[1]
  - groupby: t_key
```

Figure 5.6F

T, N

Q, Q

```
mark: "line"
encoding:
  x:
    field:
        value[0]
  y:
    field:
        value[0]
  facet:
    field: n_key
  color:
    field: n_key
  order:
    field: t_key
```

```
- encoding:
    pitch:
      field:
          value[0]
  traversal:
    - field: n_key
    - field: t_key
- encoding:
    pitch:
      field:
          value[1]
  traversal:
    - field: n_key
    - field: t_key
```

```
groupby: n_key
children:
  - groupby: value[0]
  - groupby: value[1]
  - groupby: t_key
```

Figure 5.1.7

A scatterplot. With axes Flipper Length (mm) and Body Mass (g).

binned Body Mass (g)
pitch: mean Flipper Length (mm)
Playback rate 1 x
Playback order 2700, 3000 to 6000, 6300

binned Flipper Length (mm)
pitch: mean Body Mass (g)
Playback rate 1 x
Playback order 170, 180 to 230, 231

```
{
  visual: {
    mark: "point",
    encoding: {
      x: { field: "Flipper Length (mm)" },
      y: { field: "Body Mass (g)" },
      color: { field: "Species" },
    },
  },
  audio: {
    units: [
      {
        name: "audio_unit_0",
        encoding: {
          pitch: { field: "Flipper Length (mm)", aggregate: "mean" },
        },
        traversal: [{ field: "Body Mass (g)", bin: true }],
      },
      {
        name: "audio_unit_1",
        encoding: {
          pitch: { field: "Body Mass (g)", aggregate: "mean" },
        },
        traversal: [{ field: "Flipper Length (mm)", bin: true }],
      },
    ],
    composition: "concat",
  },
}
```

```
{
  fields: [
    { name: "Species", type: "nominal", encodings: ["color"]
    {
      name: "Flipper Length (mm)",
      type: "quantitative",
      encodings: [
        "x",
        { property: "pitch", aggregate: "mean", unit: "audio_unit_0" },
      ],
    },
    {
      name: "Body Mass (g)",
      type: "quantitative",
      encodings: [
        "y",
        { property: "pitch", aggregate: "mean", unit: "audio_unit_1" },
      ],
    },
  ],
  visual: {
    mark: "point",
  },
  audio: {
    units: [
      {
        name: "audio_unit_0",
        traversal: [{ field: "Body Mass (g)", bin: true }],
      },
      {
        name: "audio_unit_1",
        traversal: [{ field: "Flipper Length (mm)", bin: true }],
      },
    ],
    composition: "concat",
  },
}
```

Figure 5.4: *Prototypes of encoding- and field-oriented specifications of a scatterplot with concatenated sonification, illustrating the role-inexpressiveness [40] of field-oriented textual specification. Color-coded spans on the left side of each text prototype show the lines of code that pertain to each modality: green represents visual while blue represents audio. A) Encoding-oriented specification groups each modality into unit specifications. B) Field-oriented specification fragments each unit's encodings across the spec.*

5.3.3 Design Rationale

Field- vs Encoding-Oriented Specification

The design of a specification language can impose constraints on a user's order of operations [40]. Conventional visualization grammars, including Vega-Lite, are *encoding-oriented*: encoding is a top-level abstraction in a Vega-Lite unit specification, and field definitions are nested within encodings. However, a consequence of encoding-oriented specification is that users must first decide what encodings they are using before assigning fields to them, requiring them to have visual idioms in mind when initially formulating their goals. This limitation is even more pronounced in the context of multimodal representations, as an author may not even have an initial choice of modality in mind. When I began designing Umwelt, I first designed it as an encoding-oriented declarative JSON language. However, the co-design process led me to explore *field-oriented* specification as an alternative: fields are top-level entities and encoding definitions are nested within fields. I felt that a field-oriented approach was amenable to multimodal authoring because an author can make localized changes to a single field definition used across multiple modalities, or switch a field's encoding from one modality to another. The increased ease of these changes enables deferred commitment to a specific representation (DG1). Comparing encoding-oriented and field-oriented specification using the cognitive dimensions framework [40], I argue that field-oriented specification reduces *viscosity* (difficulty of making changes) and increases *provisionality* (ease of exploratory prototyping).

Although field-oriented specification helped me address one of the design goals, I found that it became much more difficult to understand a specification without using modality-specific abstractions. Encoding-oriented specification is prevalent amongst existing declarative grammars because its syntax captures an important semantic property of the relationship between encodings and fields — namely, that each encoding property can only have one field assigned to it. This in turn enables the concise expression of other top-level abstractions: for instance, a unit visualization has one mark and one set of encodings. When reading a Vega-Lite spec, it is easy to understand that a mark and a set of encodings are associated together because they are contained within the same unit spec. When I switched to a field-oriented language, I found that modality-specific definitions became fragmented across field definitions. Consider the example in Figure 5.4. The encoding-oriented specification in 5.4A uses unit specs to group the functionality of each modality together. But in 5.4B's field-oriented specification, encodings belonging to the same unit specification are nested under multiple field definitions. Further, in 5.4B, modality-specific concepts like mark or traversal are not nested under any individual field, so additional verbosity or repetition must be introduced to associate these concepts with their respective units. In terms of cognitive dimensions, field-oriented specification introduces *role-inexpressiveness* [40] because it is more difficult to read a specification and clearly understand relationships and dependencies between entities affecting the same modality.

Field-oriented and encoding-oriented approaches both had affordances that felt essential but were in conflict with each other in a textual language, leading to significant tension in the design process. The co-design process led me to bridge between field- and encoding-oriented specification by designing Umwelt primarily as a structured editor interface, rather than as a textual JSON language. In the editor, the FIELDS tab allows a user to create a field-oriented specification by populating a field with encodings from any modality. Then, the user can switch to the VISUAL or AUDIO tabs to edit modality-specific attributes like mark and traversal, or perform actions that are scoped to one modality (like adding or removing a unit spec). In its internal representation (shown in Figure 5.3), Umwelt maintains both field-oriented and encoding-oriented abstractions. It links the two kinds of specification together via references. In my prototype language designs, expressing these references in a textual specification language was unwieldy and lead to repetition, but they are suitable for an interface where a user can easily navigate between two views into the same underlying spec. My eventual design for Umwelt prioritizes field-oriented specification to encourage ease of switching between modalities during exploratory specification, but also allows users to switch to encoding-oriented specification for detailed control.

Designing Umwelt as a structured editor also introduces additional benefits. An editor interface can reactively update the options it presents to a user based on its current state and can, thus, hide operations that would lead a user to invalid states. As a result, each atomic edit in the editor is a transition from one valid specification to another. In contrast, with a textual language, any time a user is partway through typing out a statement, the program will not compile. In terms of the cognitive dimensions of notation framework [40], an editor interface reduces *error-proneness* compared to the textual language, and affords users a better ability to *progressively evaluate* [40] the specification they are editing.

Aligned vs Disjoint Modalities

In a multimodal data representation, how each representation relates to the others can reflect different design priorities. For example, modalities can be *aligned* in that they redundantly encode the same information, emphasizing a cohesive insight or set of possible comparisons. Or, modalities can be *disjoint*, conveying different aspects of the data that can be synthesized together into a greater whole than the message of each individual representation. In existing systems that derive non-visual representations from the visual, the derived representations are inherently aligned with the original. But in systems like Umwelt where modalities are independent, it can be up to the author's discretion whether modalities are aligned or disjoint.

In the co-design process, thinking about aligned and disjoint modalities uncovered a tension in the design goals, where I seemingly could not simultaneously prioritize DG2 and DG3. On one hand, using visualization and sonification as *disjoint* modalities meant that I could use sonification to focus on comparisons between fields that are difficult to compare visually, or encode fields that are not present in the chart because it would be too visually overwhelming to include them. This additional expressiveness contributes to DG2, where a user can gain additional information from the use of multiple modalities together. On the other hand, using visualization and text as *aligned* modalities preserves consistency between the two representations, which is crucial for BLV users who need the textual representation to access the visualization. During the authoring process, a screen reader user needs the representations to align to verify that they are creating sensible visuals. This verification process is crucial to DG3, establishing common ground between blind and sighted users.

In sum, the textual structure can serve a dual purpose of (1) textually conveying the data and (2) making the visualization accessible. These two purposes fulfill DG2 and DG3, respectively, but it is difficult to fulfill both purposes simultaneously because one implies a disjoint representation while the other implies an aligned representation. While Umwelt could allow authors to override field definitions in the textual modality, this would cause the visualization and textual description to become disjoint. The co-design process led me to prioritize aligned visual and textual representations, and I made a decision not to expose a TEXT tab in the editor. Nonetheless, disjoint visual and textual representations is an important area for future design exploration. For instance, researchers could explore ways to enable a user to customize whether a textual structure is aligned or disjoint on-the-fly.

5.4 The Umwelt Viewer

Umwelt's viewer renders interactive multi-modal representations specified in the editor, including a visualization, a structured textual description, and a sonification. In this section, I first introduce the viewer's components and its linked interaction model. Then, I discuss my design rationale and how the viewer addresses the design goals.

5.4.1 Multi-Modal Data Representations

The Umwelt Viewer, as shown in Figure 5.2C, consists of three components: a visualization, a textual structure, and a sonification. Though there is no explicit interaction specification in the editor, each representation in the viewer is implicitly interactive. This interaction-first approach to the design of the viewer is motivated by the need to selectively attend to data. Interactive representations enable a user to select a subset of data and share that selection across multiple representations. Here, I describe each representation before discussing their interactive behavior in more detail in 5.4.2.

Visualization. Umwelt converts its internal representation into a Vega-Lite [232] specification to render a visualization (Figure 5.2.C.1). It augments this specification with additional Vega-Lite selection parameters, resulting in a visual representation that is interactive by default. For example, a user can drag on the visualization to select a rectangular region of data.

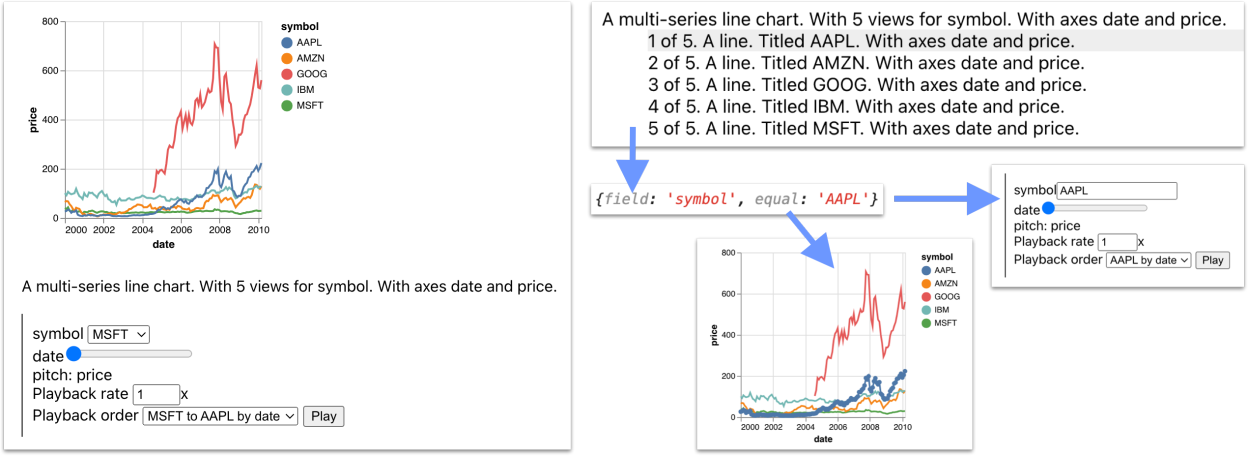
Textual structure. Umwelt renders a structured textual description (Figure 5.2.C.2) with Olli [41], an open-source library that implements Zong, Lee, Lundgard et al.'s design dimensions for screen reader experiences [299]. Olli outputs a hierarchical structure in the shape of a tree. Each node in the structure is associated with a textual description. The root of the structure gives an overall description of the data, while deeper levels in the structure apply successive filters on the data to give more granular descriptions.

The textual output does not require a visual specification, but can use visual information to augment its structure and descriptions. When there is a visual specification present, Olli structures the tree based on the visualization's encodings, and reference visual concepts in its description. As I discussed in 5.3.3, this makes the visual representation accessible for screen reader users and establishes common ground. On the other hand, when there is no visual specification, Olli outputs a relatively flat structure that allows a user to group the data by each field, and uses descriptions that do not reference visual concepts.

Sonification. Umwelt implements an interactive sonification runtime to render its audio specifications (Figure 5.2.C.3). For each audio unit specification, Umwelt renders a set of audio controls representing a single audio track. A user can press the play button (or the 'p' key on their keyboard) to play and pause the sonification. They can also interactively control their position within the sonification using input elements (i.e., dropdown menus for nominal and ordinal fields, and sliders for quantitative and temporal fields).

To help users keep track of their position in the sonification playback, the sonification runtime uses spoken announcements of data values interleaved with the sonification to communicate playback progress (exposed in the interface as an option called "audio axis ticks"). For example, in Figure 5.2C, as the sonification traverses `date` values, the system will speak the `date` value (e.g. Jan 2000) before playing the sonified segment between each axis tick. If a visualization is present, these ticks will always correspond to the visual ticks for consistency. Otherwise, they are determined by binning the key fields' domains to calculate regularly spaced intervals (or, for categorical fields, directly reading the value corresponding to each category).

Playback order is an important consideration for sonification, because different playback orders can facilitate different comparisons. For instance, in Figure 5.2C, playing through all `date` values for a given `series` before advancing to the next `series` is akin to the visual operation of reading each line left to right, getting a sense of each line's trend. On the other hand, playing through all `series` for a given `date` before advancing to the next `date` is akin to vertically comparing the values for a given x-axis position. Because the preferred order will depend on a user's goals, Umwelt determines the initial playback order by the ordering of the traversal specification and provides a dropdown menu to select an alternate playback order.



A multi-series line chart. With 5 views for symbol. With axes date and price.

symbol MSFT
date
pitch: price
Playback rate 1 x
Playback order MSFT to AAPL by date

A multi-series line chart. With 5 views for symbol. With axes date and price.
1 of 5. A line. Titled AAPL. With axes date and price.
2 of 5. A line. Titled AMZN. With axes date and price.
3 of 5. A line. Titled GOOG. With axes date and price.
4 of 5. A line. Titled IBM. With axes date and price.
5 of 5. A line. Titled MSFT. With axes date and price.

`{field: 'symbol', equal: 'AAPL'}`

symbol AAPL
date
pitch: price
Playback rate 1 x
Playback order AAPL by date

Figure 5.5: *An example of linked interaction across modalities, driven by the textual modality. Navigating to a node in the textual structure emits a query predicate. The visualization reifies this predicate as a conditional encoding, and the sonification reifies it as a filter.*

5.4.2 Coordinating Modalities with Linked Interactions

Though each modality maintains its own interactive state, Umwelt links interactions across modalities to aid analysis. Each modality has one or more interactions that define a selection over the data, and can be modeled as query predicates. For instance, a user can drag a brush over the visualization, navigate to a location or define a custom filter in the textual structure, and navigate to a position in the sonification playback. When a user performs one of these interactions on a representation, that representation updates its own state and emits a query predicate to the other representations. Each representation then reifies this predicate as some sort of effect (e.g. a transformation). Figure 5.5 shows an example of this process, driven by the textual modality. Olli associates a query predicate with each node in its structure — as a user navigates through the structure, the current node's predicate describes the data selected by the user's current position. In this example, a user navigates to a node corresponding to the predicate `{field: 'symbol', equal: 'AAPL'}`. This interaction emits the predicate to the visualization and the sonification. The visualization updates to visually highlight the selected data, and the sonification filters its domain to match the selected data.

5.4.3 Design Rationale

Highlighting vs Zooming in Non-Visual Representations

In visualization, the same user interaction could plausibly map to multiple possible effects. For example, dragging a rectangular area on a Vega-Lite scatterplot could *highlight* the data by giving it a conditional encoding (e.g., showing highlighted points in a different color). Or, that same drag interaction could *zoom* into that data (e.g. in an overview + detail interaction), resizing the viewport to only contain the selected data.

Analogously in non-visual modalities, there are multiple possible techniques for conveying the result of an interaction. Consider an example in the textual modality, which I surfaced while prototyping ways to filter a textual structure. One way of applying a filter to a textual structure is to re-scale the structure to fit the filtered data. For instance, an x-axis that originally represented a domain of 0–100 by splitting it into five nodes representing increments of 20 might be re-scaled to split a filtered domain of 50–70 into four nodes representing increments of 5. On testing this approach, co-designer Hajas compared this feature to "zooming in" on a visualization by changing its viewport. Another way of applying a filter is to leave the structure unchanged while re-flowing the structure with only the filtered data. For instance, the previous example would still have five nodes representing increments of 20, but many of the nodes would be empty after applying the filter. This approach is more analogous to "highlighting" a visualization, because the viewport remains the same but the un-selected data is de-emphasized.

Though zooming and highlighting appear to be recurring concepts across modalities, it is not clear that either is universally preferable. Currently, Umwelt's visual representation uses highlighting to convey interactive state because this is a more common interaction technique in visualization. This makes sense when considering the fact that visually, it is helpful to maintain a consistent viewport to situate a highlighted subset within the broader context of the full data. However, in the co-design process, I felt that the "zoom" interaction made more sense as a default for text, since a structure that is not scaled to the data often requires a user to navigate through extraneous nodes to find useful data. Guided by DG2, I chose these defaults per-modality according to each modality's affordances. I also considered cases where the modalities are used together — for instance, a sighted collaborator brushing on the visualization to momentarily draw a screen reader user collaborator's attention to a subset of data. However, my choice of default potentially trades off consistency across modalities — an important consideration for DG3.

Future work on interaction design for multi-modal data representations can work towards a better understanding of what types of approaches are best suited for certain situations or tasks, and how an author or end user might be able to switch between interaction techniques. And, though I conducted this initial exploration in the textual modality, future work remains to explore how interaction concepts like conditional encoding and viewport scaling extend to other non-visual modalities, like interactive sonification, in the context of a multi-modal system.

Preserving Interactive Context Across Modalities

Because the representations are designed to be used together, I wanted to enable users to smoothly switch between modalities to facilitate complementary use (DG2). This required me to think about how to maintain context when switching representations. In my initial explorations, a co-author compared the ability to select data via navigation in the textual structure with "pointing at part of a chart." I designed linked interactions so that the system could express a consistent understanding of what data the user is "pointing" at across all representations.

Another important goal of linked interaction was to establish common ground for collaboration and presentation (DG3). One of the most important uses of data is to communicate with others, and not everyone in a conversation may use the same sensory modalities. This is why, despite primarily designing Umwelt with screen readers in mind, I found it important to include a visual representation that visually conveys the state of a screen reader user's exploration in the textual structure or sonification. Conversely, the textual structure and sonification update to reflect interactions on the visualization. This also helps users think of the representations as complementary, e.g. by using one for wayfinding and the other for consuming [158] (DG2).

5.5 Evaluation: Example Gallery

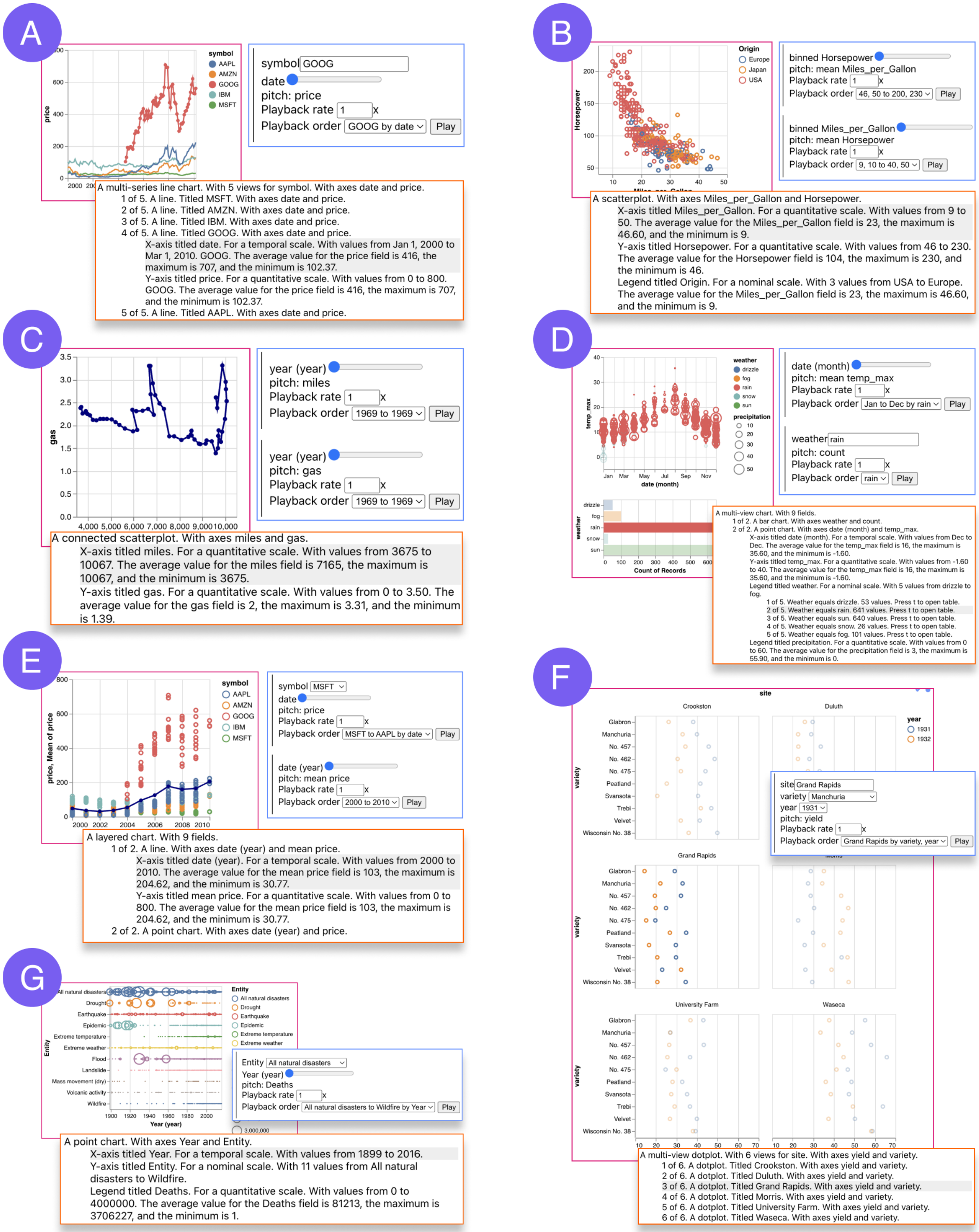
To evaluate my approach's expressive extent, I used Umwelt to create a gallery of diverse multi-modal examples representing a variety of key-value semantics (Figure 5.6). In addition to simple examples consisting of one visualization, one audio track, and one textual structure (A, G), Umwelt provides a simple view composition abstraction that can express concatenated and layered visualizations (D, E), concatenated sonifications (B, C, D, E), and multi-view textual tree structures (A, D, E, F).

In contrast to prior approaches, such as Chart Reader [262], that were constrained to a small set of chart forms (and therefore key-value semantics), Umwelt can express more complex relationships among fields. For instance, Figure 5.6.B shows a dataset with an empty key and a set of values with two quantitative fields and one nominal field, visually represented as scatterplot. Because there are two quantitative values, a user could plausibly want to look up either one by pitch. Figure 5.6.B provides two audio units so that users can choose which value field to sonify. However, since there is no key by which to look either value up, Figure 5.6.B's sonification uses binning and aggregation to transform the fields, creating a 1:1 correspondence between bins and aggregated values so that users can traverse the bins to look up a value. The result is a 2 dimensional sonification that conveys the distribution of quantitative values in the x or y orientations.

Though this is not the only possible way to sonify a scatterplot, Figure 5.6.B is illustrative of the importance of decoupling visual and non-visual specification (DG1) in order to express representations that achieve complementary goals (DG2). A system that derives audio encodings from visual encodings might re-map the `x` and `y` encodings directly to `pitch`. A user might want to bin before sonification, as shown in Figure 5.6.B, in order to get a high level sonic overview of how the data is distributed along an axis without being overwhelmed by the fluctuating values of each individual data point. But in a visualization-first system, because the visualization is not binned, a user would not be able to add binning to the sonification without first changing the visualization to a binned representation (e.g. heatmap). Yet, the user may not want to align the modalities in this way; they may want to switch between the sonification overview and a visual or textual representation of individual data points. Umwelt's approach enables a user to choose the set of representations that best suits their goals.

5.6 Evaluation: User Study

To evaluate Umwelt, I conducted remote studies with 5 expert BLV participants. Each participant met with me for two 90-minute sessions over Zoom video calling with screenshare for a total of 3 hours per participant. I split studies into two sessions to limit participant



symbol GOOG
date
pitch: price
Playback rate 1 x
Playback order GOOG by date Play

A multi-series line chart. With 5 views for symbol. With axes date and price.
1 of 5. A line. Titled MSFT. With axes date and price.
2 of 5. A line. Titled AMZN. With axes date and price.
3 of 5. A line. Titled IBM. With axes date and price.
4 of 5. A line. Titled GOOG. With axes date and price.
X-axis titled date. For a temporal scale. With values from Jan 1, 2000 to Mar 1, 2010. GOOG. The average value for the price field is 416, the maximum is 707, and the minimum is 102.37.
Y-axis titled price. For a quantitative scale. With values from 0 to 800. GOOG. The average value for the price field is 416, the maximum is 707, and the minimum is 102.37.
5 of 5. A line. Titled AAPL. With axes date and price.

binned Horsepower
pitch: mean Miles_per_Gallon
Playback rate 1 x
Playback order 46, 50 to 200, 230 Play

binned Miles_per_Gallon
pitch: mean Horsepower
Playback rate 1 x
Playback order 9, 10 to 40, 50 Play

A scatterplot. With axes Miles_per_Gallon and Horsepower.
X-axis titled Miles_per_Gallon. For a quantitative scale. With values from 9 to 50. The average value for the Miles_per_Gallon field is 23, the maximum is 46.60, and the minimum is 9.
Y-axis titled Horsepower. For a quantitative scale. With values from 46 to 230. The average value for the Horsepower field is 104, the maximum is 230, and the minimum is 46.
Legend titled Origin. For a nominal scale. With 3 values from USA to Europe. The average value for the Miles_per_Gallon field is 23, the maximum is 46.60, and the minimum is 9.

year (year)
pitch: miles
Playback rate 1 x
Playback order 1969 to 1969 Play

year (year)
pitch: gas
Playback rate 1 x
Playback order 1969 to 1969 Play

A connected scatterplot. With axes miles and gas.
X-axis titled miles. For a quantitative scale. With values from 3675 to 10067. The average value for the miles field is 7165, the maximum is 10067, and the minimum is 3675.
Y-axis titled gas. For a quantitative scale. With values from 0 to 3.50. The average value for the gas field is 2, the maximum is 3.31, and the minimum is 1.39.

date (month)
pitch: mean temp_max
Playback rate 1 x
Playback order Jan to Dec by rain Play

weather rain
pitch: count
Playback rate 1 x
Playback order rain Play

A multi-view chart. With 9 fields.
1 of 2. A bar chart. With axes weather and count.
2 of 2. A point chart. With axes date (month) and temp_max.
X-axis titled date (month). For a temporal scale. With values from Dec to Dec. The average value for the temp_max field is 16, the maximum is 35.60, and the minimum is -1.60.
Y-axis titled temp_max. For a quantitative scale. With values from -1.60 to 40. The average value for the temp_max field is 16, the maximum is 35.60, and the minimum is -1.60.
Legend titled weather. For a nominal scale. With 5 values from drizzle to fog.
1 of 5. Weather equals drizzle. 53 values. Press t to open table.
2 of 5. Weather equals rain. 641 values. Press t to open table.
3 of 5. Weather equals sun. 640 values. Press t to open table.
4 of 5. Weather equals snow. 26 values. Press t to open table.
5 of 5. Weather equals fog. 101 values. Press t to open table.
Legend titled precipitation. For a quantitative scale. With values from 0 to 60. The average value for the precipitation field is 3, the maximum is 55.90, and the minimum is 0.

symbol MSFT
date
pitch: price
Playback rate 1 x
Playback order MSFT to AAPL by date Play

date (year)
pitch: mean price
Playback rate 1 x
Playback order 2000 to 2010 Play

A layered chart. With 9 fields.
1 of 2. A line. With axes date (year) and mean price.
X-axis titled date (year). For a temporal scale. With values from 2000 to 2010. The average value for the mean price field is 103, the maximum is 204.62, and the minimum is 30.77.
Y-axis titled mean price. For a quantitative scale. With values from 0 to 800. The average value for the mean price field is 103, the maximum is 204.62, and the minimum is 30.77.
2 of 2. A point chart. With axes date (year) and price.

site Grand Rapids
variety Manchuria
year 1931
pitch: yield
Playback rate 1 x
Playback order Grand Rapids by variety, year Play

A multi-view dotplot. With 6 views for site. With axes yield and variety.
1 of 6. A dotplot. Titled Crookston. With axes yield and variety.
2 of 6. A dotplot. Titled Duluth. With axes yield and variety.
3 of 6. A dotplot. Titled Grand Rapids. With axes yield and variety.
4 of 6. A dotplot. Titled Morris. With axes yield and variety.
5 of 6. A dotplot. Titled University Farm. With axes yield and variety.
6 of 6. A dotplot. Titled Waseca. With axes yield and variety.

Entity All natural disasters
Year (year)
pitch: Deaths
Playback rate 1 x
Playback order All natural disasters to Wildfire by Year Play

A point chart. With axes Year and Entity.
X-axis titled Year. For a temporal scale. With values from 1899 to 2016.
Y-axis titled Entity. For a nominal scale. With 11 values from All natural disasters to Wildfire.
Legend titled Deaths. For a quantitative scale. With values from 0 to 4000000. The average value for the Deaths field is 81213, the maximum is 3706227, and the minimum is 1.

Figure 5.6: *Example multi-modal representations created with Umwelt, expressing a range of key-value semantics. A) Two keys and one value. B) No keys and three values. Audio units represent the quantitative values' distribution in 2 dimensions. C) One key and two values. Audio units represent looking up each value with the same key. D) Concatenated visualization and sonification. E) Layered visualization with concatenated sonification. F) Three keys and one value. G) An alternate visual representation for the same key-value semantics as A.*

fatigue, and to give participants adequate time to become familiar enough with the system and its concepts to surface insights that reflect ordinary use conditions. In the first session, participants used the viewer to analyze an example dataset using multiple modalities. In the second session, participants used the structured editor to choose from a set of example datasets and create their own multi-modal representations. Participants were compensated $250 for 3 hours. The purpose of the evaluation was primarily exploratory, seeking to form an initial understanding of how screen reader users approach multi-modal representations and their specification.

Because Umwelt is a tool designed with expert users in mind, I made an intentional choice to recruit a smaller number of participants to spend more time going in-depth with each participant. In qualitative research, the goal of selecting each qualitative case to examine is not to systematically answer descriptive questions about a population; it is to "ask how or why questions about processes unknown before the start of the study" [249]. Consequently, the goal of recruitment in the study was not to create the largest, most representative sample of a population, but to draw on participants' lived experiences to reach a "saturation" of insights — building my understanding to a point at which adding more participants stops giving me new or surprising information [249]. I found that three hours per participant with five participants enabled me to reach saturation.

Recognizing that there is a history of exploitative relationships between researchers and marginalized research participants [175, 292], I reference the participants by name throughout the paper with their consent. As scholars of citational justice in HCI note, a lack of intentional decision-making about who to acknowledge for their intellectual contributions can lead to the erasure of marginalized individuals' work and knowledge [172]. My intention with this choice is to appropriately acknowledge and credit the expertise for which I recruited. Following methodological recommendations to be specific about my target population and concept of expertise [51], my study conceives of expertise along two dimensions: screen reader experience, and data analysis experience. I describe the participants' backgrounds in Table 5.2.

5.6.1 Quantitative Results

I designed two Likert surveys to separately evaluate the user experience of the viewer and the editor. Participants responded on a scale of 1 to 5, where a higher number corresponds to an easier or more enjoyable experience. I report participants' responses in Table 5.3. The median scores suggest that participants generally found both the viewer and editor fairly easy to learn and enjoyable to interact with. According to participants, the viewer facilitates trend and pattern exploration in the data, and transitioning between modalities within the viewer is straightforward. While participants rated the editor as slightly more difficult to learn, many also expressed interest in investing more time to learn because of its capabilities. When it comes to making edits, participants found the sonification settings easy to customize, but had a harder time predicting updates in the viewer based on the changes made in the editor. In the qualitative analysis section, I will further contextualize participants' ratings.

Table 5.2: *Participants' names, demographic information, and descriptions of their screen reader and data analysis experience. Participants are referenced by name with their consent.*

Name	Age Bracket	Self-Description of Disability	Screen Reader Experience	Data Analysis Experience
Ben Mustill-Rose	20-35	Totally blind, lost sight in early teens	Proficient with NVDA and sonification add-ons	Frequently analyzes data in Python as software engineer
Ken Perry	50+	100% blind, lost sight in early 20s	Proficient with JAWS and other screen readers	Teaches Python and other programming languages, has written statistical software
Dorene Cornwell	50+	Low vision / high partial, had detached retinas in mid-life	Proficient with JAWS screen reader + ZoomText for magnification	Masters-level courses in statistics and related fields
Liam Erven	35-50	No usable vision, hearing impaired in right ear	Teaches students how to use assistive technology, including all major screen readers	Basic proficiency, uses spreadsheets
Amy Bower	50+	Low partial vision, cannot see computer screen. Declining vision since mid-20s	Proficient in JAWS (self-taught)	Does research on oceanographic data. Uses Matlab for data analysis.

Table 5.3: *Rating scores for the viewer and editor on a five-point Likert scale where 1 = Very Difficult (Very Unenjoyable) and 5 = Very Easy (Very Enjoyable). Median scores are shown in **bold**, averages in brackets [], standard deviations in parentheses ().*

Viewer	Score	Editor	Score
How easy was it to learn to use the viewer?	**4** [4] (0.71)	How easy was it to learn to use the editor?	**3** [3.4] (0.55)
After understanding how the viewer works, how enjoyable was it to interact with the data?	**5** [4.8] (0.45)	After understanding how the editor works, how enjoyable was it to edit the data representation?	**4** [3.8] (0.45)
After understanding how the viewer works, how easy was it to switch between descriptions and sonifications on-demand?	**4** [4.2] (0.45)	If you had a change you wanted to make to the data representation, how easy is it to understand how to make that change using the editor?	**4** [3.8] (0.84)
How easy was it to be able to customize the sonification settings, including playback mode, audio axis ticks, and playback rate?	**5** [4.6] (0.55)	After understanding how the editor works, how easy was it to predict how changes in the editor would affect the viewer?	**3** [3.6] (0.89)
After understanding how the viewer works, how easy was it to understand trends and patterns in the data?	**4** [4.4] (0.55)	After understanding how the editor works, how easy was it to check the result of your edits in the viewer?	**4** [4] (0.71)

5.6.2 Qualitative Results: Multi-Modal Viewer

Modalities have complementary affordances.

Participants found it useful to have multiple modalities available for many reasons, including increased optionality, modality-specific affordances, complementary uses leading to better understanding, and toggling between overview and detail.

Multiple representations as options to accommodate varying needs. Offering multiple modalities can help avoid cognitive or sensory overload. Minimizing cognitive load is a foundational principle in HCI; however, as research on accessible data analysis has shown, cognitive load can pose amplified challenges when it intersects with various disabilities [82]. As Erven noted, using only textual or tabular representations can result in "number fatigue" where the numbers "stop meaning anything." This fatigue can be compounded for people with disabilities related to attention management or memory. Having the option to switch from textual representations to sonification can potentially help provide more usable alternatives.

For users who may need to commit additional effort to use certain representations, having alternatives can also help manage sensory fatigue. Perry, who works with low-vision colleagues, suggested that they might like to "rest [their] eyes [while] flipping through the data." In these situations, being able to switch to a different representation can better accommodate an individual's needs by balancing their sensory load.

Complementary modalities enable better understanding via overview and detail. Just like sighted visualization users, studies [239, 299] have shown that BLV users follow the information-seeking heuristic of "overview first, zoom and filter, and details on demand" [244]. As Bower noted, "when people look at a graph, they look at the big picture first and then they start scrutinizing it." Participants found that sonification and textual description complement each other by effectively conveying overview and detail, respectively. Mustill-Rose noted that the textual description gave him the min, max, and average values, which are "hard if not impossible to get from sonification." On the other hand, Perry enjoyed the ability to sonify "trend lines in the data without having to go point by point." Since modalities afford different kinds of information-seeking operations, participants sought to choose the right modality for the task at hand. Switching between representations also helped participants adjust their initial assumptions about the data. For instance, Mustill-Rose listened to the sonification first and initially hypothesized that the stocks dataset contained only one data point per year. Then, he noticed that this was not the case when he explored the textual representation. He reflected that "the lesson there is to not consider just one modality."

Synchronized query predicates help users share context between modalities.

Participants valued the ability to maintain a shared query predicate while switching between modalities, which crucially helped them think of the modalities as different ways of looking into the same underlying data. Mustill-Rose described the synchronization across modalities as an "enabler" in the sense that "it's decreasing the time that it's taking me to get the data [from] the [time] period that I need" before he then "switch[es] to something else to look at it in a different way." Because the system maintained his interactive context as he switched representations, he was able to stay in the flow of his ongoing analysis. This echoes prior findings that delays caused by interactive latency during data analysis can "[disrupt] fluent interaction" and cause people to lose their train of thought during exploratory analysis [177].

Customization supports differences in task and experience.

Research has shown that customizable textual descriptions support users who have different preferences or tasks, allowing them to control the information they receive and how it's presented [158]. This was reiterated by Perry, Cornwell, and Erven, who encountered situations where they wanted to adjust the presence, verbosity, and ordering of information in text. I also found this customizability idea applicable beyond textual description, particularly for Jones et al.'s wayfinding and consuming affordances [158].

Wayfinding. The audio axis ticks feature supports wayfinding by helping users understand their progress through a temporally proceeding sonification. However, it trades off efficiency, and becomes less necessary over time as users get more familiar with the data. Mustill-Rose found himself wanting to disable the axis ticks after listening to a few sonifications. He said, "at first it was useful [...] but now that I know what I'm looking at, I feel like the [audio axis ticks] has proved its value. And now I don't need it anymore." However, once he selected a different subset of the data, he realized that it was "now useful again because I haven't explored this section." His need for this feature was situational throughout his analysis, depending on whether he was focusing on learning the layout of the data or "understanding and honing in on" the data. As a result, the ability to enable or disable the axis ticks was important to offer as a customization.

Consuming. Another important customization was the sonification's playback speed. As Cornwell noted, preferred screen reader reading speed varies widely among BLV individuals. For sonification, participants considered their base preference as well as their familiarity and task-specific needs. Perry and Bower both noted that their preferred playback speed was situational. Perry noted that he "would get used to it faster, but [he] would start slower because it gives more time to listen to each point." Slower speeds were better when initially learning about the data, and he would speed up as he became more familiar.

Multi-modal representations facilitate communication between people who rely on different senses.

In a multi-modal system, participants who were not primarily using the visualization still valued the presence of synchronized visual highlighting and references to visual concepts in the description. As BLV professionals who work with sighted colleagues, participants frequently encounter situations where they need to establish common ground with others who primarily use different senses.

Contributing confidently to data-driven discussions. As a software engineer who works with only sighted colleagues, Mustill-Rose stressed that an important goal of data analysis is to have enough information to "participate meaningfully in a discussion." At minimum, he said, he wanted to be in Zoom meetings and "not seem clueless," because as the only blind person on a team, consistently being the only person who can't comment on a topic can compound with unconscious bias to affect promotions and work opportunities. Erven echoed this sentiment, saying that "the most important thing is independence." Visual modalities are only helpful when "it's not something you need to rely on to do your work," forcing BLV users to rely sighted help. Instead, as previous work has also argued [299], accessible representations should promote user agency for self-guided analysis — and for BLV people not only to participate in, but also create and lead data-driven discussions.

Presenting to mixed audiences. In her job, Cornwell frequently makes presentations to majority sighted audiences. As a result, "visual charts are always useful" to her. She explained, "if I'm needing to talk about [the data], I can just say, look at the red line and the people who are really visual — that's an important source of interactivity for them." Additionally, synchronization between modalities plays a helpful role in presentation. For example, Cornwell imagined a hypothetical situation where she played a sonification while presenting, and sighted audience members followed along on the visualization. In this situation, having multiple modalities would make the presentation more accessible and also help communicate the data more effectively.

Collaborating across different levels of vision. Many participants frequently collaborate directly with others with different levels of vision. Cornwell mentioned working with someone who was totally blind, and thought that "sonification on a screen share" would be extremely valuable for communicating about data. Similarly, Bower felt that the visual highlighting of her selection in the textual structure and sonification would help a collaborator "get on the same page" and help them "know where [she's] looking." She drew an analogy to pointing at something on a visual chart, as a way of directing a collaborator's attention.

5.6.3 Qualitative Results: Structured Editor

Users want, but lack, interfaces for creating data representations.

Participants have existing strategies for working with data that primarily involve spreadsheets and scripting. Erven, Perry, Cornwell, and Bower reported using Excel or Google Sheets; Mustill-Rose, Perry, and Bower reported writing their own scripts in various tools, including python, octave, and matlab. However, there was consensus that these existing workflows are insufficient. Cornwell put it succinctly when describing raw data: "no one wants to read that stuff." But with the exception of Bower (who has used Highcharts Sonification Studio and SAS Graphics Accelerator), no participants could think of available tools for creating their own representations without having to write code.

End-user tools are important because they lower the technical barrier for creating representations. When comparing the Umwelt editor to writing code, Mustill-Rose said that "there's less cognitive pressure using a UI than if I was having to write code to do it." However, sometimes tools can overly complicate the process of making a simple representation. Bower said, "I don't care about instruments and timbre and all that, I just want access to a time series." Because of high up-front specification cost, some tools are too difficult to use for even simple cases. Nonetheless, Bower is interested in trying new tools for creating data representations, saying, "I'm kind of desperate for anything" that's usable and accessible.

Structured editing with default specifications reduces semantic and articulatory distance.

When participants decide to create representations, they face challenges to do with semantic and articulatory distance [146]. In HCI theory, semantic distance is the distance between a user's intentions and how these intentions are translated into the concepts provided by a user interface. Similarly, articulatory distance is the relationship between an interface's concepts and the set of physical actions a user has to take to express something in terms of those concepts.

Semantic distance. When Perry approached analyzing the penguin dataset, he initially said, "I want to compare beak length, body mass, and sex altogether — I want to see the graph for all three of these together." Though Perry knew that he wanted to specify visual encodings that would be reflected in the textual hierarchy, he did not immediately know what those encodings were. This was a problem of semantic distance, because he needed to map his goal onto the concepts provided by the user interface. Luckily, the heuristics generated a default specification for that set of fields that matched his expectations. As a result, he was still able to create the chart despite lack of familiarity with visualization concepts. However, when the system was not able to generate a default specification for Cornwell, she remarked that it was hard to figure out "which functions apply, like figuring out if I wanted it grouped by island or species." Even though she had goals in mind, it was difficult for her to translate those goals into specific encodings and field transformations. This suggests a need for future work on bridging semantic distance — for instance, by designing high-level abstractions that adhere closer to users' abstract goals, reducing the amount of translation work.

Articulatory distance. Using the editor, Mustill-Rose remarked that "if I were writing code, I'd need to think about what the end result was and what code I needed to write to achieve it at the same time." Rather than having to remember the names of functions and expressions in a textual language, Mustill-Rose was able to use commonplace HTML input elements that express atomic edits to a specification as simple button clicks or dropdown selections. However, the editor also has limitations when it comes to articulatory distance. Cornwell, who had created a chart that was not a default specification, noted that a main challenge was that "when you're looking at the fields, you have to add encodings for everything you want." When specifying multi-modal representations, there can be a lot of repetitive operations to create three outputs that are conceptually similar.

Users think in both field-oriented and encoding-oriented terms.

Throughout the specification process, I observed that participants moved between field-oriented and encoding-oriented ways of thinking. For most participants, the tendency was to begin by identifying a set of fields they were interested in. For instance, Erven commented that it felt natural to begin by "choosing the fields you want," since you "might not want all that data." When default specifications matched their expected representation, or when they only required minor edits, participants were generally content with the result that they achieved through field-oriented specification.

However, when more manual editing was required to achieve the desired output, I found that participants shifted more toward encoding-oriented specification as they envisioned specific output representations. Cornwell initially stated her goal by saying, “I want to know what species are on which island and then I want a sex distribution.” At this point in his process, she had not committed to any encoding properties or specific modalities, but was envisioning the semantics and structure of the data in terms of relationships between fields. After selecting the relevant fields, she began to add encodings, and then became somewhat stuck. When prompted to reiterate her goal, she said that she wanted to create a “bar chart with island on the x-axis and count for species for the y-axis.” At this point further into the process, she had imagined a specific visual representation, which she was attempting to decompose into encodings and then map onto editor operations.

Interestingly, Bower — who is familiar with both visualization and sonification — had a mental model that blurred the dichotomy between field- and encoding-oriented specification. She initially approached the Seattle weather dataset by selecting `date` and `temp_max`. When she tabbed down the editor to read the default specification, she noticed that Umwelt had assigned `y` and `pitch` encodings to `temp_max`. Based on her extensive previous experience with data visualization and sonification, she remarked, “I merge those in my head – I think of those as the same thing.” This suggests that even when thinking in encoding-oriented terms, Bower was reasoning about the data’s underlying key-value semantics.

5.7 Discussion and Future Work

I contribute Umwelt, an accessible authoring environment designed to de-center the visual modality in data analysis. Umwelt allows users to specify data representations, including visualization, structured textual description, and sonification, using a shared abstract data model. Unlike existing tools, Umwelt does not rely on an existing visual specification, affording users more flexibility in prototyping multi-modal representations. The editor’s state is reflected in independent visual, textual, and sonification views linked through shared interactions, encouraging complementary use of multiple modalities. In this section, I discuss potential directions for future work surfaced by Umwelt.

5.7.1 Designing Natively Non-Visual Data Representations

Differences in a representation’s modality affect how information is presented to a user, and the operations the user needs to perform to access the information. For example, a screen reader must “explicitly linearize reading a visualization” in order to narrate elements one at a time [299] — in contrast to how visual perception enables a user to move around parts of a visualization relatively freely. Similarly, researchers have compared tactile perception to “reading a map through a small tube” [126, 299]. An implication of these modality differences, as participants in the study found, is that users find some modalities inherently more suited to certain tasks than others. Further, it suggests that due to medium-specificity, it is not always possible to directly translate a data representation from one modality to another while maintaining 1:1 support for the same set of tasks.

However, existing systems for authoring non-visual representations largely attempt to directly translate source visualizations into standalone non-visual replacements. For example, while Highcharts Sonification Studio [55] successfully translates single-series line charts into equivalent sonifications, this approach breaks down for scatterplots. This is because the way a sighted user reads a scatterplot has no unambiguous analogue in the medium of sonification, which imposes a linearized traversal order over the data. In contrast, Umwelt's default specification heuristics pair scatterplots with sonifications that diverge from the visualization by adding additional binning and aggregation, in order to prioritize conveying the data's 2d distribution.

In the context of multimodal representations, Umwelt advances the idea that a representation should prioritize fit with its modality's affordances over fidelity to the visual representation. This conceptual orientation has implications for the design of future non-visual representations. For example, current approaches to tactile charts largely focus on converting visual channels to tactile ones while otherwise faithfully recreating the visualization [89]. Instead, future work could explore tactile-first designs that make more intentional use of the processual, part to whole [125] nature of tactile perception.

5.7.2 Interdependence and Relational Dimensions of Access

In designing Umwelt, I advocate for a conceptual shift in the field of accessible data visualization — focusing not only on making existing visualization accessible to BLV readers, but also on empowering BLV data analysts to independently produce their own representations and conduct self-guided data exploration. Because existing approaches that center the visual modality can sometimes create barriers or reinforce BLV users' dependence on sighted assistance, I believe a focus on independence to be an important step forward. However, in addition to independence, disability scholars have advanced *interdependence* as a complementary conceptual frame [37]. An interdependence frame acknowledges that all people constantly depend on others, and so a focus on relationships is necessary to understand how access is socially created in practice.

My initial evaluation of Umwelt surfaced ideas that suggest the need for interdependence (alongside independence) as a lens for design. For instance, participants felt that building common ground between mixed ability colleagues in workplace settings was important to their career advancement, highlighting the fact that BLV people's access needs are embedded in a social and relational context. Future work can motivate and inspire system design based on not only how BLV users can get immediate access to information in data, but also how they hope to use that information to participate in broader social processes.

Chapter 6

Discussion and Future Work

6.1 Participation and Power in Accessible Data Analysis

In developing design frameworks and designing tools for BLV screen reader users, the lens of participation and power was key in helping me understand the core issues of this problem domain. Previous approaches focused on making data visualizations accessible through standard approaches involving alt text and data tables. Chapter 4 contributed a set of design dimensions for applying screen reader affordances to the design of rich textual interfaces with structured navigation. Applying a lens of participation and power to this work highlights the importance of screen reader experiences that do not merely reproduce the data, but provide affordances for data exploration that are comparable to the reading experiences of sighted users. As a result, this work was able to help users conduct self-guided data exploration and reduce the degree to which BLV users are dependent on the editorial decisions of alt text writers in interpreting data.

While this contribution would be considered successful on its own by many standards, reflecting on participation and power uncovered new research opportunities. Often, work in accessible data visualization is about making existing charts more accessible. However, for blind and low-vision (BLV) people to equally participate in data analysis, they must not only be able to consume data representations created by others. They must also be empowered to create their own custom representations. This framing drew attention to limitations of existing systems which prioritize visual specification. In chapter 5, we designed a system that prioritizes the need for BLV users to be active producers of data analysis rather than passive consumers. With this system, BLV users are able to create their own multimodal data representations, furthering the goal of independent data exploration.

6.2 Future Work

This work suggests an opportunity for future work to begin radically rethinking the field of data visualization by de-centering visual representation. Instead, the field might consider visualization as one of many ways data analysts structure data to make it perceptible to the senses. The systems introduced in this dissertation, Olli and Umwelt, can serve as platforms for this future research. A multimodal approach to data representation might entail the following directions for research:

Modality-specific interaction design. Visualization researchers have identified abstract typologies of interaction techniques that represent fundamental operations in interactive data analysis [291]. While it is well understood how to instantiate these interaction techniques in visualization systems, it is less clear how designers might translate interactive operations into non-visual interfaces, enabling screen reader users to conduct comparable kinds of interactive analyses using modality-specific affordances.

Empirically deriving principles for effective non-visual representation. The design of data visualization systems benefits from theories of perceptual effectiveness built over time from a base of empirical evidence [200]. While there has been empirical research into non-visual perception for sonification and tactile graphics, there remain open questions about how to develop design principles to the same level of fidelity as their visual counterparts.

Designing for mixed-ability collaboration. In my work, I found that many BLV data analysts' needs are shaped by the educational or professional context they are working in. However, the implications of these social relationships on system design has yet to be explored, because most systems focus on designing for an individual analyst and do not consider the context in which they might share their insights with others. Future work may develop a further understanding of how mixed-ability collaborations happen in practice, and how these needs can inform design.

Part III

Empowered Non-Participation in Data Collection

Chapter 7

Introduction

The lack of consent or debriefing in online research has attracted widespread public distrust. How can designers create systems to earn and maintain public trust in large-scale online research? Procedural theories inform processes that enable individuals to make decisions about their participation. Substantive theories focus on the normative judgments that researchers and participants make about specific studies in context. Informed by these theories, I designed Bartleby, a system for debriefing participants and eliciting their views about studies that involved them (Chapter 8). I evaluated this system by using it to debrief thousands of participants in a series of observational and experimental studies on Twitter and Reddit. I find that Bartleby addresses procedural concerns by creating new opportunities for study participants to exercise autonomy. I also find that participants use Bartleby to contribute to substantive, value-driven conversations about participant voice and power. I conclude the chapter with a critical reflection on the strengths and limitations of reusable software to satisfy values from both procedural and substantive ethical theories.

Amidst calls for public accountability over large data-driven systems, feminist and indigenous scholars have developed *refusal* as a practice that challenges the authority of data collectors. However, because data affect so many aspects of daily life, it can be hard to see seemingly different refusal strategies as part of the same repertoire. Furthermore, conversations about refusal often happen from the standpoint of designers and policymakers rather than the people and communities most affected by data collection. In Chapter 9, I introduce a framework for *data refusal from below*—writing from the standpoint of people who refuse, rather than the institutions that seek their compliance. Because refusers work to reshape socio-technical systems, I argue that refusal is an act of design and that design-based frameworks and methods can contribute to refusal. I characterize refusal strategies across four constituent facets common to all refusal, whatever strategies are used: *autonomy*, or how refusal accounts for individual and collective interests; *time*, or whether refusal reacts to past harm or proactively prevents future harm; *power*, or the extent to which refusal makes change possible; and *cost*, or whether or not refusal can reduce or redistribute penalties experienced by refusers. I illustrate each facet by drawing on cases of people and collectives that have refused data systems. Together, the four facets of my framework are designed to help scholars and activists describe, evaluate, and imagine new forms of refusal.

Chapter 8

Designing and Evaluating Research Ethics Systems

8.1 Introduction

A series of high-profile research scandals in the past decade has led to calls for improvements and standardization in research procedures. In 2014, after public outrage about a study that altered the contents of hundreds of thousands of Facebook news feeds [170], many called for studies to include informed consent, debriefing, and a chance for participants to opt out [118]. In 2020, residents of the state of Illinois sued IBM when researchers included their online photographs, without consent, in a research dataset initially prepared by Yahoo [205, 254]. Yet surveys of research practices in social computing have found that many academics avoid consenting or informing participants because they believe it to be impractical [274].

Scholars, critics, and policymakers have argued that these research projects failed by ignoring individual autonomy [118]. By preventing people from learning about research and data collection, researchers failed to give individuals a chance to choose whether to participate in a study or choose how their data would be used. Critics argued that the needs of research and the autonomy of participants could be maintained with the right procedures in place. After all, the communication technologies that enable large-scale data collection have also enabled new design possibilities for innovations in research procedures [118]. Since then, lawmakers in the E.U. and U.S. have passed regulations that require data processors (but not academic researchers) to inform people about the data they collect and provide them with a chance to have it removed [5, 6]. In parallel, researchers have suggested that more studies include a debriefing stage, where participants are told the details of a study and given a chance to opt out [77, 118].

To advance research ethics procedures that protect individual autonomy, prevent abuses of power, and promote public trust, I introduce Bartleby: a system that delivers research ethics procedures for large-scale online studies. Bartleby provides a user interface that researchers can customize to the details of their study. Using Bartleby, researchers can automatically send each of their study participants a message directing them to a website where they can learn about their involvement in research, view what data researchers collected about them, and give feedback. Most importantly, participants can use the website to opt out and request to delete their data. The system is named after the titular character in Herman Melville's short story *Bartleby, the Scrivener*. Over the course of the story, Bartleby opts out of completing various requests. Instead, he states simply that he "would prefer not to" [194].

I designed Bartleby in response to public criticism of academic research conducted by Facebook and IBM. Because these studies lacked informed consent, people who were included in the research data likely never found out whether or not they were in the study. Scholars have suggested that the ethics of these studies could have been improved with minimal effort by debriefing participants—in other words, notifying them of their participation and offering a chance to opt out [118]. With Bartleby, researchers can automate debriefing with few adjustments to their existing research processes. By creating and deploying Bartleby in the field, I also demonstrate that large-scale debriefing can be simple and practical, despite claims to the contrary.

I also present the Bartleby system as a case study for critical thinking about the design of research ethics procedures. What does it mean for a research ethics system to be successful? Because the purpose of debriefing is to protect participants' right to autonomy, it can serve moral and procedural purposes regardless of whether any individual participant exercises their rights by opting out. For that reason, I include an extended discussion of the system in light of two kinds of ethical theories drawn from feminist and political philosophy: procedural and substantive theories. Procedural theories are concerned with the abilities and limitations of scalable, repeatable procedures to protect individual autonomy. Substantive theories are concerned with the values upheld by the research, and the use of power in deciding those values.

As a U.S.-based researcher, my frameworks for thinking about the role of autonomy in research ethics may differ from those of a global audience. Scholars have questioned the idea of universally-applicable research ethics, arguing instead that "ethical codes are never universal and are geographically sensitive" [294]. While my work is primarily informed by the institutional environment of U.S. ethics regulation and university review boards, I acknowledge the importance of respecting different values and expectations that arise when conducting research on social media platforms with global reach.

In this paper, I summarize the design challenge of creating ethics systems, describe the design of the Bartleby system, and present empirical evidence from two 2020 field studies on Twitter and Reddit involving 4,766 and 1,342 participants. I also review the design of the system through the lens of procedural and substantive theories of ethics. In addition to presenting the Bartleby system, my work demonstrates how procedural and substantive theories can guide the design and evaluation of research ethics systems.

8.2 Debriefing Participants and Opting Out of Social / Behavioral Research

Debriefing is a research ethics procedure that happens at the end of a study, after data collection has concluded [260]. During debriefing, researchers notify participants that they were involved in research and disclose information about study procedures. Participants are informed about the data researchers collected, and have an opportunity to exercise their agency by opting out and withdrawing their data from the study.

Debriefing serves a distinct purpose from other procedures. For example, informed consent is a procedure that happens before a study begins. Participants are given information about what will happen if they are involved in research, and can give or withdraw their consent based on that information. While debriefing and informed consent are not mutually exclusive, debriefing is unlike informed consent in that it can be used in study designs where participants are unaware of their inclusion in data collection. In this paper, I focus on debriefing due to the fact that this kind of research is increasingly common on social media platforms. While there is a broader ongoing conversation in research ethics about whether the increasing prevalence of large-scale data collection on unaware individuals is acceptable [290], my work starts from the premise that there exist at least some cases where valid methodological reasons prevent the use of informed consent. For instance, it is impractical to seek prior informed consent for a study that observes Twitter discussions between certain dates, because researchers cannot know who will participate in discussions ahead of time. Furthermore, knowledge about the study might also bias the behavior of subjects who are being observed, affecting the validity of the results. In some studies, consent procedures might introduce selection bias—some people might be more likely to consent than others due to group membership. Indeed, researchers in political science and philosophy have studied the moral significance of ethics procedures other than prior informed consent—such as proxy consent [144] and hypothetical consent [90]—due to their necessity in certain practical situations.

Debriefing has most commonly been used offline in fields like psychology and behavioral economics to manage the ethics of deception-based studies [130]. Because this research has usually happened in-person in a lab, participants know that they are part of research. However, they are not informed about the true nature of the research until researchers debrief them in-person, providing an immediate opportunity to ask questions and address harms. Online research creates a new situation where participants may not be outright deceived, but potentially never become aware of their inclusion in research unless debriefed. I use the term *non-consented research* to distinguish this situation from other research designs where debriefing has previously been used.

Even though debriefing is a well-understood ethics procedure, it is rarely used in non-consented online social science research [77]. In non-consented research, debriefing serves two important purposes: informing participants and creating opportunities for them to opt out. Informing research subjects of their participation, providing opt out opportunities, and opening communication between subjects and researchers all serve to increase the agency of participants over their involvement in research. For observational studies, debriefing can be thought of as retroactive informed consent. It could be argued that debriefing has equivalent moral significance to informed consent in observational studies, because data could potentially remain un-processed until consent is given and can be deleted when consent is withdrawn. However, for field experiments, post-hoc debriefing cannot undo participant exposure to interventions. Debriefing is not an equivalent replacement for upfront informed consent in field experiments, though it is one of the better alternatives when consent is not an option.

As an ethics procedure, debriefing cannot always address participant concerns about the normative content of the research. In prior empirical work on participant perceptions of research ethics, adding debriefing to study designs that lacked consent did not significantly change the perceived acceptability of research [77]. More debriefing, then, does not necessarily resolve ethical issues with studies that are objectionable for reasons other than whether or not they were consensual.

Nonetheless, scholars have argued that making debriefing more commonplace in online research would be an achievable yet significant advancement in research ethics. Writing about the Facebook Emotion Contagion study, Grimmelmann argues that "standardized debriefings could easily have been given via email or private message to the users who were unwittingly drafted into the studies" [118]. Systems like Bartleby that provide debriefing at scale contribute to research ethics by lowering the technical burden required for researchers to implement ethics procedures. Grimmelman argues that "as it becomes easier to do more for participants, researchers should, because there is less and less reason not to" [118]. Prior work has highlighted how shared tools that assist ethical compliance have benefited participants [45]. Bartleby makes it less likely that researchers will forgo debriefing due to impracticality or technical burden, which will create more opportunities for participants to exercise individual autonomy. While greater access to debriefing opportunities will not solve all problems in research ethics, debriefing improves participant agency compared to the alternative.

8.3 Bartleby: Research Debriefing System

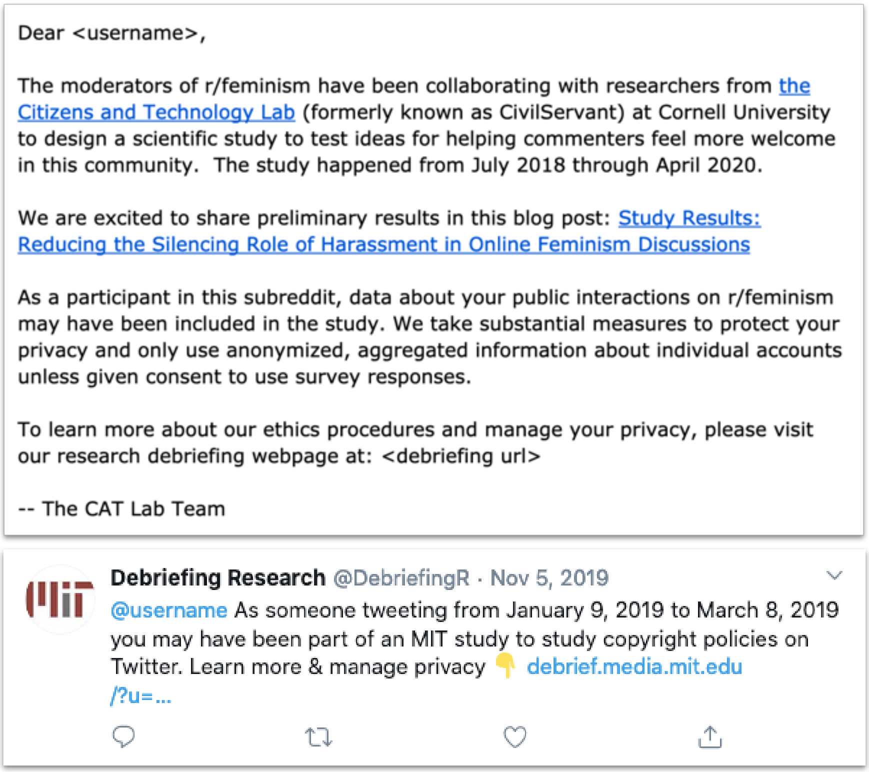
Bartleby is a system that automates debriefing for large-scale social and behavioral experiments online. Bartleby consists of a message-sending script to invite participants to start the debriefing process, and a website that provides debriefing information and an opt out form. It is available as open source software under an MIT license at: *https://github.com/jonathanzong/bartleby.*

8.3.1 Design Values

HCI designers and researchers routinely grapple with the values, ethics, and politics of technologies [243]. In the process of creating Bartleby, I listed values that I think research ethics systems should be designed toward. Different systems may reflect these values in different ways, and to varying degrees. Whatever their goal, the designers of any research ethics system will encounter questions of informedness, agency, and scale.

Informedness. The goal of debriefing procedures is to provide people with the capacity to make an informed decision about research participation. Informedness is thought of as a state of understanding that people can achieve given enough information and guidance [222]. Debriefing interfaces can inform participants about the purpose of the research and what data was collected about them. They may also provide ways for participants to ask questions to further guide decision-making. The process of informing and clarifying is normally facilitated by informed consent, which makes it essential to debriefing in non-consented research.

Agency. Debriefing procedures provide participants with control over their involvement in research. Participants in debriefing must be able to withdraw from the experiment by opting out. They can also give feedback to researchers and address any harms that may


Dear <username>,

The moderators of r/feminism have been collaborating with researchers from the Citizens and Technology Lab (formerly known as CivilServant) at Cornell University to design a scientific study to test ideas for helping commenters feel more welcome in this community. The study happened from July 2018 through April 2020.

We are excited to share preliminary results in this blog post: Study Results: Reducing the Silencing Role of Harassment in Online Feminism Discussions

As a participant in this subreddit, data about your public interactions on r/feminism may have been included in the study. We take substantial measures to protect your privacy and only use anonymized, aggregated information about individual accounts unless given consent to use survey responses.

To learn more about our ethics procedures and manage your privacy, please visit our research debriefing webpage at: <debriefing url>

-- The CAT Lab Team

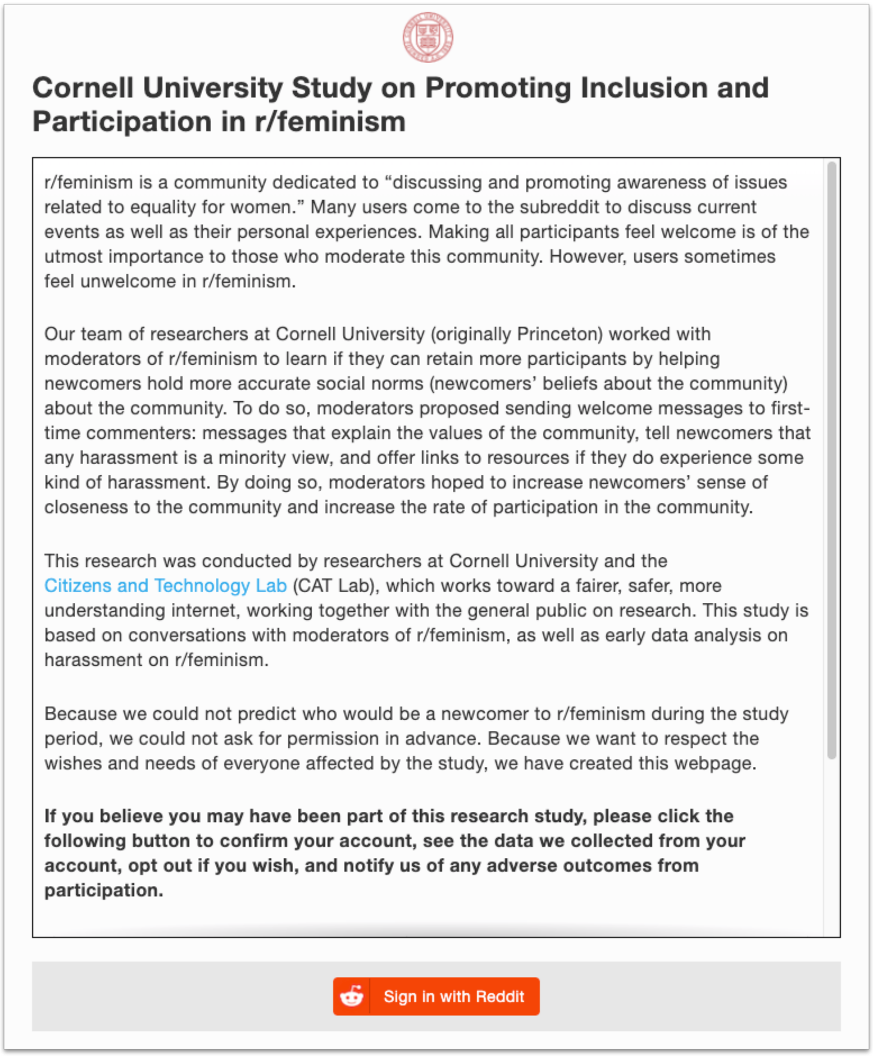
Debriefing Research @DebriefingR · Nov 5, 2019

@username As someone tweeting from January 9, 2019 to March 8, 2019 you may have been part of an MIT study to study copyright policies on Twitter. Learn more & manage privacy 👇 debrief.media.mit.edu/?u=...

Figure 8.1: *Debriefing invitation messages sent to participants eligible for debriefing. The messages included a link to the debriefing login page for each study.*

arise. These mechanisms of accountability ensure participants can exercise their right to individual autonomy. University IRBs also typically provide language in consent forms that lets participants know that they can contact the IRB if they suspect researcher misconduct. Debriefing procedures can provide pathways for participants to access other, more powerful procedures if necessary. Even if these other procedures are never activated, making them available to participants can increase their agency.

Scale. As more people are included in online research, researchers interested in exploring the potential for large sample sizes to contribute to knowledge must also think about how to protect the autonomy of large numbers of people. As Gillespie notes, scale is more than size; "scale is about ... how a process can be proceduralized such that it can be replicated in different contexts, and appear the same" [110]. In my work, I see questions of scale arising whenever ethics procedures are called on to manage the autonomy of more people than a research team could interact with on an individualized basis, without the help of automation. Ethics procedures can outline standard practices that can be repeated efficiently and reliably for many people, across many studies. Procedures must navigate the tension between being customizable enough to adapt to different people and studies, while being standardized enough to be reusable.



Cornell University Study on Promoting Inclusion and Participation in r/feminism

r/feminism is a community dedicated to "discussing and promoting awareness of issues related to equality for women." Many users come to the subreddit to discuss current events as well as their personal experiences. Making all participants feel welcome is of the utmost importance to those who moderate this community. However, users sometimes feel unwelcome in r/feminism.

Our team of researchers at Cornell University (originally Princeton) worked with moderators of r/feminism to learn if they can retain more participants by helping newcomers hold more accurate social norms (newcomers' beliefs about the community) about the community. To do so, moderators proposed sending welcome messages to first-time commenters: messages that explain the values of the community, tell newcomers that any harassment is a minority view, and offer links to resources if they do experience some kind of harassment. By doing so, moderators hoped to increase newcomers' sense of closeness to the community and increase the rate of participation in the community.

This research was conducted by researchers at Cornell University and the Citizens and Technology Lab (CAT Lab), which works toward a fairer, safer, more understanding internet, working together with the general public on research. This study is based on conversations with moderators of r/feminism, as well as early data analysis on harassment on r/feminism.

Because we could not predict who would be a newcomer to r/feminism during the study period, we could not ask for permission in advance. Because we want to respect the wishes and needs of everyone affected by the study, we have created this webpage.

If you believe you may have been part of this research study, please click the following button to confirm your account, see the data we collected from your account, opt out if you wish, and notify us of any adverse outcomes from participation.

Figure 8.2: *Experiment-specific login page for the debriefing software. People can log in with their social media account to verify their inclusion in the study and view debriefing information.*

8.3.2 Debriefing User Experience

The Bartleby user flow has three steps: the invitation message, the login page, and the debriefing page.

Invitation message. Participants enter the debriefing process when they receive an invitation message notifying them about the research study (Figure 8.1) and inviting them to debrief. The message includes basic information about the purpose of the study and a link to the debriefing website. If the participant clicks the link in the message, they arrive at the login page for the study.

Login page. Each study that uses Bartleby for debriefing will have its own login page. On this page, participants can view a more detailed explanation of the research question, who the researchers are, and why they are being debriefed (Figure 8.2). At the bottom of the page, they can click the button to log in with their social media account. Bartleby uses the Reddit and Twitter APIs to log users in. This means that the platforms handle authentication, and Bartleby only handles a username and public account information once authenticated. Once logged in, the system matches their username against a list of study participants stored in a database. If the account attempting to log in is not in this list—for example, if a study participant forwarded their debriefing message to an account that was not included—that account is ineligible for debriefing. Ineligible participants are redirected to a page stating that they were not included in the study. However, eligible participants will be redirected to the debriefing page containing information pertaining to their account (Figure 8.3).

Debriefing page. The debriefing page confirms the participant's involvement in the study and documents the specific actions that researchers took (Figure 8.3A). Participants can view a table showing exactly what data the researchers collected (Figure 8.3B). The table comes with a description of why the data collection was necessary, and how data will be used in the future. Below the table, there is a section on opting out of the study. Participants can read information about the effects of opting out, and decide whether to check the opt out checkbox (Figure 8.3C). They can return and update their decision at any point before the date listed at the bottom of this section. The cutoff time is intended to allow researchers to submit their work for publication without constantly needing to check back. If participants are curious to know the results of the study, researchers may include instructions for following up (Figure 8.3D). This is especially useful for studies where results have practical value for the studied population. In the case that participants feel that they have experienced harm from being included in the study, the debriefing interface includes contact information for the university IRB (Figure 8.3E). Finally, participants can fill out an optional survey to give feedback on the study and the debriefing interface at the bottom of the page (Figure 8.3F).

8.3.3 Researcher User Experience

Bartleby requires some initial configuration to set up a database, store database credentials in config files, define API keys to interface with Reddit and Twitter, and set up a public-facing web server to serve the login and debriefing pages for each study.

To use Bartleby to debrief a study, the researcher populates the Bartleby database with records of people who were included in the study. These records include the minimal amount of information required to debrief: the user id on the platform where the study occurred (i.e. Twitter handle or Reddit username), and the user's associated study data.

The researcher also defines templates for the invitation message, login page, and debriefing page. These templates are specific to a particular study. For example, the login page template should describe the study, and the debriefing page template should include information on how to report ethics violations at the researcher's institution.

When the researcher is ready to debrief, they use a script to send an invitation message for each user record in the database. When a user logs in, their debriefing page template will be filled in with the study data associated with their username in the database.

8.3.4 Supporting Multiple Studies

Because I intend for researchers to use the system on an ongoing basis, Bartleby supports debriefing for multiple concurrent studies. Researchers can define experiments in the Bartleby database, and associate eligible participant accounts with those experiments. Each experiment has its own base URL on the web server. Base URLs start with a randomly generated unique identifier, so that the existence of other experiments cannot be revealed by sequentially guessing URLs from a known experiment. Each experiment is also associated with a template directory, which contains template files to render for the experiment's landing and debriefing pages. In these template files, researchers can customize the language for each experiment page. When a participant logs into a debriefing page, Bartleby renders the data collected on that account for that study into a table on the page.

8.3.5 Data Removal Procedure

In digital systems, data collection and circulation are usually ongoing over long periods of time. Because people's preferences change over time, maintaining consent over time is a general problem in the use of research data, as datasets are shared publicly by researchers and adapted for new and unexpected purposes. In my field deployments, I found that everyone who engaged with the system did so in the first few days after receiving the debriefing invitation and never returned. In practice, this method does give people some flexibility to change their mind but encounters difficulties upholding autonomy past a certain amount of time.

When people opt out of my studies, I overwrite the entries in the research dataset for that person so their data is not included and researchers retain a record that a participant opted out. I also mark people as opted-out in email lists and other databases to prevent them from being included in future analyses or communications related to the study.

How other researchers handle data removal procedures for their own studies will depend on contextual judgments of risks and benefits to participants or society. Risks will be influenced by the sensitivity of the data and possibility of de-anonymization. Analysis of benefits will need to consider the minimum amount of data needed to facilitate scientific reproducibility. In the event that certain sub-populations are more likely to opt out, data removal could potentially introduce skew affecting the validity of the results. In this case, researchers must consider whether and how to communicate about uncertainty while respecting participant privacy preferences.

8.4 Field Deployment of the Debriefing System

To test Bartleby's effectiveness, I used the system to debrief two large-scale causal studies of online behavior. For both studies, I used the Bartleby system to message participants about their involvement in the study after it was complete. Participants could log in to the Bartleby system to receive more information about the data researchers collected and its intended use. They could also choose to opt out of the study and provide further feedback.

Table 8.1: *Accounts Participating in Twitter Debriefing*

Period	Participants	Contacted	Logged in	**Login rate**	Opted out	**Opt-out rate**
8/25—11/05/2019	4766	3631	3	0.00083	1	0.00028

8.4.1 Debriefing an Observational Study on Twitter

In a large-scale observational quasi-experiment on Twitter, my collaborators collected 5,171,111 public posts made by 9,818 accounts that had received legal action for allegedly violating copyright law [64, 189]. This study was designed by a group of legal scholars, social scientists, and computer scientists based on prior ethnographic and survey research, without consulting participants about the specific study design. To identify eligible accounts, they scraped public records of legal notices, linked them with specific Twitter account IDs, and queried the public Twitter API to retrieve their public statements. They collected public tweets over the course of the 23 days before and after they received the notice. To support analysis, they then created an anonymized, aggregated record of the number of tweets per day for each account, removing reference to specific days from the final dataframe. The final analysis examined differences in the daily rate of tweets before and after receiving a legal notice.

They developed this observational study because they believe a randomized trial would violate the principles of beneficence and justice. A field experiment that randomly assigned people to different law enforcement conditions could disproportionately expose some people to tens of thousands of dollars in legal penalties. Because they were not assigning people to receive the intervention, they could not consent people before they received a legal notice. Furthermore, they hypothesized that receiving a notice would cause people to participate less on Twitter. If they had sought consent afterward, their requests might only be seen by people who were not deterred by legal action, leading them to mistakenly under-estimate the damage of copyright enforcement to people's participation online. This study was reviewed by the MIT IRB, who granted them permission to waive informed consent and required them to debrief participants. They were also required to store all public Twitter data and legal notices in an encrypted datastore.

I used Bartleby to send debriefing invitation messages to a random sample of 4766 study participants. The message sender script sent tweets from an account that presented itself as a research debriefing account, with a university logo as its profile picture. The tweets @-mentioned the participant and included a link to the Bartleby page for the study. I did not use direct messages because most Twitter accounts are configured to only receive direct messages from accounts that they follow. While these tweets are publicly viewable on the debriefing account, they will not appear in the timelines of recipients' followers unless those followers also follow the debriefing account. The additional privacy risk the debriefing account introduces is small because there are already public records databases of copyright notices. Because the study is relatively low risk, my collaborators and I decided the benefit of sending debriefing invitations was worthwhile for this particular study. Out of the 4766 accounts I designated for debriefing, numerous accounts were not contactable because they had been suspended, deleted, or had otherwise become unavailable in the time between data

Table 8.2: *Accounts Participating in Reddit Debriefing*

Period	Participants	Contacted	Logged in	**Login Rate**	Opted out	**Opt-out rate**
6/25–07/31/2020	1342	1177	10	0.0085	3	0.0025

Table 8.3: *Direct Message Replies to Reddit Debriefing Invitation*

Period	Participants	Contacted	Replies	Unique users	**User reply rate**
6/25–7/06/2020	1342	1177	23	22	0.019

collection and debriefing. I successfully sent debriefing invitations to 3631 accounts. As reported in Table 8.1, 3 accounts logged into the debriefing system and 1 opted out.

During the debriefing process, I became aware of Twitter's "Quality Filter" feature, which filters notifications from "duplicate Tweets or content that appears to be automated" [87]. It is possible that some participants did not receive notifications about the debriefing tweet. However, the filter would not have affected all participants, and the algorithm's exact criteria are opaque. Because this feature would affect all attempts at debriefing on Twitter, my results are still of value for understanding Twitter debriefing and opt out rates.

8.4.2 Debriefing a Field Experiment in a Reddit Community

I also used Bartleby to debrief participants in a field experiment hosted by an online feminism discussion community [63]. In this study, which included feminist and anti-feminist participants, my collaborators randomly assigned first-time commenters to a control group or to an intervention group that received a private message. Several weeks later, they both collected data on public participant activity and sought consent for participation in a survey. After an observation period, they created an anonymized, aggregated dataset of participant activity in the community and merged it with survey responses. They then conducted analyses within the full observation sample and within the subset of participants that completed the survey. Finally, I used Bartleby to send participants a link to a debriefing experience, providing them an opportunity to opt out.

In this study, the researchers followed procedures of co-design and participatory hypothesis testing [187, 188], where community representatives were involved in the research process from inception to debriefing. They held a day-long workshop with community representatives to identify the general research area, co-designed the study over several months with the community, and presented the final study design for approval by the community before submitting it to review by the Princeton and Cornell University IRBs.

In the final design, the researchers decided to use a debriefing process for participants who experienced the intervention and whose public activity was observed for the study. Researchers and community representatives decided together on debriefing rather than consent because the intervention was designed to support first-time participants, because the minimal, short-term risks were reversible, and because it was impossible to anticipate who would participate before the study began.

Using Bartleby, I sent direct messages to all 1342 study participants. These debriefing invitation messages included study information and a link to the Bartleby login page for the study (Figure 8.1). Out of the 1342 participant accounts, a small number of accounts had been deleted in the time between data collection and debriefing. I successfully sent debriefing invitations to 1177 accounts. As reported in Table 8.2, 10 accounts logged into the debriefing system and 3 opted out.

In addition to receiving feedback from those who logged into the debriefing system, I also received 23 direct messages from 22 participants via replies to the debriefing invitation message (Table 8.3). Some messages were positive, thanking researchers for conducting the study. Others included harsh criticism and profanities. I observed that many messages to the debriefing account were directed at the moderators of the Reddit community, whom participants did not distinguish me from as a researcher.

8.5 Discussion

8.5.1 Procedural and Substantive Theories in Research Ethics

How academics think about research ethics is shaped by the underlying ethical theories they are working with. For example, when U.S. institutions established their approach to research ethics with the Belmont Report, they were responding to human rights violations through the lens of existing theories of research ethics [270]. Mid-20th century ethicists described research ethics as a balancing act between the individual autonomy of participants and the common good that scientists were expected to pursue [222]. Consequently, when developing the model of research ethics in the U.S., the authors of the Belmont Report cited principles protecting individual rights including respect for persons. They also advanced principles that guided scientists toward the common good, including beneficence and justice [270]. By working from principles of individual autonomy and the common good, the authors of the Belmont Report were able to develop guidelines, regulations, and bureaucracies that govern research ethics in the U.S. to this day.

Feminist philosophers Mackenzie and Stoljar have identified two overarching kinds of ethical theories at play in discussions of autonomy and the common good: procedural theories and substantive theories [181]. These theories have been particularly important for feminist ethics, which has sought to reconcile individual autonomy with the collective concerns of structural oppression. Similar distinctions have been made between theories of justice in political philosophy [218], and between procedural ethics and "ethics in practice" in medicine [119]. *Procedural theories* focus on the details of procedures—standardized, repeatable steps that can be automated by a software system or a bureaucracy—that protect individual rights, such as a right to individual autonomy. Procedural theories are often called "content-neutral" because these theories do not treat the content of a person's specific actions and decisions as relevant to whether they are moral. Procedural theories focus on how well those actions and decisions happened within the structure of pre-defined ethics procedures that protect autonomy, such as informed consent. *Substantive theories* focus on the idea that procedural theories "must be supplemented by some non-neutral condition" [181]. In other words, substantive theories argue that the actions and decisions people make, even within procedures such as informed consent, must account for moral ideas such as the common good—ideas that cannot necessarily be standardized into a procedure. For example, if a participant gives informed consent to a study that is against their own interests or that poses a threat to others in society, substantive theorists would question whether that consent is sufficient moral justification for the study to happen.

Researchers and their institutions apply procedural theories to research ethics when they design and implement ethics procedures. For example, researchers comply with regulatory requirements by submitting their plans to an institutional review board (IRB) for review. IRBs will often require researchers to implement specific procedures, such as informed consent, designed to protect participant autonomy and enforce oversight processes. When U.S. researchers "believe that approval by institutional review boards (IRBs) is sufficient for addressing ethical considerations," [59] those beliefs about research ethics can be described in purely procedural terms. When regulators mandate certain procedures and institutional arrangements, the power of researchers over participants is guided and restrained by researchers' focus on procedural compliance. When new needs arise as research methods change, regulators revise policies governing ethics procedures—such as in 2018 when the U.S. government made changes to the Common Rule [271].

Procedural theories can guide valuable progress in research ethics. University IRB staff have struggled with the fact that the "Common Rule ... does not provide appropriate guidance for the realities of research with online data" [275]. Even in the absence of institutional requirements, online researchers and their participants can benefit from improvements in the design and use of ethics procedures for large-scale behavioral research online. For example, many offered procedural criticisms of the Facebook Emotion Contagion study, arguing that it should have included consent or debriefing [118]. By introducing a large-scale debriefing procedure using a system like Bartleby, the researchers would have made the morally significant improvement of offering an opportunity to exercise autonomy where none was previously afforded.

Researchers apply substantive theories to research ethics when they have value-driven conversations about a study's content, design, and other ethically-relevant issues—regardless of any procedures employed. For example, a substantive ethics conversation on the Facebook study would cover topics including mental health risks, collective risks associated with large-scale attempts at social influence, and the nature of people's relationship with Facebook. These substantive conversations often depend on the content and context of a specific study. These conversations are also affected by potentially different views about normative concepts (such as the common good) held by participants and researchers.

Even when researchers and participants agree on the importance of normative concepts in substantive ethics (such as harm, beneficence, and justice) they can still disagree on how to understand those concepts and their relative importance. As prior work has shown, participants are active agents, with their own agendas separate from researchers' plans [142]. As these conflicts are ignored or negotiated, the power structures and power imbalances that enable those moves are also a concern of substantive ethics. Researchers applying substantive theories would ask how participants are able to exercise voice and power (if at all) in normative discussions and decisions about the design, implementation, and uses of research.

While substantive theories help people ask context-specific questions about a study, these theories also enable conversations about power relations between participants and researchers that can apply across multiple studies. Guided by these substantive questions, researchers and communities have worked to redesign how interpersonal and institutional power relations are structured in research [188]. Participatory research, co-design processes [230], refusal [36], and empirical work on participant expectations [77, 96] all provide ways to surface substantive issues in research. Because none of these models can prescribe solutions to normative questions in research, they should not be treated as procedural checkboxes.

Just as researchers work to balance or weave together individual autonomy and the common good, they should also see procedural and substantive theories as complementary. Rather than supplanting each other, these theories provide resources for combining equally-important considerations inherent in research ethics. For example, substantive theorists make an important critique that the content-neutrality of procedural theories provides a necessary but insufficient account of the ethics of a situation. Rather than abolish procedures that respect individual autonomy, researchers should do further work to consider questions of power.

Because designing systems that automate ethics procedures often involves creating structures of power [286], attempts to design new systems for research ethics can benefit from both procedural and substantive ethical theories. In this paper, I use the two kinds of theories to conduct complementary analyses that situate the Bartleby system within a broader design space of research ethics systems that might balance individual and collective concerns differently.

8.5.2 Procedural Ethics Evaluation of Bartleby

When researchers ask whether introducing debriefing procedures via Bartleby increases the ability of participants to exercise their individual autonomy, they are applying procedural theories to the evaluation of research ethics systems. Bartleby protects autonomy by offering each participant access to the information and interface controls required to make an informed opt out decision. In most non-consented online research, research participants have no opportunity to exercise autonomy over their participation. They are often never even told that researchers collected their data. Because Bartleby makes it easier for researchers to implement and scale debriefing, the system creates new opportunities for participants to exercise autonomy that did not previously exist.

In my field deployments, I found that some participants logged into the system and made an active choice to either remain in the study or opt out. Because procedural theories are content-neutral, the debriefing system is considered successful in procedural terms regardless of what their choice was. It does not matter whether they opted out or remained in the study; the system is successful because they made an informed choice that otherwise would not have been possible.

I argue that Bartleby protects participant autonomy even if no people use the system. To illustrate this with an example, consider the case of a facial recognition dataset where people can have their images removed on an opt out basis. Imagine that everyone is offered a reasonable opportunity to opt out, but nobody chooses to exercise that ability. Now imagine a different project where nobody is ever offered the ability to opt out. The outcomes are the same (no one opts out), but the protections to individual autonomy in these two situations are very different. The fact that Bartleby offers participants a choice that would not otherwise be available is morally significant to the ethics of the study.

Opt out procedures are successful if the people who would have opted out, when given the best opportunity to do so, actually do opt out. Because opt out rates must be interpreted in the context of both risks and benefits to participants, and the overall size of the study, there is no normatively desirable opt out rate. For a given study, low participation in debriefing could accurately reflect participant preferences (especially for studies with low risk). I believe this is the case for my studies, which collected low-risk data and are motivated by reasonable common good arguments. The opt out rates observed in my studies may also not be considered low for a study with more total participants, because the absolute number of people who opt out would be much higher. Bartleby can also be used with studies that pose greater risk to participants. In these cases, I might hypothesize that opt out rates will be higher.

Ethics procedures may introduce barriers to autonomy if a system does not do enough to include people in the process. The counterfactual of whether or not people would have opted out under other circumstances is difficult to measure, because it is difficult to study non-participants. For example, people who declined to use Bartleby would likely not respond to a survey about why they declined. In the case of non-consented studies that would use Bartleby, the alternative to debriefing is not informed consent but rather an absence of procedures that protect individual autonomy.

Although my university ethics board received no complaints about either study during my debriefing tests, Bartleby would also be considered successful if people had reported me for researcher misconduct. Guided by procedural theories, I consider a procedure successful when accountability mechanisms are available and activated when needed. If people had complained about the study after hearing about it through Bartleby, or if there had been a public scandal because people found out about the study, or if the study had been shut down, or if I had lost my research position—all of these would have been procedural successes.

In addition to improving the implementation and adoption of ethics procedures, the Bartleby system could also support empirical research on the design of ethics interfaces. Researchers using Bartleby could, for example, conduct field experiments that vary features of the debriefing interface to test their effectiveness and usability. Researchers might also interview participants to learn about the relationship between the content of research and the effectiveness of ethics procedures—for example, whether more people opt out of studies perceived to be higher risk. As consumer privacy regulations like the GDPR continue to prioritize consent procedures, research from Bartleby and related systems could influence the design of ethics procedures beyond academic research.

8.5.3 Substantive Ethics Evaluation of Bartleby

Although the two studies in my field deployment used debriefing procedures in the same way, they were different in significant substantive ways. For instance, the participants' prior relationships to each other and to the researchers differed between the two studies. A procedural account of these two studies would describe them very similarly—through a procedural lens, the studies used the exact same debriefing process. The lens of substantive ethics reveals differences in the exercise of power between participants and researchers that may lead to differences in normative judgments on the ethics of those studies.

For researchers thinking in terms of substantive theories, the normative content of the research is important to ethics—not simply the content-neutral procedures that were implemented. If researchers and participants both agree that benefits of the study outweigh the risks of using the data, the research might proceed uncontested. However, interpretations of risk and benefit are subjective. When researchers make decisions without involving participants, they wield a large amount of power over participants through this exclusion. In my analysis, I narrate the different substantive ethics concerns surfaced from the two studies where I tested Bartleby.

In the Twitter study, which included individuals who have experienced copyright takedowns, researchers did not have access to organizations or entities who could speak and act on behalf of participants. Consequently, they were unable to include participants in the design of the study or account for their voices and perspectives beyond evidence from exploratory qualitative research with people similar to those in the study. As a result, the Bartleby system offered the study's only opportunity for input or power from research participants.

Since few people in the study responded to debriefing, my collaborators and I have relied on our own intuitions and the oversight of our IRB boards to address substantive questions. The Twitter study was observational, collected minimal data, and used appropriate anonymization and data storage practices. Harms from unintended disclosure of our research dataset would also be minimal—the count of tweets someone made in a time period is not sensitive in the same way that, for example, medical information is sensitive. As a result, I can reasonably argue that the study is minimal risk. The study also has a reasonable common good argument, as knowledge about the effects of automated legal notices on the exercise of speech rights could inform future policy discussions. Because the study is relatively innocuous, I might be less concerned that Bartleby had low usage.

In contrast, the Reddit study involved a community in conflict—a large feminism discussion community where anyone can join discussions and where anti-feminists attempt to disrupt conversations. People who participate in this community accept governance by community, so my collaborators worked with moderators to co-design research questions and study procedures. Because moderators understand the norms and preferences of the community, researchers obtained what Humphreys calls *proxy consent* [144], asking moderators to grant consent on behalf of the community they represent. Since community representatives reviewed, influenced, and approved the study design, they contributed to decisions about substantive ethics concerns. Independently of the number of people who used Bartleby, the community representatives' knowledge of norms and values gave me more confidence that the research was respecting and managing risk for all participants, including those who did not debrief.

Working with community members is a helpful way to surface substantive issues in research, but participatory research should not be understood as just another one-size-fits-all procedure. Community-based proxy consent, for example, is not sufficient for all studies because communities are usually heterogenous. It's not always clear how researchers should position themselves when different groups within the community disagree [164].

In the Reddit study, debriefing enabled me to maintain respect for individual autonomy alongside community engagement. The study included people who self-identify as feminists, people who do not, and people who identify as anti-feminists. Among those who identified as feminists, some did not identify closely with the community, arguing that the community's moderators align with ideological positions that they do not agree with. The research co-design process was also unable to include the views of anti-feminists whose purpose was to disrupt the community and cause them harm. The researchers held the normative position that harassment and disruption were not legitimate goals to uphold in the research. Consequently, when deciding to conduct research focused on protecting community from harassment, they committed to a power structure that aligned them with the community's moderators. But they also wished to respect individual autonomy.

Within this complex situation, the Bartleby system provided opportunities for voice and agency among individuals who disagreed with researchers' ethical judgments and those of the community. When debriefing participants, I learned that some people disagreed that the moderators' use of power was legitimate. Several participants who had been banned by moderators sent private messages to the Bartleby system complaining about their treatment by the community. When they saw the study and were included in its procedures, they were reminded of community moderators' continued power over them. They then used the debriefing system to object not to the study, but to that deeper structure of power. Procedurally, those messages were irrelevant. Substantively, they were essential.

These messages also suggest potential risks to researchers arising from increased visibility due to the use of systems like Bartleby. Although U.S. research ethics regulation was written to protect participants from abuse of power by researchers, internet scholars have written about the risk of online communities conducting organized harassment and abuse against researchers. Writing about the challenges of researching far-right online spaces, Massanari notes that "[power] asymmetry is, in part, due to the visibility of those being targeted and the relative invisibility of those who are perpetrating the attacks" [186]. Indeed, researchers have chosen not to debrief research subjects for reasons ranging from "a demonstrated propensity for online harassment" [198] to "[an indicated] strong desire to be left alone" [143]. In these cases, debriefing could cause more harm than benefit. Harms to researchers from online harassment disproportionately impact marginalized scholars based on factors including gender, race, chosen research topic, and career expectations of online visibility [113, 186, 253]. The decision to use Bartleby must account for these power dynamics, especially when institutions are still learning how to protect researchers from online harassment. When I decided to study and debrief the Reddit community, I followed guidance from a Data & Society report on "Best Practices for Conducting Risky Research" [185]. For example, team members had discussions about possible risks and followed cybersecurity guides to remove personal information from the internet to reduce risk of doxxing. These examples of situations where debriefing would be inappropriate demonstrate that even well-designed procedures require substantive conversations about when they should be used.

As I found, debriefing systems can also contribute to conversations about substantive ethics even if they seem like mere box-checking exercises. As I saw in the Reddit study, the Bartleby system protected individual autonomy while surfacing substantive issues by enabling participants to voice concerns. I found that debriefing can make researchers aware of contrasting values held by different participants, informing how researchers think and act on their normative values and uses of power.

8.5.4 Passive Non-Participation, Sovereignty, and Non-Alienation

In my deployment of Bartleby, I observed aspects of online, non-consented research that drew my attention to two risks to autonomy: default inclusion and passive non-participation.

In many offline studies in controlled settings, people are not enrolled into studies by default without their consent. If people ignore requests for informed consent, they will not be included in research without their active involvement. However, in online non-consented research, people are *included by default* in research as they participate in online public spheres. People must actively opt out in order not to participate in research.

Non-consented research and debriefing procedures introduces the possibility of *passive non-participation* [56]. People who log into the debriefing system will either decide to opt in as active participants or opt out to become active non-participants. In contrast, passive non-participants do not respond to debriefing and do not actively reason about participation. When people are included in research but are either unaware of their involvement or disinterested in managing a relationship they did not initiate, they do not engage in research ethics on the terms laid out by researchers employing procedures or automated systems.

In my field deployments of Bartleby, I observed that most people were passive non-participants—that is, they did not log into the system to complete debriefing. Distinguishing between opt out and passive non-participation is important because critics of Bartleby may question whether the system truly protects the autonomy of passive non-participants. Opt-out is morally significant because it involves an act of communication, which makes the result of their informed autonomous decision known to me. Passive non-participants do not communicate with me, so I have no way of knowing whether they made an informed decision. I cannot distinguish participants who saw the debriefing invitation and chose to ignore it from participants who never saw the debriefing invitation at all. Access to the ability to opt out via the debriefing invitation is what protects autonomy. So if participants never saw the invitation, one might say that morally it is similar to if the invitation was never made.

How should researchers think about how to protect the autonomy of passive non-participants? Enoch, an ethicist, argues that concerns about autonomy can actually reflect two distinct concerns: a concern for sovereignty, and a concern for non-alienation [90]. *Sovereignty* concerns are about individuals having control over choices that affect them. When theorists and activists apply the standard of affirmative consent [150] to issues such as sexual politics, they are responding to the concern for sovereignty. Once someone has asserted their sovereignty and communicated their decision about a request for consent, that decision is final. *Non-alienation* concerns are instead articulated in terms of a person's deep commitments. Medical caregivers must navigate the concern for non-alienation when, for example, treating someone who is unconscious and cannot affirmatively consent. Imagine the unconscious person has religious commitments that disallow certain medical interventions. Going against those commitments, even if it might save their life, would violate their autonomy. This is because people's deep commitments are intimately tied to their sense of self.

A theory of distinct autonomy concerns helps researchers untangle their worries about passive non-participation in debriefing when using Bartleby. According to Enoch, when researchers talk about autonomy they are sometimes concerned with sovereignty, sometimes with non-alienation, and sometimes with both [90]. If researchers are primarily concerned with sovereignty, debriefing procedures are not enough to address this concern when passive non-participation is possible. Because people cannot be forced to participate in procedures, researchers do not know what sovereign decision they would have made about their own participation.

If researchers are primarily concerned with non-alienation, passive non-participation might be less of a problem depending on the normative content of the research. For example, it's unlikely that counting how many tweets someone made during a time period or advancing public understanding of the effect of DMCA takedowns on free speech is against anyone's deep commitments. It may be against someone's preferences, but likely does not threaten their sense of self. Making feminism discussion communities more welcoming could be against anti-feminists' deep commitments. However, as people who participate in that community, anti-feminist commenters have made a sovereign decision to subject themselves to the governance structure and community expectations of that forum.

Because Bartleby is a system designed to protect participant autonomy in online non-consented research, its design must navigate the issues that arise in this category of research. For situations where researchers are mostly concerned with sovereignty, Bartleby can only protect the autonomy of active participants (whether or not they opt out). For situations where researchers are mostly concerned with non-alienation, the autonomy of passive non-participants is sensitive to substantive issues in the normative content of the research. In most situations, researchers will want to protect autonomy out of concern for both sovereignty and non-alienation. For the reasons I have discussed so far, using Bartleby is an improvement over not using it in such cases.

8.5.5 Debriefing and Spam Filters

Ethics procedures might constitute a form of spam. In my Twitter field deployment, the platform's algorithmic "Quality Filter" potentially affected whether subjects received my debriefing invitations in a way that researchers are currently unable to quantify. The feature is designed to filter out notifications from automated spam accounts. Twitter does not distinguish between debriefing messages and other automated communications for purposes such as marketing or disinformation, a normative stance motivated by values of authenticity in content moderation. From the platform's perspective, these messages are "high-volume unsolicited ... mentions," which constitute platform manipulation [7].

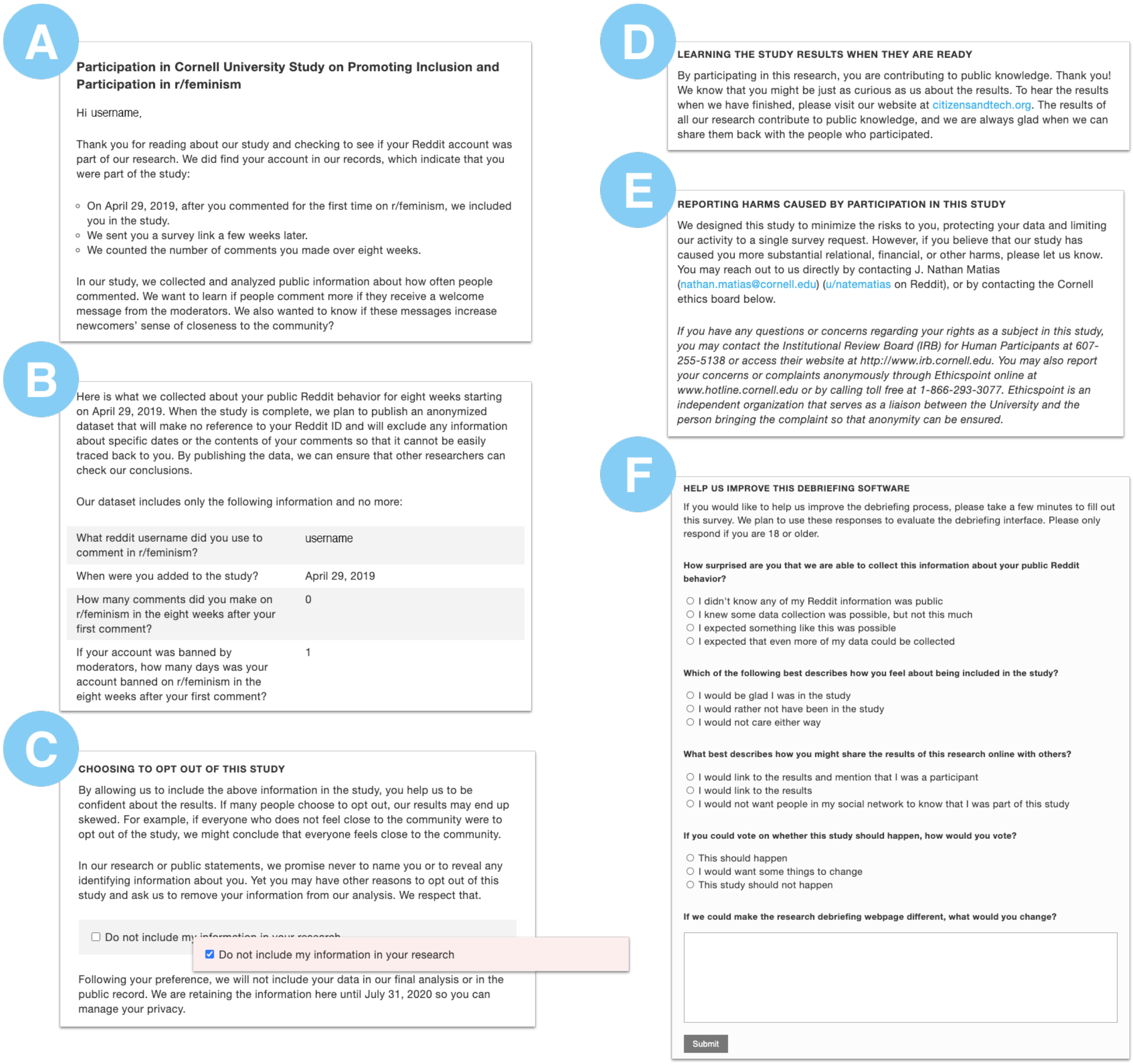
Anti-spam efforts likely complicate the problems of passive non-participation. Spam filters create situations where someone could potentially have exercised their right to opt out, but was prevented from doing so by an algorithm that the neither researchers nor participants control. This problem is widely applicable beyond social media and includes any digitally-mediated communication—such as email-based recruitment, where algorithms can down-weight ethics procedures in the inbox or relegate them to a spam folder.

In contrast to those for whom the platform filters out debriefing messages, other passive non-participants see the debriefing invitation but choose not to respond. For these subjects, debriefing is an unwanted burden on their time and attention. If more researchers adopt a norm of debriefing, the volume of requests could grow substantially. Scholars have noted that the more entities collect and use personal data, the less feasible it is for individuals to manage their privacy separately with each entity [251]. This creates a difficult trade-off for researchers, who want to reach people who will be helped by debriefing (especially those who would be helped but are prevented from engaging due to anti-spam algorithms) but do not want to lose trust by gaining reputations as spammers.

These issues suggest a need for researchers to situate thinking about autonomy in a broader discussion about socio-technical systems. For researchers thinking about debriefing procedures, spam filters highlight the fact that participants' individual choices—usually thought of as a product of individual autonomy—are a product of the interaction between individual autonomy, platform algorithms, and possible value-driven differences between what researchers and participants think of as spam. The burden created by a large volume of individual debriefing requests also highlights the need for more research into collective approaches for managing autonomy.

8.6 Conclusion

With Bartleby, I contribute an open-source research ethics system that provides an interface for delivering debriefing procedures alongside large-scale online research. Researchers who use Bartleby in non-consented research will offer participants more opportunities to exercise autonomy than would otherwise be available. In evaluating this contribution, I underscored the importance of bringing multiple complimentary theories to bear when interrogating the both the promise and limitations of research ethics system design. At a time when few researchers invite any kind of public voice into the research process, I believe that similar creative conversations among design, empirical research methods, and feminist philosophy can advance research ethics and increase public trust in research.



Participation in Cornell University Study on Promoting Inclusion and Participation in r/feminism

Hi username,

Thank you for reading about our study and checking to see if your Reddit account was part of our research. We did find your account in our records, which indicate that you were part of the study:

- On April 29, 2019, after you commented for the first time on r/feminism, we included you in the study.
- We sent you a survey link a few weeks later.
- We counted the number of comments you made over eight weeks.

In our study, we collected and analyzed public information about how often people commented. We want to learn if people comment more if they receive a welcome message from the moderators. We also wanted to know if these messages increase newcomers' sense of closeness to the community?

Here is what we collected about your public Reddit behavior for eight weeks starting on April 29, 2019. When the study is complete, we plan to publish an anonymized dataset that will make no reference to your Reddit ID and will exclude any information about specific dates or the contents of your comments so that it cannot be easily traced back to you. By publishing the data, we can ensure that other researchers can check our conclusions.

Our dataset includes only the following information and no more:

What reddit username did you use to comment in r/feminism?	username
When were you added to the study?	April 29, 2019
How many comments did you make on r/feminism in the eight weeks after your first comment?	0
If your account was banned by moderators, how many days was your account banned on r/feminism in the eight weeks after your first comment?	1

CHOOSING TO OPT OUT OF THIS STUDY

By allowing us to include the above information in the study, you help us to be confident about the results. If many people choose to opt out, our results may end up skewed. For example, if everyone who does not feel close to the community were to opt out of the study, we might conclude that everyone feels close to the community.

In our research or public statements, we promise never to name you or to reveal any identifying information about you. Yet you may have other reasons to opt out of this study and ask us to remove your information from our analysis. We respect that.

☐ Do not include my

☑ Do not include my information in your research

Following your preference, we will not include your data in our final analysis or in the public record. We are retaining the information here until July 31, 2020 so you can manage your privacy.

LEARNING THE STUDY RESULTS WHEN THEY ARE READY

By participating in this research, you are contributing to public knowledge. Thank you! We know that you might be just as curious as us about the results. To hear the results when we have finished, please visit our website at citizensandtech.org. The results of all our research contribute to public knowledge, and we are always glad when we can share them back with the people who participated.

REPORTING HARMS CAUSED BY PARTICIPATION IN THIS STUDY

We designed this study to minimize the risks to you, protecting your data and limiting our activity to a single survey request. However, if you believe that our study has caused you more substantial relational, financial, or other harms, please let us know. You may reach out to us directly by contacting J. Nathan Matias (nathan.matias@cornell.edu) (u/natematias on Reddit), or by contacting the Cornell ethics board below.

If you have any questions or concerns regarding your rights as a subject in this study, you may contact the Institutional Review Board (IRB) for Human Participants at 607-255-5138 or access their website at http://www.irb.cornell.edu. You may also report your concerns or complaints anonymously through Ethicspoint online at www.hotline.cornell.edu or by calling toll free at 1-866-293-3077. Ethicspoint is an independent organization that serves as a liaison between the University and the person bringing the complaint so that anonymity can be ensured.

HELP US IMPROVE THIS DEBRIEFING SOFTWARE

If you would like to help us improve the debriefing process, please take a few minutes to fill out this survey. We plan to use these responses to evaluate the debriefing interface. Please only respond if you are 18 or older.

How surprised are you that we are able to collect this information about your public Reddit behavior?

○ I didn't know any of my Reddit information was public
○ I knew some data collection was possible, but not this much
○ I expected something like this was possible
○ I expected that even more of my data could be collected

Which of the following best describes how you feel about being included in the study?

○ I would be glad I was in the study
○ I would rather not have been in the study
○ I would not care either way

What best describes how you might share the results of this research online with others?

○ I would link to the results and mention that I was a participant
○ I would link to the results
○ I would not want people in my social network to know that I was part of this study

If you could vote on whether this study should happen, how would you vote?

○ This should happen
○ I would want some things to change
○ This study should not happen

If we could make the research debriefing webpage different, what would you change?

Figure 8.3: *The components of the debriefing page. (A) Information about inclusion in the study and the purpose of the research. (B) Table showing what data was collected on the participant, with description of anonymization procedures. (C) Information on risks and benefits, with checkbox for opting out. (D) Details on following up about the study results. (E) Contact information for university review board. (F) Survey on improving the debriefing webpage.*

Chapter 9

Data Refusal as Design

9.1 Introduction

Refusal is a practice of saying "no" to how data is collected or used, and rejecting the processes, goals, or authority of data collectors. A First Nations community might force governments and academics to follow community-defined research policies [97]. A family might try to refuse e-commerce data collection [272]. A citizens group might sue a government agency for engaging in domestic surveillance [88]. In our time, corporations and governments continue to collect data to run systems that profoundly affect every element of people's daily lives. Because these systems reach across many different domains, acts of refusal can take on many different forms. As a result, it can be difficult to see the diverse acts of refusal undertaken by individuals and collectives across society as instances of a broader movement.

To begin to describe these actions, feminist and indigenous scholars have developed *refusal* as a broad concept for understanding the agency of the people whose lives are affected by data regimes [25, 26, 36, 62, 106, 137, 192, 210, 226, 247, 258, 295]. These concepts matter because while activists and community leaders are using the idea of refusal to build shared conversations and to explain their work. Yet, as new ways to use and collect data continue to be invented, so do new ways to refuse. How can people both understand existing practices as acts of refusal, and also think systematically about how refusal responds to the design of data systems?

To address this question, I develop the idea of *data refusal from below* by presenting a design framework for illuminating the spectrum of refusal. The options available to communities are shaped by policymakers, computer scientists, and designers who influence what kinds of data refusal to require, allow, or prevent. Communities navigate these technical and institutional constraints to arrive at creative ways to resist, disrupt, or pursue alternatives. I recognize the process of discovery that enables new refusals as a generative practice of design. To support refusers in this work, I adopt the idea of a "prism of refusal" from Benjamin [36] to characterize refusal strategies across four constituent facets: *autonomy*, *time*, *power*, and *cost*. These facets are considerations that apply to all refusal, whatever strategies are used. I illustrate each of these facets with cases of people who have refused systems of data power. By considering these four facets, I hope marginalized communities can envision and organize refusals that more directly meet community needs and advance a just world.

Starting from the standpoint of people most affected by data collection, I write about and for those who are typically excluded from design decisions about data-driven systems. Scholarship on refusal has largely focused on actions that designers and policymakers can take to create change within powerful institutions. However, when research on data refusal primarily focuses on the standpoint of small groups of influential people, scholars risk sidelining the goals and perspectives of the vulnerable and marginalized. Furthermore, they risk entrenching theories of change that rely on the decisions and goodwill of elites, rather than supporting people's agency to shape their relationships to data regimes. For activists, this framework offers a way to explain their work within and alongside marginalized communities as part of a broader movement of refusal. It also offers a way to think about how their acts of refusal fit into a broader terrain of possible actions. For designers and scholars, this framework offers a way to understand the actions of refusers as a form of participation in technology design. Just as data systems affect people's everyday lives, the agency people exert within systems — including through their non-participation — exert pressure on the design of new technology in response to changing behaviors and collective actions.

While there is an important tradition of research from below in a global context, this paper has a North American focus. As US-based researchers, my point of view on questions of data refusal may differ from the values or needs of a global audience. Many of the cases of refusal I draw on are responses to U.S., Canadian, and European regulatory regimes that have outsized influence on large North American tech firms with global reach. While my work is informed by these particular conversations, I acknowledge the importance of geographic and cultural context in shaping what approaches to refusal are possible or practical.

In this paper, I advance the idea that refusal can be theorized as an act of design, and provide a framework for thinking about current and future acts of refusal. In **??**, I establish an initial scope for the concept of *data refusal from below*. Drawing on conversations in feminist standpoint theory, I argue that acts of refusal undertaken by the marginalized must be understood differently from refusal by the institutionally privileged. In section 9.2, I argue that refusal can be thought of as an act of designing alternate social configurations. Seeing refusal as design — and by extension, refusers as designers — creates opportunities for design-oriented theory and methods to contribute to the continually-evolving practice of refusal. In section 9.3, I ground my conceptual framework contribution in the methods and approaches of computer science and design. Frameworks serve *descriptive*, *evaluative*, and *generative* purposes by giving scholars language to describe and compare artifacts in the world, and imagine new design possibilities. The paper's main contribution appears in section 9.4, where I present a framework consisting of four facets of refusal: *autonomy*, *time*, *power*, and *cost*. When introducing each facet, I use real-world cases to highlight important considerations that apply generally when using the framework to analyze instances of refusal. Finally, in section 9.5, I explicate the descriptive, evaluative, and generative uses of the framework, reflect on how designers can learn from acts of refusal, and articulate a vision for a politics of refusal in a world shaped by large institutions.

9.2 Refusal as Design

To understand refusal as design, one must understand the nature of design as a form of socio-technical power that people resist and also take up in the act of resistance. Herbert Simon calls design an "act of envisioning possibilities and elaborating them" [129]. Just as Simon writes of planners who "took whole societies and their environments as systems to be refashioned," [129] the systems I am concerned with in this paper use the products of mass data collection to reshape politics, economics, and culture. Information scholars have long argued that data systems should be understood not as purely technical, but socio-technical — best understood "as systems and devices embedded in larger material and social networks and webs of meaning" [202]. Design of this kind is not just carried out by individual experts or engineers, but by an assemblage of actors across the tech industry, governments, and civil society.

Because refusal involves deliberate acts to influence socio-technical systems of power, I see refusal itself as a form of design. For example, when activists refuse the use of facial recognition systems for mass surveillance, they are imagining a different arrangement of power that includes the typical domains of design: services, policies, and user experiences. For Benjamin, refusal is "seeded with a vision of what can and should be" [36]. Similarly, Simon points out that "everyone designs who devises courses of action aimed at changing existing situations into preferred ones" [129]. When refusers think strategically about the actions they need to take to bring about better technological futures, they are engaging in a design process that is pragmatically oriented and considers alternatives. Rather than designing physical artifacts, they are designing refusals as social artifacts. Instead of physical objects, these artifacts are individual and collective actions that challenge existing social relations.

Like all design, refusal also is concerned with relationships and obligations between designers and users. As Flores et al. argue, "technology is not the design of physical things. It is the design of practices and possibilities to be realized through artifacts" [100]. When people are affected by systems that they did not ask for, design imposes relations and obligations that are unwanted. Anthropologists have theorized this creation of obligation through the lens of gift-giving. For Marcel Mauss, gifts are never freely given; they carry with them a hierarchy of power, expectations, and responsibilities between giver and recipient [182]. Sociologists Fourcade and Kluttz have pointed out that enrollment into digital systems often involve a "give-to-get" relationship that "masks the structural asymmetry between giver and gifted" [101]. Saying no to a gift, then, involves refusing the social order that the gift implies and upholds. Analogously, when people refuse technical systems and data collection, they are reshaping an implied relationship—the commodification and appropriation of their data under the pretext of a fully reciprocal relationship.

Because people refuse systems with an alternative in mind, refusal is an important form of participation in the design process, which expert designers never fully control. McGranahan argues that "to refuse can be generative and strategic, a deliberate move toward one thing, belief, practice, or community and away from another" [192]. Ruha Benjamin notes that the "prism of informed refusal sets out to explore the capacity for resisting and reimagining ... without at once romanticizing or valorizing resistance as an inherent 'good'" [36]. For

Benjamin, resistance and refusal are not end goals, but pragmatic and transient forms of agency along the way to a longer term set of societal outcomes. Seeta Peña Gangadharan argues that "when marginalized people refuse technologies, they imagine new ways of being and relating to one another in a technologically mediated society" [106].

By framing refusal as a form of design, my work explores an opportunity to apply methods and theories from design to support refusal in practice. I offer a framework for describing, evaluating, and generating acts of refusal using the four facets of *autonomy*, *time*, *power*, and *cost*. My framework encourages designers and activists to consider each of these facets as component design characteristics of refusal.

9.3 Methods: Design Dimensions in Computer Science

In this paper, I offer a conceptual framework for refusal of data systems that I hope can be used by people planning acts of refusal. Framework papers illuminate a series of cases with relevant scholarship to reveal how that scholarship can inform further research and design. Computer scientist Michel Beaudouin-Lafon offers three central uses for conceptual frameworks: *descriptive*, *evaluative*, and *generative* power [33]. Descriptive power refers to how well a framework provides language to describe existing cases. A usefully-descriptive framework would provide a consistent set of concepts for analyzing multiple cases of refusal, even if they seem very different. Evaluative power refers to how a framework helps make comparisons between features of different cases. A framework with clear evaluative power would help designers use core concepts to assess the relative strengths and weaknesses of two cases of refusal. Finally, generative power refers to the usefulness of a theory to identify unmet needs and inform the creation of new cases. A framework with strong generativity would help people develop new approaches to refusal that have not yet been realized in practice.

Computer science and design researchers, particularly in HCI, use conceptual frameworks to describe, evaluate, and generate designs. Their goal is to put theory into practice by using the framework to inform future designs. In particular, Shaowen Bardzell has articulated the importance of generative applications of feminist theory to feminist HCI and interaction design [28]. Frameworks serve a similar purpose for scholars and activists in the social sciences. For instance, Sasha Costanza-Chock has drawn from social movement theories to develop a framework for imagining online activism tactics [229]. For Ruha Benjamin, refusal is a "prism" that situates actions "within a more comprehensive spectrum of human agency vis-à-vis technoscience" [36]. Benjamin's idea of a prism of refusal is an apt analogy for the purpose of the framework in this paper. By separating instances of refusal into the components of a spectrum, researchers can study refusal more clearly while also supporting people to imagine new kinds of resistance in "a much larger terrain of action and negotiation" [36].

In this paper, I offer a framework of refusal as design with four facets: *autonomy*, *time*, *power*, and *cost*. In contrast to work taxonomizing forms of resistance to data collection [48, 273], I am not listing discrete categories that differentiate instances of refusal from one another. Instead, my framework separates out four common aspects that constitute any act of data refusal. Each example of refusal includes all four facets. These facets can be used analyze cases independently, and also to draw attention to ideas that emerge at the intersection of multiple facets.

I developed this framework by curating real-world cases of refusal from academic and journalistic sources, and identifying characteristics of each approach that could generalize to other cases. I also referenced existing critical scholarship on refusal, especially the Feminist Data Manifest-No [62] and Benjamin's work on informed refusal [36], with particular focus on discussions of what refusal can accomplish and who is able to refuse. As I fleshed out parts of the framework, I iteratively sought out additional cases to illustrate contrasting positions along relevant dimensions. For example, after reviewing a case involving efforts to uphold individual autonomy at the expense of other values, I made sure to consider cases with contrasting approaches to autonomy. I settled on the four facets when I felt that they were able to usefully represent the considerations and critiques associated with the cases in academic and journalistic discussions. The cases I include in the paper are not meant to be an exhaustive representation of refusal practices. This is because technology, and people's responses to technology, continue to evolve over time. However, I did make an effort to include examples that exemplify multiple representative ways of thinking about autonomy, time, power, and cost.

9.4 Four Facets of Refusal

This paper describes four facets of refusal to consider when evaluating, describing, and imagining data refusal. When people develop new forms of data refusal, they will need to develop answers to these four questions.

9.4.1 Autonomy: Individual and Collective

The concept of autonomy is important to conversations about refusal, since the desire and capacity to refuse are linked with the idea of self-determination. Autonomy is typically defined as the capacity to freely make informed choices. The authors of the Feminist Data Manifest-no, for example, express their commitment to "research cultures that promote data autonomy and SELF-representation" [62]. In the U.S., individual consent rose to prominence in the 20th century as the dominant method for establishing practical protections to individual autonomy in research and medical contexts [222]. Regulations such as the European Union's General Data Protection Regulation (GDPR), Brazil's Lei Geral de Proteção de Dados (LGPD), and the California Consumer Privacy Act (CCPA) have drawn from this framework to establish individual consent as a central virtue of data protection [183].

As the current wave of privacy regulations demonstrate, prioritizing *individual autonomy* leads to forms of data management that involve offering people choices that they can accept or refuse. In the U.S., the close relationship between autonomy and individual consent was established in the mid-20th century when U.S. medical researchers and bio-ethicists turned to the idea of consent in their search for ways to prevent the atrocities of Nazi scientists who exploited victims of the Holocaust [208]. On the frontiers of science, risks and harm are difficult to estimate and scientists tend to justify their work as a common good. As Rothman summarizes, "human experimentation pitted the interests of society against the interests of the individual" [222, Chapter 5]. Informed consent provided a versatile procedure for individual interests to be protected from the consequences of utilitarian arguments from scientists.

Because harms can often be collective, any act of refusal must reckon with both individual and collective autonomy. In the theory of privacy as contextual integrity, the philosopher Helen Nissenbaum argues that privacy depends on upholding context-specific norms beyond the individual. These norms are influenced by individual data subjects as well as data collectors, infrastructures of information flow, and the relationships between all involved entities [201]. In parallel, feminist philosophers have noted that accounts of individual autonomy have often struggled to reconcile the importance of self-determination with the fact that individuals are "socially embedded and ... shaped by a complex of intersecting social determinants" [181]. To reconcile this tension, feminist scholars have advanced the concept of "relational autonomy," a concept that acknowledges how individuals are interconnected, their actions are socially embedded, and their choices are influenced by power relations [181]. Because autonomy can be an individual or collective action in context, any endeavor of refusal involves questions about whose autonomy is at work.

Individual autonomy protects individuals' interests against those of data collectors.

Though many people consider large-scale corporate, government, and academic data collection to be common knowledge, others are surprised and upset to learn that their data is collected and used without their personal knowledge or consent [96]. For example, when the public learned that some Facebook users had their feeds altered as part of Facebook' Emotion Contagion study in 2014, many people expressed that they felt exploited or manipulated [120]. Scholars pointed out the lack of consent—either prior or retroactive—as one of the central ethical issues with the study [118]. A procedure for consent or retroactive opt-out, scholars argued, would have provided Facebook users with an opportunity to exercise their autonomy and potentially refuse participation in the study.

Individual consent represents a fruitful area for the design of refusal. Researchers and policymakers have worked to establish individual consent procedures as a norm and legal requirement for online data collection. Feminist HCI researchers have suggested that for data collectors to fully support individual agency and autonomy, their operationalization of affirmative consent must be voluntary, informed, revertible, specific, and unburdensome [150]. System designers have also used design as a form of enquiry in the articulation and evaluation of ethical frameworks including procedural and substantive theories of autonomy and consent [296, 297]. For example, in response to regulatory requirements created by the GDPR, companies have introduced consent management platforms to provide users with information about web trackers and solicit their individual preferences about tracking [204]. In another example, the Consentful Tech Project develops guidelines and resources around digital consent, and has produced a curriculum guiding tech designers to "build consentful cultures and technologies" [269]. Designers have used these guidelines to prototype a user flow for signing a collective letter that also allows signatories to revoke their signature just as easily [93].

Individual autonomy cannot account for the conflicting, equally legitimate interests of different people.

When a single decision about data collection affects multiple individuals, individual consent cannot account for everyone's autonomy. For example, photographs often depict more than one person. The person who takes the photo might not even appear in the image, and the person uploading a photo to an online service might be another person altogether. An individual consent process will only ask one of these people for their preferences, and that person might not have the right to make decisions for the others. Barocas and Levy use the term *privacy dependencies* to refer to the ways "our privacy depends on the decisions and disclosures of other people" [30]. Amy Hasinoff notes that these problems are especially pronounced when thinking about consent and the circulation of intimate imagery online [124]. Refusal strategies that prioritize individual autonomy—such as opting out of consent to data collection—are meant to protect a single person's welfare against the interests of data collectors. But data collection commonly complicates where the "boundaries of a person" can be drawn [83]. When multiple people have legitimate interests in the circulation of data, no individual-choice consent decision can even record everyone's interests—especially if they don't all agree.

Even if a single instance of data collection respects the autonomy of everyone the data is collected from, procedures like individual consent cannot account for the collective outcomes of that data's use. Many uses of data also affect people across society, not only those directly involved in data collection. While a single photo on its own has limited uses, large photographic datasets can be used by facial recognition algorithms that are incorporated into abusive policing and immigration systems. Because consent is only meant to help people manage risks to themselves, it cannot prevent risks to people who are not yet in a dataset or to society at large.

For many privacy scholars, the need to protect individual autonomy presents a seemingly intractable "consent dilemma" [251]. When problems arise from individual choice, legal scholars tend to think that the only alternative is the power of the state. Solove notes that a dilemma arises because "the most apparent solution — paternalistic measures — even more directly denies people the freedom to make consensual choices about their data" [251]. Relational theories of collective autonomy offer a further option that prioritizes autonomy while accounting for risks and benefits beyond the individual.

Collective autonomy balances the interests of both individuals and society.

When people come together to collectively manage how their data is collected and used, they are conceiving of their autonomy as relational by entwining their own interests with that of a larger group. As Salomé Viljoen argues, "data isn't collected solely because of what it reveals about us as individuals. Rather, data is valuable primarily because of how it can be aggregated and processed to reveal things (and inform actions) about groups of people" [225].

Data cooperatives are an example of organizations that allow their members to collectively manage the terms of data collection and use. Cooperatives—such as those formed by consumers, tenants, or employees—comprise a "group that perceives itself as having collective interests, which it would be better to pursue jointly than individually" [9]. The Salus Coop is an example of a data cooperative formed around health research data. Members donate their own health data to the co-op, and contribute to medical research while retaining control over their own data. Through the use of a common good data license, the co-op specifies terms that researchers must follow in order to use members' anonymized data [155]. The Native BioData Consortium [264–266] is another example organization that is led by indigenous geneticists who work to keep the storage and use of biological samples local to tribal communities. Cooperatives protect the rights and autonomy of individual members by setting clear terms for how their contributed data can be used, while also permitting health research that might provide collective benefits.

To imagine new legal mechanisms for implementing collective notions of consent and autonomy, privacy scholars have argued that data processing entities should have a fiduciary obligation to act in the interests of the people whose data they manage. For example, Balkin suggests that the same legal duties of care and loyalty that apply to doctors and lawyers ought to apply to companies and providers that work with data [24]. Delacroix and Lawrence instead argue that a multitude of third-party "data trusts" could negotiate with data collectors according to terms collectively determined by the trust's beneficiaries, "introducing an independent intermediary between data subjects and data collectors" [76]. Similarly, RadicalXChange's proposal for Data Coalitions would "establish tightly regulated collective bargaining entities ... [to] pursue their Members' varying interests from a vastly better bargaining position" than if they were to bargain individually [152]. These proposals rely on models of autonomy that allow individuals to delegate the capacity to consent to a representative third party, which social scientists have called "representative consent" [144]. Critics of data trusts and co-ops note that these proposals attempt to fit data into existing legal frameworks for governing property or labor relations, which subjects people and their data to pre-existing extractive and coercive market conditions [224]. Nonetheless, this line of research suggests possible ways collective data management can balance individual autonomy with the common good.

9.4.2 Time: Reactive or Proactive

Designers of data refusal activities must also consider the timescale in which refusal operates. As Solove writes, "privacy is an issue of long-term information management" [251]. Many cases of refusal from informed consent to class action lawsuits occur after an attempt at data collection. These reactive approaches are possible when mechanisms of data collection procedures and their possible harms are known. Even after harms occur, future refusal can sometimes redress those harms for the people affected. Yet as new developments in data collection, use, and disclosure continue to be invented over time, refusal can be mobilized to proactively prevent future harm.

Reactive refusal responds to harms that have already occurred.

Because the power of data is often wielded without even informing those involved, the people harmed by data collection often find themselves reacting to harms after they occur. In January 2019 when IBM published roughly a million photos of unsuspecting individuals to create a facial recognition dataset they called Diversity in Faces, many people were surprised and upset to discover their photos were included [205]. From 2004-2014, some people who published photographs to Flickr (a photo-sharing site) provided them under a Creative Commons license that allowed the reuse of those photos by third parties. Whether or not the people in those photographs consented, their likeness was passed from company to company over the years. Yahoo purchased Flickr, published a public photography dataset in 2016, [261], and IBM later re-purposed these images to train facial recognition algorithms. These photos were quickly adopted by other large tech companies hoping to improve their products [216]. It is likely that immigration offices, law enforcement, and other institutions are using algorithms trained on the faces of people who never knew their photos had been shared and never had a chance to refuse.

In response to public outcry over the Diversity in Faces dataset, IBM assured the public that they "will work with anyone who requests a URL to be removed from the dataset" [205]. Although IBM will not inform people if their image appears in the dataset, the company requires people to send a precise copy of the image they want the company to remove. If that person is unable to submit a copy of every image of them that IBM holds, the company will retain their biometric data. Because IBM defines the process and controls information about the process, they can create a reasonable-sounding procedure that is nearly impossible to opt out of.

Because IBM's process puts the burden on the public to check whether their images were included, people in the state of Illinois instead filed a class action lawsuit against the company [8]. This lawsuit represents a collective refusal of IBM's actions by people in Illinois. By asking for $5,000 in damages for each person whose rights were ignored, the lawsuit is designed to deter IBM and other firms from collecting, storing, and disseminating data that puts people at risk without their consent.

IBM's data removal process and the Illinois class action lawsuit are both refusal strategies that can only happen after a harm has been discovered by the public. These strategies help those affected by a current instance of data collection. But if a similar scandal happens in the future, the people affected would have to seek recourse all over again. Though scholars have put forth arguments for and against class actions as a way to deter future wrongdoing [99], lawsuits will still only be able to react to new data collection scandals in the future if they do occur.

The effects of data-driven harm occur on extended timescales.

Over a long period of time, the same dataset can change hands and be copied and reused beyond its original purpose. When Flickr launched in 2004, it was a small startup based in Vancouver. In 2005, Yahoo acquired Flickr and moved all of its data from servers in Canada to servers in the United States, which changed the set of laws that applied to the data [1]. In 2007, Yahoo shut down the Yahoo Photos service and provided people with a way to transfer their photos to Flickr. People who transferred their photos to Flickr would have exposed their photos to inclusion in the facial recognition dataset despite originally using a different service. Ten years after Flickr launched and likely after many people had forgotten about their early uploads, Yahoo aggregated 100 million photos into a public research dataset [261]. Yahoo researchers published an academic report on the dataset that has been cited in hundreds of computer vision research papers. Then in 2017, Yahoo was acquired by Oath (renamed in 2019 to Verizon Media). In 2018 Oath sold Flickr to SmugMug, another photo sharing service. In the meantime, IBM researchers downloaded a copy of the dataset from Yahoo and modified it to create the Diversity in Faces dataset released in 2019.

Because datasets can be copied, remixed, and recirculated endlessly, addressing harm reactively requires ongoing vigilance. Throughout this convoluted chain, people's non-commercial copyright agreements with Flickr became less influential as their data proliferated through multiple hands. Even if everyone who had a copy of a photo could be contacted, it would be difficult to force every derivative dataset to comply with the original copyright agreement, let alone consent that was never sought.

A one-time decision based on information available at the time a dataset is created cannot address all subsequent developments. Originating in biomedical contexts, informed consent is a one-time decision at the point of data collection that forms an agreement about how the data will be used. But as Sedenberg and Hoffmann [236] point out, within the biomedical model it is "unclear how to assess harm or potential risk outside of physical interventions or in-person interactions, as with the possibility of reidentified data or harms that may occur on extended timescales." As the IBM case illustrates, data also passes through different hands after collection, where it can be copied, stored, and combined with other data in ways that cannot be anticipated at the point of consent. In an analysis of machine learning datasets that had been retracted for ethical concerns, Peng et al. found that the data not only continued to have wide availability and use in research, but also that researchers lack the infrastructure to track derivative datasets to understand their impact [209]. Because new developments in data collection, use, and disclosure continue to be invented over time, all possible harms cannot be anticipated by a one-time decision.

Proactive refusal tries to prevent future harm.

When it's apparent that technologies will continue to produce further unacceptable harm, activists have turned to refusal strategies that prevent new cases of harm before they occur. In the U.S., the movement against government face surveillance is an important case for understanding the importance of proactive refusals. In January 2020, Robert Williams was wrongfully arrested on the basis of an incorrect match by a police department's facial recognition system [133]. Despite the fact that the company that provides this technology to the police "does not formally measure the systems' accuracy or bias," the police used the flawed match as a primary factor in the decision to arrest Williams. Because of the arrest, Williams was forced to miss work, had to pay a bond to be released, and faced potential shame and embarrassment in his social circle. In response, the American Civil Liberties Union (ACLU) has filed suit against the Detroit Police Department on Williams' behalf. While a reactive lawsuit may help Williams, it cannot reverse the harm that was done to him. If police continue their use of surveillance technologies, cases like Williams' will only happen more and more often. How can advocates proactively prevent new cases where police abuse facial recognition from occurring in the future?

Activists and advocacy groups have used strategic litigation and legislation to ban the government use of facial recognition and force corporations to stop building and selling facial recognition systems [160]. In 2019, San Francisco became the first U.S. city to pass a bill preventing local government agencies from procuring or using facial recognition, or using information from external facial recognition systems [68]. Since then, a handful of other U.S. cities have passed similar legislation, and some state and federal regulations are under consideration [50]. By reducing police use of facial recognition, activists are reducing the capacity of the state to make future misidentifications that lead to arrests.

Proactive refusal involves concerted efforts to prevent irreversible harms by addressing the conditions that bring them about. In Joy Buolamwini's testimony at a congressional hearing on facial recognition technology, she notes that even when the police correct mistakes resulting from facial recognition mis-identifications, the "damage [is] done" irreversibly [49]. In 2020, IBM announced that it would stop developing or researching facial recognition technology [211]. In the future, IBM and other tech companies could be deterred from collecting facial biometric data—and repeating the same data scandal—by regulation and ongoing public scrutiny.

9.4.3 Power: Foreclosing Possibility or Creating Possibility

People engage in refusal because there is something they want to prevent or change. That's why any act of refusal must grapple with questions of power, which feminist scholars have defined as "the capacity to produce a change" [156]. It is "the power of ability, of choice and engagement. It is creative; and hence it is an affecting and transforming power but not a controlling power" [228]. In my work, I focus on power from the standpoint of refusers rather than data collectors. For corporations and governments who seek data to increase their control over the public, power is a form of domination. But for those engaged in refusal, power is expressed as an ability to resist and reconfigure their relationships with data collectors. Refusal is powerful in a given situation when it attempts to change the behavior of data collectors, causing them to engage with refusers in a new way or preventing them from engaging at all.

Some offers to refuse are designed to placate instead of create change.

Because refusal involves potentially coming into conflict with powerful institutions, it can appear to be a good thing when data collectors offer well-defined processes for people to opt out or voice dissent. However, data collectors often use these pathways of dissent to limit refusers' agency by directing them away from other unspoken possibilities. Ruha Benjamin conceives of refusal as a privileged form of political action because "the capacity to refuse rests upon a prior condition of possibility—that one has been offered something in the first place" [36]. When data collectors give people limited power to choose, it's arguably better than no choice at all. But if data collectors frame choices such that all possible options benefit them at the chooser's expense, then people must find ways to refuse on different terms.

The choice of whether to opt out of facial recognition training data is an example of a choice that ultimately limits agency. As facial recognition becomes increasingly integrated into policing practices that disproportionately target communities of color, people of color might reasonably choose to opt out of training the systems that will be used to surveil them. Under the California Consumer Privacy Act (CCPA), companies like IBM, Amazon, and Clearview AI who produce facial recognition systems are required to offer individual consumers the ability to remove their own data from these training sets. Hypothetically, if enough people do this, then training sets could become systematically biased against those who are more likely to opt out. The option to opt-out sets up a lose-lose situation. Opt out, and you will potentially exacerbate already-pervasive racial bias in facial recognition. Opt in, and your image will be used to fine-tune a facial recognition system that a company is likely trying to sell to immigration enforcement agencies. Engaging with the choice in either direction might distract refusers from seeking other

Rejecting the broader injustice of surveillance and overpolicing is not within the terms of refusal offered by these companies. The question of opting out treats as given the idea that companies will continue to build these systems. Even if any particular individual opts out, companies and governments will simply acquire more vulnerable people's faces to build their systems—as Google did by targeting homeless people for facial recognition data [139], or as NIST did by using images of immigrants, children, and dead people [165]. Fundamentally, these opt-out systems do not provide the option of challenging or constraining companies' power to pursue data collection for facial recognition.

Some kinds of refusal create pathways to systemic change.

Some acts of refusal reconfigure systems of power entirely, beyond individual relationships or circumstances. For example, a long history of unethical research conducted on indigenous communities in Canada has included researchers who performed non-consented nutrition research on indigenous children that led to deaths [180], used blood samples for secondary research without consent, and supplied information on dissident indigenous groups to repressive regimes [98]. Research and data collection on indigenous communities has happened through "imperial eyes"[250], excluding these communities from participation in data collection processes, governance, and ethics oversight. As a result, data collection on indigenous people often fails to serve their interests or help communities answer questions about their own well-being. Scholars and practitioners of *indigenous data sovereignty* have explored ways for indigenous communities to retain access and control over their own data as a way to reduce colonial dependency on settler states by collecting data that better reflect the values of indigenous groups—including by gathering information about disparities that settler states refuse to provide [250]. In this context, asserting sovereignty refuses extractive forms of data collection in favor of alternative methods that prioritize indigenous peoples' right to self-determination.

To provide guidance on how First Nations data should be used, the First Nations Information Governance Centre (FNIGC) developed a set of principles in 1998 called OCAP (ownership, control, access, and possession) [97]. The OCAP principles state that communities collectively own information and cultural knowledge pertaining to them; have the right to control the collection, use, and disclosure of that information; must maintain access to that information no matter where it's circulated; and must be able to assert legal jurisdiction over their data [98]. To demonstrate the use of OCAP in practice, the FNIGC operationalized the principles to guide the design of the First Nations Regional Health Survey (FNRHS), "the only First Nations-governed, national health survey in Canada that collects information about First Nation on-reserve and northern communities" [97]. As the first national survey to be fully designed by First Nations representatives, it now serves as a primary information source informing healthy policy decisions relevant to First Nations communities. The development of OCAP represents one of many parallel efforts led by indigenous communities to assert self-determination in community-driven data practices [127].

Without action from groups involved with the FNIGC, the Canadian government would not have considered First Nations collective data autonomy to be possible. In fact, some national legislation in Canada continues to conflict with OCAP principles [97]. The FNIGC have changed how research on and with First Nations is conducted in Canada by pragmatically creating opportunities to exercise data sovereignty through partnerships and policy. In this context, exercising power is not merely engaging with what has been offered, but "attempting to negotiate the terms of one's engagement" [36]. Exercising refusal involves more than saying no to existing options on the table. Refusal also involves creating new options and changing the landscape of engagement going forward.

9.4.4 Cost: Accepting Cost or Reducing / Redistributing Cost

Acts of refusal can involve significant costs to the people and groups who engage in it. I use the term *cost* to refer to things people must sacrifice in the course of refusal—such as effort, time, money, safety, or social capital. Researchers have found that introducing even small costs can make people less likely to opt out of data collection—even if they previously expressed a desire for privacy [20]. I also consider more abstract costs that create barriers to refusal, such as when hidden information or specialized knowledge is required to make an informed refusal.

Scholars who study social movements have studied cost in efforts to explain why one person might participate in social movements while another may not. Sociologist Doug McAdam draws a distinction between "low- and high-risk/cost activism," arguing that engagement in high-cost activism are affected by both individual and structural factors [190]. For instance, some people may be individually more risk-tolerant, or have more resources that enable them to take risks. Costs are also structural, since not everyone has the same support from peers and institutions to take risks. Due to these intersecting considerations, higher-cost forms of refusal are less available in ways that reproduce individual and societal inequality. People also experience costs differently depending on their personal situation and social position.

While cost can deter people from refusal, refusal also has the potential to change how cost is distributed amongst groups of people. For example, some forms of collective refusal exhibit network effects, becoming less individually costly as more people join in. Refusal that relies on the development of infrastructure can have a high one-time cost that is amortized over long-term use by many people. Refusal can even uncover hidden information that reduces barriers to future refusal by others. When addressed collectively, the costs of refusal can be altered alongside the relationship between data collectors and refusers.

Refusal has costs for individuals and communities.

The costs of refusal are also borne by the people connected to those who refuse. In 2014, sociologist Janet Vertesi decided to prevent companies from learning about her pregnancy and her children after birth [272]. Vertesi's family collectively decided that the only way to hide this knowledge from companies was to prevent data about her pregnancy from being collected at all. Since many parts of everyday life rely on commercial data collection, Vertesi and her family decided to build and maintain alternative infrastructures in a practice she calls "digital homesteading". Because companies use purchase data for targeted advertising, Vertesi started using cash for all transactions. To stay connected with others, she built a phone using open source hardware to avoid data collection from commercial apps and operating systems. Her family shares data on their own servers over a home network. Vertesi also had to convince friends and family not to post or mention her children online. Her family chose to be isolated from community networks in some cases and decided to "unfriend" a family member who didn't realize that so-called private messages are visible to corporate platforms [272].

Vertesi's attempt to restructure digital power in her household came with individual and social costs. Vertesi needed the technical expertise, time, and money to set up her own devices and networks. Refusal disconnected her from important friendships and led her to lose contact with some family members. Though she decided to bear the costs of refusal, Vertesi's case illustrates Benjamin's observation that "refusing the terms set by those who exercise authority in a given context is only the first (and at times privileged) gesture in a longer chain of agency that not everyone can access" [36].

Costly refusal can entrench inequality.

The authors of the Feminist Data Manifest-No write that "not everyone can safely refuse or opt out without consequence or further harm" [62]. When refusal is costly, cost is experienced unequally in ways that reflect structural inequality in society. For example, the U.S. Customs and Border Protection (CBP) agency has deployed biometric face scanning technology at airports across the country. The agency scans international travelers' faces on entry and exit, and compares the scans to images stored in visa, passport, and immigration databases [104]. This program poses risks to privacy and civil liberties, and CBP has already experienced a data breach that leaked the information of thousands of people [85]. But despite the risks that CBP's face scanning program poses to travelers, refusing the program could be even more individually costly. U.S. citizens can opt out of face scanning, though airlines typically fail to disclose that this option exists [104]. The choice is unavailable to non-citizens. When citizens do opt out, they are rerouted into private search and screening procedures that involve additional time and scrutiny. As Ruha Benjamin points out, "it is coercive to say one has a choice, when one of those choices is automatically penalized" [36]—yet, these are the terms of refusal offered to international travelers. These procedures are known for discriminatory application, particularly against people who appear to be Muslim or from the Middle East. Benjamin further notes that "rebuffing the authority of the state as exercised through technoscience causes individuals to experience the underside (or outside) of biological citizenship . . . in which refusal is always, already guilty" [36]. When the choice to refuse means opposing the power of the state, whether or not a refusal is successful is often a secondary consideration to the high cost of refusing at all.

Some refusal strategies reduce or redistribute the costs of future refusal.

Some refusal strategies allow people to help others refuse more easily in the future. This can occur when collective action benefits from network effects. Consider Tor, a volunteer-run system for using the internet anonymously. A person accessing an online service through Tor will have their internet traffic encrypted, interleaved with other people's traffic, and routed through a series of distributed relays. Each relay is only able to decrypt enough information to know where to pass the data next, making it difficult for web services to trace the origin of a request. People who volunteer to run relays help forward each other's requests to conceal the origins of their data. Tor is an example of refusal by obfuscation [48]. Obfuscators accept that data collection will occur given the difficulty of opting out, and instead refuse to allow data to be identifiable to data collectors. Tor only works when many people run relays, and its effectiveness at obfuscation improves as more people use it. Therefore, it distributes risk across many people and reduces the individual risk of deanonymization for each additional participant. Further, its benefits can be shared by people who don't need to know who each other are or imagine themselves as a collective.

Another example of refusal creating new opportunities for others is the Stop LAPD Spying Coalition's public records request and subsequent lawsuit against Los Angeles Police Department [88]. In response to the LAPD's expanding use of surveillance technology, the Coalition conducted surveys and focus groups with community members, finding that clear majorities of participants felt that the police should not be using predictive policing [256]. The coalition requested information on data policing programs that involved creating lists of neighborhoods and individuals that would be the focus of increased policing. The programs had been canceled after an internal audit concluded that the lists were racially biased and lacked oversight. Yet the details of the lists had not been released to the public. By using legal and public pressure to acquire records about data policing, Stop LAPD Spying Coalition was able to bear the costs of refusal on behalf of the individuals targeted on the list. With records of how the lists were used, the Coalition can help individuals navigate the negative effects of their inclusion—which could involve further acts of refusal.

9.5 Discussion

9.5.1 Uses of the Framework of Data Refusal from Below

The framework of data refusal from below can be used for descriptive, evaluative, and generative purposes.

Descriptive power. The framework can be used to describe existing forms of refusal with shared language, no matter how different they may appear. For example, in the case of the IBM facial recognition dataset described in section 9.4, some users may have opted out by asking IBM to remove their images from the dataset. This refusal is individual because it relies on people to make decisions about their own autonomy. It is reactive, because it responds to IBM's acquisition and use of data after the fact. The act of opting out alone has little power to change IBM's business model. And finally, people experience few costs by opting out. Going to a website and filling out a form to opt out at first glance has little in common with the actions of digital homesteaders or the movement around Indigenous Data

Sovereignty. However, as we've shown, these cases can also be described along the same dimensions. Having better descriptions of refusal cases is important because the language people use to describe cases focuses their attention toward the mechanisms that cause refusals to be successful or unsuccessful.

Evaluative power. The framework can also facilitate comparison between different cases of refusal along the four dimensions. Consider the movement to ban facial recognition in the U.S., which can be described as a collective, proactive, high power, high cost form of refusal. When compared to the approach of individual opt-out in the IBM case, facial recognition bans are advantageous along the facets of autonomy, time, and power. However, pursuing these has higher costs in terms of time and collective effort from activists, advocacy groups, and lawmakers. Because the framework enables these comparisons along four non-overlapping dimensions, people are able to have clearer conversations about values — what kinds of refusals are more or less normatively desirable.

Generative power. Feminist scholars have long noted the importance of refusal as a generative practice [26, 107, 192, 247, 267]. In design frameworks, a highly generative framework is also *operationalizable*, which is to say that it speaks both to *what* — what are the desirable qualities of something that doesn't yet exist — and also the *how* — what are steps a designer could take to create that thing. An example of something with low generative power is a typology based on observed cases, because it's hard to infer things yet to be imagined from a list of existing types. An example of something with high generative power is a formal model of a system which can describe a combinatorial space of possibilities in terms of discretely adjustable dimensions. My framework certainly has generative capabilities, but falls short of the ideal — because I am studying a highly situated social phenomenon, it is unlikely that I would be able to produce a simple formal model of refusal the same way I would of a programming language.

For design researchers, the main generative use of my framework is to identify gaps in the refusal design space. The cases I examined all involved trade-offs between the four facets of autonomy, time, power, and cost. Thinking systematically using these facets, researchers can ask questions about whether there exist refusal strategies that approach the ideal: collective, proactive, powerful, and low-cost. Even if the framework's generative power is limited in the sense that it doesn't tell researchers how to design such a refusal strategy, the facets help them imagine what could be possible. If such ideal strategies don't exist, the facets may inform conversations about what practical factors lead to these considerations being in tension.

More broadly, the framework can help refusers think about what kinds of approaches are appropriate for their goals given their pragmatic constraints. People may often feel trapped at a fixed position along each of the refusal facets. For example, people included in facial recognition datasets often don't know each other, and don't have existing infrastructure in place to collectively organize — constraining them to primarily rely on individual autonomy. In other cases, people are not in financial or legal positions to bear high costs of refusal, constraining them to consider what can be accomplished at low cost. Staking out certain positions along some of the facets while considering open-ended possibilities along other facets, refusers can narrow the search space when considering a wider range of actions available to them.

9.5.2 Lessons for Designers: Refusal as a Prompt for Design

What can designers (broadly construed) learn from the framework of data refusal from below? In addition to inspiring new social relations, refusal can also motivate creative new software designs. Sasha Costanza-Chock notes that in many endeavors, the question " 'what's wrong?' drives our pursuit of 'what if?' " [72]. In addition to driving social change, attempts to refuse corporate information systems have also generated fundamental advances in various computing fields. Scholars and activists have created grassroots computer networking infrastructure [134], new theoretical approaches and software plugins for information security [48], and novel research software for community-led experiments [188], to name a few.

These examples of refusal-inspired design also challenge conventional wisdom in computer science about the relationship between design and technology use. Computer scientists and designers who study technology non-use often assume that non-users will become future users with improvements to technology design [223, 231]. That is, they assume that if they implement certain features or streamline certain experiences, the number of users will grow. This standpoint views refusers as either passive resources to be mined for design ideas, or problems to be solved through design. In reality, refusal from below requires active exertion of agency that exceeds what software is designed to offer. Refusal expresses the dissonance between the limited forms of agency that software imposes on users, and the fuller extent of their desires and intentions. In some cases, this dissonance can be resolved with improvements to software design that more faithfully model a user's intention. But in others, they point to differences in substantive values [297] that directly challenge the power relations that a system enables or the politics it expresses — where improving the software means acting against the refusers' interests. This is clearest in the facial recognition double-bind discussed in 9.4.3. A view that exclusion from technology "always and necessarily involves inequality and deprivation" [223] implies that communities of color will be better served by facial recognition systems with improved subgroup accuracy. But as Wyatt reminds us, it is important to "distinguish between 'have nots' (the excluded and the expelled) and 'want nots' (the resisters and the rejecters)" [223]. The anti-facial recognition movement built on many refusals has demonstrated the desirability of exclusion compared to forcible inclusion. In this light, designers should see refusal as a prompt for reflecting on the broader socio-technical context of a system, and how software acts upon people with different subject positions.

9.5.3 Working Toward Institutionalized Refusal

While many cases of refusal are developed ad hoc in response to specific harms that arise, the long term sustainability of refusal requires contending with institutions. Given the scale and reach of governments and corporations developing data systems, what might equally powerful institutional forms oriented around refusal look like? Huybrechts et al. use the term institutioning to highlight the idea that institutions are "highly dynamic and contested spaces where change is not only imposed from the outside but also generated from within," and how movements of participatory design and co-design alter institutional frames [148]. In her writing on refusal, Ruha Benjamin argues that there is a need to *institutionalize* refusal [36] to support people's capacity to collectively organize, consider long-term solutions, and challenge power.

The need to contend with institutions is clear in Ruha Benjamin's case study of a UK Border Agency project using genetic ancestry tests to screen asylum applicants [36]. Even though the Border Agency's ancestry tests were flawed, any individual who refused to comply risked increased scrutiny and potential deportation by the U.K. government. Indeed, caseworkers were "encouraged to regard refusal to submit samples with suspicion" [36]. At the border, Benjamin observed the need for "'second-hand refusals' of those speaking on behalf of asylum seekers" [36]. These refusals came from refugee advocates and academic scientists who publicly criticized the project. For Benjamin, the case at the border "underscores the need to institutionalize informed refusal rather than leaving it to already vulnerable individuals to question scientific and state authority" [36].

Cases of refusers working toward institutionalized refusal can be studied in terms of autonomy, time, power, and cost.

Autonomy. As the work of the Stop LAPD Spying Coalition shows, advocacy groups embedded in local communities can establish structures that help them act as an extension of refusers' autonomy. Benjamin argues that institutionalized refusal should not redirect agency away from refusers toward small groups of institutional elites. Benjamin argues that "rather than rely[ing] on 'secondhand refusals' by public advocates, watch-dogs, and whistleblowers, it is vital to cultivate norms and develop mechanisms that allow those who are targeted by a particular initiative to voice dissent on their own terms" [36]. For example, in the Coalition's campaign against predictive policing, their advocacy was informed by focus groups with community members and their materials included direct quotes from participants voicing their personal concerns in their own words.

Time. Advocates from the ACLU, Algorithmic Justice League, and other organizations working to ban facial recognition are pushing government institutions to change the temporality of refusal. Benjamin writes in the context of genetic testing that institutionalized refusal might involve, at minimum, "greater onus [...] on institutions to incorporate the concerns and insights of prospective research subjects and tissue donors upstream, far in advance of recruitment" [36]. In the context of facial recognition, this means enacting policy that limits harmful government use of facial recognition so that the onus is not on individuals to refuse after the fact.

Power. The FNIGC's OCAP principles, and their use in interactions with the Canadian government, are an example of building power to create institutional change. As Benjamin notes, "[institutions'] norms and practices, as well as existing social hierarchies, place pressure on people to defer to authority" [36]. Rather than defer to the authority of Canadian data collectors, the FNIGC's principles outline desired terms of engagement that asserted their sovereignty and established new practices for data collection, use, and circulation.

Cost. Data cooperatives and other worker-led organizations show how the costs of refusal can be redistributed through collective organizing. In 2018 after prolonged worker organizing, Google canceled Project Maven, which was a contract to produce surveillance software for the U.S. military. Since then, in many cases of worker organizing at Google, activists have experienced retaliation from the company [116]. Although the costs of refusal in this context are still apparent, many workers who participated in these movements were more able to do so because they were not the only ones speaking out. Collective organizing with an eye toward institutional change can push institutions to "actively and genuinely support the choice to refuse participation," which protects individuals who otherwise "are required to risk the fallout from acting autonomously" [36].

My framework helps illuminate the challenges people face today while suggesting aspirational goals for the future. Ultimately, the call to institutionalize refusal is a call to imagine and sustain new forms of collectivity and solidarity over the socio-technical systems that shape our lives.

Chapter 10

Discussion and Future Work

10.1 Participation and Power in Ethical Data Collection

In developing systems and frameworks for consent and refusal in data ethics, the lens of participation and power helped to highlight the strengths and limitations of software for automating ethics and to draw attention to non-participation as a creative act. In chapter 8, I designed a system for automating research ethics procedures for social media data collection. Researchers have approached this problem through a primarily procedural lens, arguing that procedures like consent and opt-out are needed to support individual autonomy. Reflecting on participation and power, I realized that this perspective was incomplete without considering the normative dimensions of the research. In evaluating Bartleby using both procedural and substantive theories, I found that understanding Bartleby's substantive implications required me to consider how power affects the negotiation of normative values between researchers and research subjects.

In response to the limitations of Bartleby for thinking about collective harms, I introduce in chapter 9 a framework that conceptualizes actions people take both individually and collectively to actively withdraw participation in systems as actions oriented toward reshaping those systems. This framework considers socio-technical forms of participation and power as refusers use whatever means available to manage their inclusion in data systems.

10.2 Future Work

This work suggests an opportunity for future work to continue to explore the implications of refusal as an act of design. This may involve designing systems and data infrastructure to support people and communities who are putting data refusal into practice. Future directions may include:

Designing a Bartleby for collective autonomy. Though Bartleby usefully helped manage individual opt out for online social media research, many of its limitations revolve around the idea of individual autonomy as a privacy management paradigm. Future work may explore how designers can create ethics systems that put collective notions of autonomy into practice — for instance, through representative or delegated forms of consent.

Adversarial research ethics. Bartleby operates under the assumption that researchers and research subjects have generally aligned ideas of the common good, even if they differ on how those values are expressed. However, as I found in my work, Bartleby is not well suited to studies that involve higher risk to participants or to researchers — especially when researchers are studying sensitive issues like online extremism. Future work might consider how to think about the ethics of data collection in situations where participants' individual autonomy may be in direct conflict with researchers' autonomy, and/or when there is a compelling public-interest reason to study people who do not want to be studied.

Part IV

Discussion

Chapter 11

Discussion and Future Work

This dissertation presents research in the domains of **accessible data analysis** and **data ethics** that approaches designing empowering systems from two directions: (1) creating systems that broaden participation in empowering creative processes and (2) protecting users' individual and collective agency to reshape disempowering processes.

11.0.1 Summary of Thesis Contributions

Part II instantiated questions of participation and power in the domain of accessible data analysis by demonstrating how the lens of empowered participation can shape design decisions. Chapter 3 introduced the idea of empowered participation as a way to guide design for blind and low-vision screen reader users. Chapter 4 presented a set of design dimensions for screen reader interfaces. These design dimensions are motivated by the need for screen reader user experiences that support self-guided data exploration comparable to sighted user experiences. In contrast to prior best practices that provide static descriptions affording limited user agency, interfaces informed by these dimensions use keyboard-navigable structures that offer descriptions at multiple levels of detail. Chapter 5 expanded on this work to introduce Umwelt, an accessible authoring interface for multimodal data representations that include visualization, structured textual description, and sonification. Unlike previous systems that center the visual modality by requiring users to create non-visual representations by first specifying a visualization, Umwelt derives its multimodal output from a shared data abstraction. Consequently, it supports BLV users in actively participating in the production of data representations by addressing limitations of visual-centric authoring systems. Chapter 6 summarizes this work and explores how the idea of empowered participation can suggest directions for future research in accessibility.

Part III examines participation and power in the domain of data ethics. Here, empowered non-participation is a powerful way for people to shape data collection systems by withholding or withdrawing their participation. Chapter 7 introduces empowered non-participation and its implications for the ethics of data collection. Chapter 8 presents a system called Bartleby, which helps researchers manage the ethics of large-scale social media research by automating an ethics procedure known as debriefing. Debriefing involves a retroactive consent process where users are notified about their involvement in research and given a chance to opt out. To evaluate the success of Bartleby in two field deployments, I draw a distinction

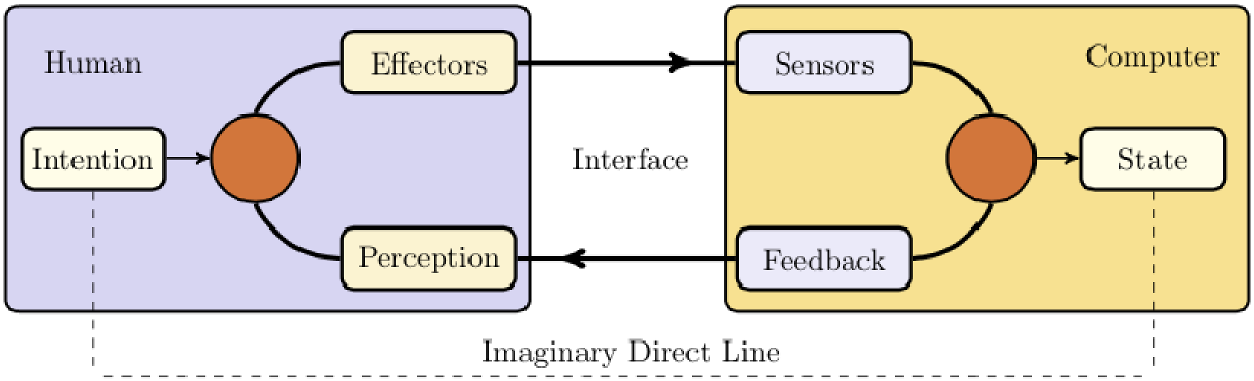


Figure 11.1: *A diagram [263] conveying the conventional model of human-computer interaction as a self-contained feedback loop.*

between *procedural* and *substantive* theories of ethics and discuss Bartleby's implications through both perspectives. Chapter 9 examines the practice of data refusal, an idea scholars have used to conceptually unify disparate practices of "saying no" to data collection. Although computer scientists have typically seen acts of refusal as antithetical to design, I argue that data refusal is a generative act oriented toward reshaping socio-technical systems. I introduce a design framework with four characteristics of refusal, and discuss the uses of this framework to understand, evaluate, and imagine new cases of refusal in practice. Chapter 10 summarizes this work and explores how empowered non-participation can suggest future research in data ethics and refusal.

11.0.2 Directions for Future Work

In this section, I discuss how the contributions my thesis open potential avenues for future work in HCI.

HCI as a Science of Socio-Technical Interaction

In HCI, interaction is conventionally understood as a feedback loop between an individual user and a self-contained system Figure 11.1. This loop is typically depicted as a symmetrical process in which input and output is exchanged. While this model has been foundational for HCI research, it lends itself to a focus on individual user interactions. Consequently, it is less helpful for understanding how those individual users engage in society, or how computer systems are shaped by their use in social processes.

A participation and power lens instead calls attention to the idea that the important part of a human-computer interaction is often how the user wants the computer to facilitate their participation (or non-participation) in a broader social process. This dissertation's contributions in accessibility built on the insight that BLV people want to engage with data with interfaces that are not merely accessible, but facilitate comparable experiences of self-guided exploration and active production of data analysis in order to participate in data-driven discussions. Similarly, this dissertation's contributions to data ethics and refusal recognize that people's engagement with socio-technical systems extends beyond what is typically legible as computer use. Rather, people enact their social and political goals through a number of pragmatic strategies that involve unintended uses or non-uses of technology, as well as advocacy and institution-building.

This work suggests a need to reimagine interaction as a feedback loop between a socially-embedded user and a socio-technical system composed of software, people, and institutions. In this view, the role of a designer is not only to produce software artifacts, but also to arrange social relationships within multi-actor systems.

Embodiment as a Design Consideration for Participation and Power

Though accessibility and data ethics have seemingly little overlap, core issues in both domains appear to share a conceptual grounding in questions of embodiment and design. In accessibility and disability studies, embodiment is a core idea for emphasizing the specificity of people's experiences that arise from bodily differences. Disability scholars argue against the idea of a universally normative body from which a disabled body deviates. Instead, they acknowledge that every person's body has a different set of abilities and limitations — the sum of which interacts with its environment in a way that may differ from others. For example, Rosemarie Garland-Thomson uses the concept of *misfitting* to describe the way that disability emerges from incompatible interactions between bodies and society [108]. In other words, for Garland-Thomson, the cause of disability cannot be located purely in the body or the environment, but in the interaction between people and structures that are designed and built.

Applying ideas about embodiment to HCI uncovers limitations of the conventional way the field thinks about human subjects. Computers sense the world through inputs and outputs, but HCI researchers have also conceptualized people as I/O machines. Influenced by cognitive science and cybernetics, the field theorizes interaction as a feedback loop between a user and a system [234]. In this model, the user is essentially an information processing machine. The user has a sense input (e.g. vision), a control (some cognitive map of their goals and intentions), and an articulatory output (e.g. the ability to move a computer mouse).

The term "user," though seemingly referencing personhood, is best understood as the particular way HCI's underlying theoretical framework constructs and universalizes a subject. HCI researchers constructed this model in order to make the concept of a person operationalizable in computer systems. To be understood by machines, humans had to conform to a machine-like schema of input and output. As a result, Lasse Scherffig writes, "the human trained to perform in front of the computer became the model for the thinking human in general—a human acting as a computer" [234]. In other words, the idea of a "user" attempts to universally stand in for all people, but in reality only captures a specific, limited understanding of what a person is.

A perspective informed by a disability studies lens on embodiment instead recognizes that people will not conform to this idealized normative "user" model due to bodily difference. Computer science educators have illustrated the universal body that it assumes—"how the computer sees us"—as a single eye augmented with a single finger (Figure 11.2) [206]. As bizarre as it looks, the eye-finger-subject is illustrative of the way the field of HCI thinks about the human sensorium in terms of interface modalities. The eye and ears represent the

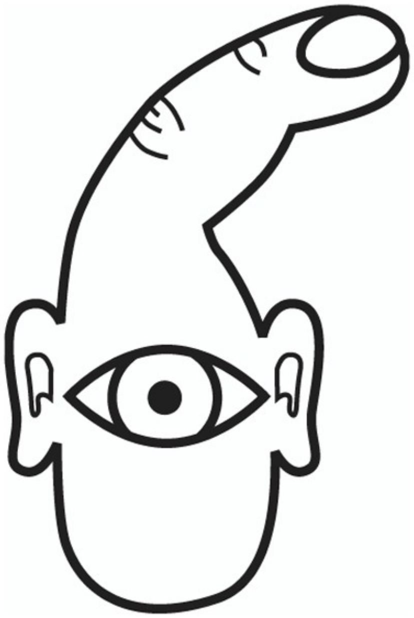


Figure 11.2: *"How the computer sees us," an illustration from a computing textbook [206] that depicts how a computer assumes the presence of a certain sensory-motor apparatus when modeling humans as information processors.*

human perceptual capacities that computers often use to output data, by rendering it visible or audible. The single finger represents a primary way computers receive human input: through pointing, or through the mechanical actuation of mouse and keyboard buttons. Recognizing that this model will not describe everyone suggests that there is no one-size-fits-all approach to user-centered design. Instead, approaches to design should aspire to be flexible enough to support a user's goals in the way that best adapts to their cognitive and sensory faculties.

Embodiment is also an important theme in scholarship on data ethics, privacy, and surveillance, which considers how data is produced from the measurement of the body. As people perform ordinary activities on their computers through pointing, typing, and other forms of input. These movements often generate additional data as software records measurements of activity during everyday use—often without users' knowledge. Logs of mouse movements, records of keystrokes, amount of time spent on a webpage; Melissa Gregg compares this excess data to sweat, which "literalizes porosity" and is a "means by which the body signals its capacity to 'affect and be affected'" [117]. Biometric data collected in the background of computer use is then used to select, differentiate, identify, and classify people.

Biometric profiles exemplify the process through which computers model and process humans as data objects—more precisely, objects assembled from the accumulation of data. For instance, proponents of digital psychiatry claim to be able to use biometric signals to diagnose and pathologize [157]. As a result, a market for biometric software that collects large amounts of data on key press timing has emerged in digital healthcare. This software models the user as a collection of behavioral facts. It defines logical criteria through which computers can define selections of users on the basis of these facts. As anthropologist Beth

Semel notes, “diagnoses also operate as vectors of social control” as people are partitioned into categories of well and unwell, deserving and undeserving of clinical attention [237]. Inclusion and exclusion in these selection criteria consequently affect people’s ability to navigate digitally-managed healthcare systems.

Instantiating embodiment as a design consideration in seemingly-different domains can help designers think about how technology mediates participation in society. Lessons from the disability and surveillance contexts demonstrate that interaction with everyday software and broader socio-technical systems (like healthcare) is affected by how those systems model the human body. In future work, ideas like embodied difference and bodily autonomy can inform approaches to design that improve upon conventional approaches to embodiment in human-computer interaction.

Bibliography

[1] Data moving to U.S. very soon!, June 2005. URL https://web.archive.org/web/20050615020306/https://www.flickr.com/forums/help/8280/. Publication Title: Flickr.

[2] Community IRBS & Research Review Boards: Shaping the Future of Community-Engaged Research. Technical report, Albert Einstein College of Medicine, The Bronx Health Link and Community-Campus Partnerships for Health, 2012. URL https://www.ccphealth.org/wp-content/uploads/2017/10/Shaping_the_Future_of_CEnR.pdf.

[3] Bronx Community Research Review Board » What is CERA?, 2016. URL http://bxcrrb.org/cera-project/what-is-cera/.

[4] SAS Graphics Accelerator, 2017. URL https://support.sas.com/software/products/graphics-accelerator/index.html.

[5] California Consumer Privacy Act (CCPA), October 2018. URL https://oag.ca.gov/privacy/ccpa.

[6] General Data Protection Regulation (GDPR), 2018. URL https://gdpr-info.eu/.

[7] Platform manipulation and spam policy, September 2019. URL https://help.twitter.com/en/rules-and-policies/platform-manipulation.

[8] Tim Janecyk et al. v. International Business Machines Corp., 2020. URL https://www.bloomberglaw.com/iframe/document/X1Q6O5FJFL82/download.

[9] Exploring legal mechanisms for data stewardship. Technical report, Ada Lovelace Institute, March 2021. URL https://www.adalovelaceinstitute.org/wp-content/uploads/2021/03/Legal-mechanisms-for-data-stewardship_report_Ada_AI-Council-2.pdf.

[10] SAS Graphics Accelerator. SAS Graphics Accelerator Customer Product Page, 2018. URL https://support.sas.com/software/products/graphics-accelerator/index.html.

[11] Philip E Agre. Toward a Critical Technical Practice: Lessons Learned Trying to Reform AI. In Geoffrey C. Bowker, Susan Leigh Star, William Turner, and Les Gasser, editors, *Social science, technical systems, and cooperative work: beyond the great divide*, Computers, cognition, and work. Lawrence Erlbaum Associates, Mahwah, N.J, 1997. ISBN 978-0-8058-2402-5 978-0-8058-2403-2.

[12] Andrea Akkad, Clare Jackson, Sara Kenyon, Mary Dixon-Woods, Nick Taub, and Marwan Habiba. Patients' perceptions of written consent: questionnaire study. *Bmj*, 333(7567):528, 2006. Publisher: British Medical Journal Publishing Group.

[13] Frances K. Aldrich and Linda Sheppard. Tactile Graphics In School Education: Perspectives From Pupils. *British Journal of Visual Impairment*, 19(2):69–73, 2001. ISSN 0264-6196. doi:10.1177/026461960101900204. URL https://doi.org/10.1177/026461960101900204.

[14] Alice Wong and Ed Yong. What Counts as Seeing. *Orion Magazine*, 2022. URL https://orionmagazine.org/article/ed-yong-alice-wong-interview/.

[15] Nancy Amick and Jane Corcoran. Guidelines for Design of Tactile Graphics. Technical report, American Printing House for the Blind, 1997. URL https://www.aph.org/research/guides/.

[16] Sareeta Amrute. Of Techno-Ethics and Techno-Affects. *Feminist Review*, 123(1):56–73, November 2019. ISSN 0141-7789. doi:10.1177/0141778919879744. URL https://doi.org/10.1177/0141778919879744. Publisher: SAGE Publications.

[17] Sherry R. Arnstein. A Ladder Of Citizen Participation. *Journal of the American Institute of Planners*, 35(4):216–224, July 1969. ISSN 0002-8991. doi:10.1080/01944366908977225. URL https://doi.org/10.1080/01944366908977225.

[18] A. Arora, S. Rajagopalan, N. Shafiq, P. Pandhi, A. Bhalla, D. P. Dhibar, and S. Malhotra. Development of tool for the assessment of comprehension of informed consent form in healthy volunteers participating in first-in-human studies. *Contemporary clinical trials*, 32(6):814–817, 2011. Publisher: Elsevier.

[19] Susan Athey, Christian Catalini, and Catherine Tucker. The Digital Privacy Paradox: Small Money, Small Costs, Small Talk. Working Paper 23488, National Bureau of Economic Research, June 2017. URL http://www.nber.org/papers/w23488. Series: Working Paper Series.

[20] Susan Athey, Christian Catalini, and Catherine Tucker. The Digital Privacy Paradox: Small Money, Small Costs, Small Talk. *MIT Sloan Research Paper No. 5196-17*, page 34, 2018.

[21] Richard P. Bagozzi. The Legacy of the Technology Acceptance Model and a Proposal for a Paradigm Shift. *Journal of the Association for Information Systems*, 8(4), 2007. ISSN 1536-9323. doi:10.17705/1jais.00122. URL https://aisel.aisnet.org/jais/vol8/iss4/12.

[22] Catherine M. Baker, Lauren R. Milne, Jeffrey Scofield, Cynthia L. Bennett, and Richard E. Ladner. Tactile graphics with a voice: using QR codes to access text in tactile graphics. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, ASSETS '14, pages 75–82, New York, NY, USA, October 2014. Association for Computing Machinery. ISBN 978-1-4503-2720-6. doi:10.1145/2661334.2661366. URL https://doi.org/10.1145/2661334.2661366.

[23] Catherine M. Baker, Lauren R. Milne, Ryan Drapeau, Jeffrey Scofield, Cynthia L. Bennett, and Richard E. Ladner. Tactile Graphics with a Voice. *ACM Transactions on Accessible Computing (TACCESS)*, 8(1):3:1–3:22, 2016. ISSN 1936-7228. doi:10.1145/2854005. URL http://doi.org/10.1145/2854005.

[24] Jack M Balkin. Information Fiduciaries and the First Amendment. *UC Davis Law Review*, 49(4):52, 2016.

[25] Chelsea Barabas. To Build a Better Future, Designers Need to Start Saying 'No', October 2020. URL https://onezero.medium.com/refusal-a-beginning-that-starts-with-an-end-2b055bfc14be.

[26] Chelsea Barabas. Refusal in Data Ethics: Re-Imagining the Code Beneath the Code of Computation in the Carceral State. *Engaging Science, Technology, and Society*, 8 (2):35–57, September 2022. ISSN 2413-8053. doi:10.17351/ests2022.1233. URL https://estsjournal.org/index.php/ests/article/view/1233. Number: 2.

[27] Kimberly A. Barchard and John Williams. Practical advice for conducting ethical online experiments and questionnaires for United States psychologists. *Behavior Research Methods*, 40(4):1111–1128, November 2008. ISSN 1554-351X, 1554-3528. doi:10.3758/BRM.40.4.1111. URL http://link.springer.com/10.3758/BRM.40.4.1111.

[28] Shaowen Bardzell. Feminist HCI: taking stock and outlining an agenda for design. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, page 1301, Atlanta, Georgia, USA, 2010. ACM Press. ISBN 978-1-60558-929-9. doi:10.1145/1753326.1753521. URL http://portal.acm.org/citation.cfm?doid=1753326.1753521.

[29] Shaowen Bardzell and Jeffrey Bardzell. Towards a feminist HCI methodology: social science, feminism, and HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 675–684, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 978-1-4503-0228-9. doi:10.1145/1978942.1979041. URL https://doi.org/10.1145/1978942.1979041. event-place: Vancouver, BC, Canada.

[30] Solon Barocas and Karen Levy. Privacy Dependencies. *Washington Law Review*, 95 (2):555–616, 2020. URL https://heinonline.org/HOL/P?h=hein.journals/washlr95&i=561.

[31] Stephen Barrass and Gregory Kramer. Using Sonification. *Multimedia Systems*, 7(1): 23–31, 1999. ISSN 1432-1882. doi:10.1007/s005300050108. URL https://doi.org/10.1007/s005300050108.

[32] Eric P.S. Baumer and M. Six Silberman. When the implication is not to design (technology). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2271–2274, Vancouver BC Canada, May 2011. ACM. ISBN 978-1-4503-0228-9. doi:10.1145/1978942.1979275. URL https://dl.acm.org/doi/10.1145/1978942.1979275.

[33] Michel Beaudouin-Lafon. Designing interaction, not interfaces. In *Proceedings of the working conference on Advanced visual interfaces*, AVI '04, pages 15–22, New York, NY, USA, May 2004. Association for Computing Machinery. ISBN 978-1-58113-867-2. doi:10.1145/989863.989865. URL http://doi.org/10.1145/989863.989865.

[34] Richard A. Becker, William S. Cleveland, and Ming-Jen Shyu. The Visual Design and Control of Trellis Display. *Journal of Computational and Graphical Statistics*, 5 (2):123–155, 1996. ISSN 1061-8600, 1537-2715. doi:10.1080/10618600.1996.10474701. URL http://www.tandfonline.com/doi/abs/10.1080/10618600.1996.10474701.

[35] Ruha Benjamin. *People's science: Bodies and rights on the stem cell frontier*. Stanford University Press, 2013.

[36] Ruha Benjamin. Informed Refusal: Toward a Justice-based Bioethics. *Science, Technology, & Human Values*, June 2016. doi:10.1177/0162243916656059. URL https://journals.sagepub.com/doi/10.1177/0162243916656059.

[37] Cynthia L. Bennett, Erin Brady, and Stacy M. Branham. Interdependence as a Frame for Assistive Technology Research and Design. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 161–173, Galway Ireland, October 2018. ACM. ISBN 978-1-4503-5650-3. doi:10.1145/3234695.3236348. URL https://dl.acm.org/doi/10.1145/3234695.3236348.

[38] Jacques Bertin. *Semiology of Graphics*. University of Wisconsin Press, 1983. ISBN 978-0-299-09060-9. URL https://dl.acm.org/doi/10.5555/1095597.

[39] S. Bhansali, N. Shafiq, S. Malhotra, P. Pandhi, Inderjeet Singh, S. P. Venkateshan, S. Siddhu, Y. P. Sharma, and K. K. Talwar. Evaluation of the ability of clinical research participants to comprehend informed consent form. *Contemporary clinical trials*, 30 (5):427–430, 2009. Publisher: Elsevier.

[40] A. F. Blackwell, C. Britton, A. Cox, T. R. G. Green, C. Gurr, G. Kadoda, M. S. Kutar, M. Loomes, C. L. Nehaniv, M. Petre, C. Roast, C. Roe, A. Wong, and R. M. Young. Cognitive Dimensions of Notations: Design Tools for Cognitive Technology. In Meurig Beynon, Chrystopher L. Nehaniv, and Kerstin Dautenhahn, editors, *Cognitive Technology: Instruments of Mind*, Lecture Notes in Computer Science, pages 325–341, Berlin, Heidelberg, 2001. Springer. ISBN 978-3-540-44617-0. doi:10.1007/3-540-44617-6_31.

[41] Matthew Blanco, Jonathan Zong, and Arvind Satyanarayan. Olli: An Extensible Visualization Library for Screen Reader Accessibility. In *IEEE VIS Posters*, 2022. URL http://vis.csail.mit.edu/pubs/olli/.

[42] Tal Boger, Steven B. Most, and Steven L. Franconeri. Jurassic Mark: Inattentional Blindness for a Datasaurus Reveals that Visualizations are Explored, not Seen. In *IEEE Transactions on Visualization & Computer Graphics (Proc. IEEE VIS)*, 2021. doi:10.1109/VIS49827.2021.9623273. URL https://ieeexplore.ieee.org/abstract/document/9623273.

[43] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. D3: Data-Driven Documents. *IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis)*, 2011. doi:10.1109/TVCG.2011.185. URL http://vis.stanford.edu/papers/d3.

[44] Alice Boxhall, James Craig, Dominic Mazzoni, and Alexander Surkov. The Accessibility Object Model (AOM), 2022. URL https://wicg.github.io/aom/.

[45] Cristian Bravo-Lillo, Serge Egelman, Cormac Herley, Stuart Schechter, and Janice Tsai. You Needn't Build That: Reusable Ethics-Compliance Infrastructure for Human Subjects Research. In *CREDS 2013 (Workshop)*, May 2013. URL https://www.microsoft.com/en-us/research/publication/you-neednt-build-that-reusable-ethics-compliance-infrastructure-for-human-subjects-research/.

[46] Matthew Brehmer and Tamara Munzner. A Multi-Level Typology of Abstract Visualization Tasks. In *IEEE Transactions on Visualization & Computer Graphics (Proc. IEEE VIS)*, 2013. doi:10.1109/TVCG.2013.124. URL http://ieeexplore.ieee.org/document/6634168/.

[47] Anke Brock, Philippe Truillet, Bernard Oriola, and Christophe Jouffrais. Usage Of Multimodal Maps For Blind People: Why And How. In *ACM International Conference on Interactive Tabletops and Surfaces (ISS)*, ITS '10, pages 247–248, New York, NY, USA, 2010. ISBN 978-1-4503-0399-6. doi:10.1145/1936652.1936699. URL https://doi.org/10.1145/1936652.1936699.

[48] Finn Brunton and Helen Nissenbaum. *Obfuscation: a users guide for privacy and protest*. MIT Press, 2016.

[49] Joy Buolamwini. Face: The Final Frontier of Privacy — Full Spoken Congressional Testimony May 22, 2019, July 2019. URL https://medium.com/@Joy.Buolamwini/face-the-final-frontier-of-privacy-full-spoken-congressional-testimony-may-22-2019-ff8607df045b.

[50] Joy Buolamwini. We Must Fight Face Surveillance to Protect Black Lives, June 2020. URL https://onezero.medium.com/we-must-fight-face-surveillance-to-protect-black-lives-5ffcd0b4c28a.

[51] Alyxander Burns, Christiana Lee, Ria Chawla, Evan Peck, and Narges Mahyar. Who Do We Mean When We Talk About Visualization Novices? In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pages 1–16, New York, NY, USA, April 2023. Association for Computing Machinery. ISBN 978-1-4503-9421-5. doi:10.1145/3544548.3581524. URL https://dl.acm.org/doi/10.1145/3544548.3581524.

[52] Matthew Butler, Leona M Holloway, Samuel Reinders, Cagatay Goncu, and Kim Marriott. Technology Developments in Touch-Based Accessible Graphics: A Systematic Review of Research 2010-2020. In *ACM Conference on Human Factors in Computing Systems (CHI)*, number 278, pages 1–15. New York, NY, USA, 2021. ISBN 978-1-4503-8096-6. URL https://doi.org/10.1145/3411764.3445207.

[53] Susanne Bødker. Third-wave HCI, 10 years later—participation and sharing. *Interactions*, 22(5):24–31, August 2015. ISSN 1072-5520. doi:10.1145/2804405. URL https://dl.acm.org/doi/10.1145/2804405.

[54] Carole Cadwalladr and Emma Graham-Harrison. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*, March 2018. ISSN 0261-3077. URL https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

[55] Stanley J. Cantrell, Bruce N. Walker, and Øystein Moseng. Highcharts Sonification Studio: An Online, Open-Source, Extensible, and Accessible Data Sonification Tool. In *Proceedings of the 26th International Conference on Auditory Display (ICAD 2021)*, pages 210–216, Virtual Conference, June 2021. International Community for Auditory Display. ISBN 978-0-9670904-7-4. doi:10.21785/icad2021.005. URL http://hdl.handle.net/1853/66348.

[56] Nathalie Casemajor, Stéphane Couture, Mauricio Delfin, Matthew Goerzen, and Alessandro Delfanti. Non-participation in digital media: toward a framework of mediated political action. *Media, Culture & Society*, 37(6):850–866, September 2015. ISSN 0163-4437. doi:10.1177/0163443715584098. URL https://doi.org/10.1177/0163443715584098.

[57] Kathy Charmaz. *Constructing Grounded Theory*. Sage Publications, London ; Thousand Oaks, Calif, 2006. ISBN 978-0-7619-7352-2 978-0-7619-7353-9. URL https://uk.sagepub.com/en-gb/eur/constructing-grounded-theory/book235960.

[58] Cindy Chen, Pou-I Lee, Kevin J. Pain, Diana Delgado, Curtis L. Cole, and Thomas R. Campion. Replacing Paper Informed Consent with Electronic Informed Consent for Research in Academic Medical Centers: A Scoping Review. *AMIA Summits on Translational Science Proceedings*, 2020:80–88, May 2020. ISSN 2153-4063. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7233043/.

[59] Emil Chiauzzi and Paul Wicks. Digital Trespass: Ethical and Terms-of-Use Violations by Researchers Accessing Data From an Online Patient Community. *Journal of Medical Internet Research*, 21(2), February 2019. ISSN 1439-4456. doi:10.2196/11985. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6403524/.

[60] Jinho Choi, Sanghun Jung, Deok Gun Park, Jaegul Choo, and Niklas Elmqvist. Visualizing for the Non-Visual: Enabling the Visually Impaired to Use Visualization. *Computer Graphics Forum (EuroVis)*, 38(3):249–260, June 2019. ISSN 0167-7055, 1467-8659. doi:10.1111/cgf.13686. URL https://onlinelibrary.wiley.com/doi/10.1111/cgf.13686.

[61] Pramod Chundury, Biswaksen Patnaik, Yasmin Reyazuddin, Christine Tang, Jonathan Lazar, and Niklas Elmqvist. Towards Understanding Sensory Substitution for Accessible Visualization: An Interview Study. *IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE VIS)*, pages 1–1, 2021. ISSN 1077-2626, 1941-0506, 2160-

9306. doi:10.1109/TVCG.2021.3114829. URL https://ieeexplore.ieee.org/document/9552177/.

[62] M. Cifor, P. Garcia, T.L. Cowan, J. Rault, T. Sutherland, A. Chan, J. Rode, A.L. Hoffmann, N. Salehi, and L. Nakamura. Feminist Data Manifest-No, 2019. URL https://www.manifestno.com/.

[63] Citizens and Technology Lab. Freedom of Expression in Marginalized Groups Online, June 2020. URL https://citizensandtech.org/research/expression-in-marginalized-groups/.

[64] Citizens and Technology Lab. Automated Law Enforcement, May 2021. URL https://citizensandtech.org/research/automated-law-enforcement/.

[65] Danielle Keats Citron and Daniel J. Solove. Privacy Harms. SSRN Scholarly Paper ID 3782222, Social Science Research Network, Rochester, NY, February 2021. URL https://papers.ssrn.com/abstract=3782222.

[66] E F Codd. Further Normalization of the Data Base Relational Model. Technical Report RJ909, IBM, San Jose, California, 1971.

[67] Enrico Coiera and Roger Clarke. e-Consent: The design and implementation of consumer consent mechanisms in an electronic environment. *Journal of the American medical informatics association*, 11(2):129–140, 2004. Publisher: BMJ Group BMA House, Tavistock Square, London, WC1H 9JR.

[68] Kate Conger, Richard Fausset, and Serge F. Kovaleski. San Francisco Bans Facial Recognition Technology. *The New York Times*, May 2019. ISSN 0362-4331. URL https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html.

[69] Andrea Cornwall and Rachel Jewkes. What Is Participatory Research? *Social Science & Medicine*, 41(12):1667–1676, 1995. ISSN 0277-9536. doi:10.1016/0277-9536(95)00127-S. URL https://www.sciencedirect.com/science/article/pii/027795369500127S.

[70] Sasha Costanza-Chock. Design Justice, A.I., and Escape from the Matrix of Domination. *Journal of Design and Science*, July 2018. doi:10.21428/96c8d426. URL https://jods.mitpress.mit.edu/pub/costanza-chock.

[71] Sasha Costanza-Chock. *Design Justice: Towards an Intersectional Feminist Framework for Design Theory and Practice*. MIT Press, Cambridge, MA, 2020. URL https://papers.ssrn.com/abstract=3189696.

[72] Sasha Costanza-Chock. Directions for Future Work: From #TechWontBuildIt to #DesignJustice. In *Design Justice*. February 2020. ISBN 978-0-262-04345-8. URL https://design-justice.pubpub.org/pub/ev26fjji/release/1.

[73] Nick Couldry and Ulises A. Mejias. Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject. *Television & New Media*, 20(4):336–349, May 2019. ISSN 1527-4764. doi:10.1177/1527476418796632. URL https://doi.org/10.1177/1527476418796632. Publisher: SAGE Publications.

[74] Scott Davert and AppleVis Editorial Team. What's New In iOS 13 Accessibility For Individuals Who Are Blind or Deaf-Blind, 2019. URL https://www.applevis.com/blog/whats-new-ios-13-accessibility-individuals-who-are-blind-or-deaf-blind.

[75] Lilian de Greef, Dominik Moritz, and Cynthia Bennett. Interdependent Variables: Remotely Designing Tactile Graphics for an Accessible Workflow. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '21, pages 1–6, New York, NY, USA, October 2021. Association for Computing Machinery. ISBN 978-1-4503-8306-6. doi:10.1145/3441852.3476468. URL http://doi.org/10.1145/3441852.3476468.

[76] Sylvie Delacroix and Neil D Lawrence. Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance. *International Data Privacy Law*, 9 (4):236–252, October 2019. ISSN 2044-3994. doi:10.1093/idpl/ipz014. URL https://doi.org/10.1093/idpl/ipz014. _eprint: https://academic.oup.com/idpl/article-pdf/9/4/236/32744015/ipz014.pdf.

[77] Scott Desposato. Subjects and Scholars' Views on the Ethics of Political Science Field Experiments. *Perspectives on Politics*, 16(3):739–750, September 2018. ISSN 1537-5927, 1541-0986. doi:10.1017/S1537592717004297. URL https://www.cambridge.org/core/journals/perspectives-on-politics/article/subjects-and-scholars-views-on-the-ethics-of-political-science-field-experiments/1A152D85B81C5F9D7FE82CA37494FEDC. Publisher: Cambridge University Press.

[78] Paul DiMaggio, Eszter Hargittai, Coral Celeste, and Steven Shafer. Digital inequality: From unequal access to differentiated use. *Social inequality*, pages 355–400, 2004. Publisher: Russell Sage Foundation.

[79] Evanthia Dimara and Charles Perin. What is Interaction for Data Visualization? *IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE VIS)*, 26(1):119 – 129, 2020. doi:10.1109/TVCG.2019.2934283. URL https://hal.archives-ouvertes.fr/hal-02197062.

[80] Carl DiSalvo. Critical Making as Materializing the Politics of Design. *The Information Society*, 30(2):96–105, March 2014. ISSN 0197-2243. doi:10.1080/01972243.2014.875770. URL https://doi.org/10.1080/01972243.2014.875770.

[81] Donald E. Stokes. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Brookings Institution Press, August 1997. ISBN 978-0-8157-8178-3.

[82] Doug Schepers. Designing Data for Cognitive Load. In Jonathan Schwabish, Sue Popkin, and Alice Feng, editors, *Centering Accessibility in Data Visualization*. Urban Institute, December 2022. URL https://www.urban.org/sites/default/files/2022-12/Do%20No%20Harm%20Guide%20Centering%20Accessibility%20in%20Data%20Visualization.pdf.

[83] Edward S. Dove, Susan E. Kelly, Federica Lucivero, Mavis Machirori, Sandi Dheensa, and Barbara Prainsack. Beyond individualism: Is there a place for relational autonomy in clinical practice and research?:. *Clinical Ethics*, April 2017. doi:10.1177/1477750917704156. URL https://journals.sagepub.com/doi/10.1177/1477750917704156.

[84] Graham M. Dresden and M. Andrew Levitt. Modifying a standard industry clinical trial consent form improves patient information retention as part of the informed consent process. *Academic Emergency Medicine*, 8(3):246–252, 2001. Publisher: Wiley Online Library.

[85] Drew Harwell and Geoffrey A. Fowler. U.S. Customs and Border Protection says photos of travelers were taken in a data breach. *Washington Post*, June 2019. ISSN 0190-8286. URL https://www.washingtonpost.com/technology/2019/06/10/us-customs-border-protection-says-photos-travelers-into-out-country-were-recently-taken-data-breac

[86] Anthony Dunne and Fiona Raby. *Speculative Everything: Design, Fiction, and Social Dreaming*. The MIT Press, 2013. ISBN 978-0-262-01984-2. URL https://www.jstor.org/stable/j.ctt9qf7j7.

[87] @EmilLeong. New Ways to Control Your Experience on Twitter, August 2016. URL https://blog.twitter.com/en_us/a/2016/new-ways-to-control-your-experience-on-twitter.html.

[88] Emmanuel Morgan. Group that sued LAPD over controversial data policing programs claims victory. *Los Angeles Times*, December 2019. URL https://www.latimes.com/california/story/2019-12-10/stop-lapd-spying-coalition-announces-lawsuit-victory-against-lapd.

[89] Christin Engel, Emma Franziska Müller, and Gerhard Weber. SVGPlott: an accessible tool to generate highly adaptable, accessible audio-tactile charts for and from blind and visually impaired people. In *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA '19, pages 186–195, New York, NY, USA, June 2019. Association for Computing Machinery. ISBN 978-1-4503-6232-0. doi:10.1145/3316782.3316793. URL https://dl.acm.org/doi/10.1145/3316782.3316793.

[90] David Enoch. Hypothetical Consent and the Value(s) of Autonomy. *Ethics*, 128(1):6–36, October 2017. ISSN 0014-1704, 1539-297X. doi:10.1086/692939. URL https://www.journals.uchicago.edu/doi/10.1086/692939.

[91] Lynn Chaikin Epstein and Louis Lasagna. Obtaining informed consent: form or substance. *Archives of Internal Medicine*, 123(6):682–688, 1969. Publisher: American Medical Association.

[92] Sheena Erete, Aarti Israni, and Tawanna Dillahunt. An intersectional approach to designing in the margins. *Interactions*, 25(3):66–69, April 2018. ISSN 1072-5520. doi:10.1145/3194349. URL https://doi.org/10.1145/3194349. Place: New York, NY, USA Publisher: Association for Computing Machinery.

[93] Erika Harano, Grant Chinn, Boaz Sender, and Una Lee. FRIES: What good consent looks like in sign-up processes, 2019. URL https://privacy.shorensteincenter.org/fries.

[94] Danyang Fan, Alexa Fay Siu, Wing-Sum Adrienne Law, Raymond Ruihong Zhen, Sile O'Modhrain, and Sean Follmer. Slide-Tone and Tilt-Tone: 1-DOF Haptic Techniques for Conveying Shape Characteristics of Graphs to Blind Users. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–19, New York, NY, USA, April 2022. Association for Computing Machinery. ISBN 978-1-4503-9157-3. doi:10.1145/3491102.3517790. URL https://dl.acm.org/doi/10.1145/3491102.3517790.

[95] Danyang Fan, Alexa Fay Siu, Hrishikesh Rao, Gene Sung-Ho Kim, Xavier Vazquez, Lucy Greco, Sile O'Modhrain, and Sean Follmer. The Accessibility of Data Visualizations on the Web for Screen Reader Users: Practices and Experiences During COVID-19. *ACM Transactions on Accessible Computing*, 16(1):4:1–4:29, March 2023. ISSN 1936-7228. doi:10.1145/3557899. URL https://dl.acm.org/doi/10.1145/3557899.

[96] Casey Fiesler and Nicholas Proferes. "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society*, 4(1):2056305118763366, January 2018. ISSN 2056-3051. doi:10.1177/2056305118763366. URL https://doi.org/10.1177/2056305118763366.

[97] First Nations Information Governance Centre. Barriers and Levers for the Implementation of OCAP™. *International Indigenous Policy Journal*, 5(2), April 2014. ISSN 1916-5781. doi:10.18584/iipj.2014.5.2.3. URL https://ojs.lib.uwo.ca/index.php/iipj/article/view/7430.

[98] First Nations Information Governance Centre. Pathways to First Nations' data and information sovereignty. In Tahu Kukutai and John Taylor, editors, *Indigenous Data Sovereignty*. ANU Press, 1st edition, November 2016. ISBN 978-1-76046-030-3. doi:10.22459/CAEPR38.11.2016.08. URL http://press-files.anu.edu.au/downloads/press/n2140/pdf/ch08.pdf.

[99] Brian T. Fitzpatrick. Do Class Actions Deter Wrongdoing? *SSRN Electronic Journal*, 2017. ISSN 1556-5068. doi:10.2139/ssrn.3020282. URL https://www.ssrn.com/abstract=3020282.

[100] Fernando Flores, Michael Graves, Brad Hartfield, and Terry Winograd. Computer systems and the design of organizational interaction. *ACM Transactions on Information*

Systems, 6(2):153–172, April 1988. ISSN 1046-8188. doi:10.1145/45941.45943. URL http://doi.org/10.1145/45941.45943.

[101] Marion Fourcade and Daniel N Kluttz. A Maussian bargain: Accumulation by gift in the digital economy. *Big Data & Society*, 7(1):2053951719897092, January 2020. ISSN 2053-9517. doi:10.1177/2053951719897092. URL https://doi.org/10.1177/2053951719897092. Publisher: SAGE Publications Ltd.

[102] Erik Frøkjær and Kasper Hornbæk. Cooperative Usability Testing: Complementing Usability Tests With User-Supported Interpretation Sessions. In *ACM Extended Abstracts on Human Factors in Computing Systems (CHI)*, CHI EA '05, pages 1383–1386, New York, NY, USA, 2005. ISBN 978-1-59593-002-6. doi:10.1145/1056808.1056922. URL https://doi.org/10.1145/1056808.1056922.

[103] Mamoru Fujiyoshi, Akio Fujiyoshi, Hiroshi Tanaka, and Toru Ishida. Universal Design Tactile Graphics Production System BPLOT4 for Blind Teachers and Blind Staffs to Produce Tactile Graphics and Ink Print Graphics of High Quality. In Klaus Miesenberger and Georgios Kouroupetroglou, editors, *Computers Helping People with Special Needs*, volume 10897, pages 167–176. Springer International Publishing, Cham, 2018. ISBN 978-3-319-94273-5 978-3-319-94274-2. URL http://link.springer.com/10.1007/978-3-319-94274-2_23.

[104] Allie Funk. I Opted Out of Facial Recognition at the Airport—It Wasn't Easy. *Wired*, July 2019. ISSN 1059-1028. URL https://www.wired.com/story/opt-out-of-facial-recognition-at-the-airport/. Section: tags.

[105] Maya Indira Ganesh and Emanuel Moss. Resistance and refusal to algorithmic harms: Varieties of 'knowledge projects'. *Media International Australia*, 183(1):90–106, 2022. doi:10.1177/1329878X221076288. URL https://doi.org/10.1177/1329878X221076288. _eprint: https://doi.org/10.1177/1329878X221076288.

[106] Seeta Gangadharan. Digital Exclusion: A Politics of Refusal. In Lucy Bernholz, Hélène Landemore, and Rob Reich, editors, *Digital Technology and Democratic Theory*, pages 113–140. University of Chicago Press, 2021. doi:doi:10.7208/9780226748603-005. URL https://doi.org/10.7208/9780226748603-005.

[107] Patricia Garcia, Tonia Sutherland, Niloufar Salehi, Marika Cifor, and Anubha Singh. No! Re-imagining Data Practices Through the Lens of Critical Refusal. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–20, November 2022. ISSN 2573-0142. doi:10.1145/3557997. URL https://dl.acm.org/doi/10.1145/3557997.

[108] Rosemarie Garland-Thomson. Misfits: A Feminist Materialist Disability Concept. *Hypatia*, 26(3):591–609, 2011. ISSN 0887-5367. URL https://www.jstor.org/stable/23016570. Publisher: Hypatia, Inc.

[109] Elisabeth Geier. 5 Important Considerations for Evaluating IRB Systems, May 2019. URL https://www.badrabbit.com/news/2019/5/2/evaluating-irb-systems.

[110] Tarleton Gillespie. Content moderation, AI, and the question of scale. *Big Data & Society*, 7(2):205395172094323, July 2020. ISSN 2053-9517, 2053-9517. doi:10.1177/2053951720943234. URL http://journals.sagepub.com/doi/10.1177/2053951720943234.

[111] A. Jonathan R. Godfrey and M. Theodor Loots. Advice From Blind Teachers on How to Teach Statistics to Blind Students. *Journal of Statistics Education*, 23(3):null, 2015. ISSN null. doi:10.1080/10691898.2015.11889746. URL https://doi.org/10.1080/10691898.2015.11889746.

[112] Goldenfein, Jake. 'Dropouts': An Interview with Lily Irani, J. Khadijah Abdurahman and Alex Hanna. In *Economies of Virtue – The Circulation of 'Ethics' in AI*, number 46 in Theory on Demand. Institute of Network Cultures, Amsterdam, 2022. ISBN 978-94-92302-96-0. URL https://networkcultures.org/blog/publication/economies-of-virtue-the-circulation-of-ethics-in-ai/.

[113] Chandell Gosse, George Veletsianos, Jaigris Hodson, Shandell Houlden, Tonia A. Dousay, Patrick R. Lowenthal, and Nathan Hall. The hidden costs of connectivity: nature and effects of scholars' online harassment. *Learning, Media and Technology*, 46(3):264–280, July 2021. ISSN 1743-9884. doi:10.1080/17439884.2021.1878218. URL https://doi.org/10.1080/17439884.2021.1878218. Publisher: Routledge _eprint: https://doi.org/10.1080/17439884.2021.1878218.

[114] Bryan Gould, Trisha O'Connell, and Geoffrey Freed. Effective Practices for Description of Science Content within Digital Talking Books. Technical report, The WGBH National Center for Accessible Media, 2008. URL https://www.wgbh.org/foundation/ncam/guidelines/effective-practices-for-description-of-science-content-within-digital-talking-books.

[115] Erhardt Graeff. The Responsibility to Not Design and the Need for Citizen Professionalism. In *Tech Otherwise*, May 2020. doi:10.21428/93b2c832.c8387014. URL https://techotherwise.pubpub.org/pub/vizamy14.

[116] Nico Grant. Google Employee Who Played Key Role in Protest of Contract With Israel Quits. *The New York Times*, August 2022. ISSN 0362-4331. URL https://www.nytimes.com/2022/08/30/technology/google-employee-israel.html.

[117] Melissa Gregg. Inside the Data Spectacle. *Television & New Media*, 16(1):37–51, January 2015. ISSN 1527-4764. doi:10.1177/1527476414547774. URL https://doi.org/10.1177/1527476414547774. Publisher: SAGE Publications.

[118] James Grimmelmann. The Law and Ethics of Experiments on Social Media Users. *Colorado Technology Law Journal*, 13:219, 2015. doi:10.31228/osf.io/cdt7y. URL https://osf.io/cdt7y.

[119] Marilys Guillemin and Lynn Gillam. Ethics, Reflexivity, and "Ethically Important Moments" in Research. *Qualitative Inquiry*, 10(2):261–280, April 2004.

ISSN 1077-8004. doi:10.1177/1077800403262360. URL https://doi.org/10.1177/1077800403262360. Publisher: SAGE Publications Inc.

[120] Blake Hallinan, Jed R Brubaker, and Casey Fiesler. Unexpected expectations: Public reaction to the Facebook emotional contagion study. *New Media & Society*, 22(6): 1076–1094, June 2020. ISSN 1461-4448. doi:10.1177/1461444819876944. URL https://doi.org/10.1177/1461444819876944. Publisher: SAGE Publications.

[121] Aimi Hamraie. Designing Collective Access: A Feminist Disability Theory of Universal Design. *Disability Studies Quarterly*, 33(4), 2013. ISSN 2159-8371, 1041-5718. doi:10.18061/dsq.v33i4.3871. URL http://dsq-sds.org/article/view/3871.

[122] Donna Haraway. Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist studies*, 14(3):575–599, 1988. Publisher: JSTOR.

[123] Sandra Harding. *Sciences from Below: Feminisms, Postcolonialities, and Modernities*. Duke University Press, June 2008. ISBN 978-0-8223-4259-5. doi:10.1215/9780822381181. URL https://doi.org/10.1215/9780822381181.

[124] Amy Adele Hasinoff. *Sexting Panic*. University of Illinois Press, 2015. ISBN 978-0-252-03898-3. URL http://www.jstor.org/stable/10.5406/j.ctt13x1kx6.

[125] Lucia Hasty. Teaching Tactile Graphics, 2017. URL http://www.perkinselearning.org/videos/webcast/teaching-tactile-graphics.

[126] Lucia Hasty, Janet Milbury, Irene Miller, Allison O'Day, Pather Acquinas, and Diance Spence. Guidelines and Standards for Tactile Graphics. Technical report, Braille Authority of North America, 2011. URL http://www.brailleauthority.org/tg/.

[127] Ashley Hayward, Erynne Sjoblom, Stephanie Sinclair, and Jaime Cidro. A New Era of Indigenous Research: Community-based Indigenous Research Ethics Protocols in Canada. *Journal of Empirical Research on Human Research Ethics*, 16(4):403–417, October 2021. ISSN 1556-2646. doi:10.1177/15562646211023705. URL https://doi.org/10.1177/15562646211023705. Publisher: SAGE Publications Inc.

[128] Jeffrey Heer and Ben Shneiderman. Interactive Dynamics for Visual Analysis. *Communications of the ACM*, 55(4):45–54, 2012. ISSN 0001-0782, 1557-7317. doi:10.1145/2133806.2133821. URL https://dl.acm.org/doi/10.1145/2133806.2133821.

[129] Herbert A Simon. *The Sciences of the Artificial*. MIT Press, Cambridge, MA, USA, 1969. ISBN 978-0-262-19051-0.

[130] Ralph Hertwig and Andreas Ortmann. Deception in Experiments: Revisiting the Arguments in Its Defense. *Ethics & Behavior*, 18(1):59–92, March 2008. ISSN 1050-8422. doi:10.1080/10508420701712990. URL https://doi.org/10.1080/10508420701712990. Publisher: Routledge _eprint: https://doi.org/10.1080/10508420701712990.

[131] Tucker Higgins. Supreme Court Hands Victory To Blind Man Who Sued Domino's Over Site Accessibility. *CNBC*, 2019. URL https://www.cnbc.com/2019/10/07/dominos-supreme-court.html.

[132] Highcharts. Accessibility Module, 2021. URL https://highcharts.com/docs/accessibility/accessibility-module.

[133] Kashmir Hill. Wrongfully Accused by an Algorithm. *The New York Times*, June 2020. ISSN 0362-4331. URL https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html.

[134] Arne Hintz and Stefania Milan. At the margins of Internet governance: grassroots tech groups and communication policy. *International Journal of Media & Cultural Politics*, 5(1):23–38, March 2009. ISSN 1740-8296. doi:10.1386/macp.5.1-2.23_1. URL http://www.ingentaconnect.com/content/10.1386/macp.5.1-2.23_1.

[135] Mark Hochhauser. The informed consent form: document development and evaluation. *Drug information journal*, 34(4):1309–1317, 2000. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

[136] Anna Lauren Hoffmann. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, 22(7):900–915, June 2019. ISSN 1369-118X, 1468-4462. doi:10.1080/1369118X.2019.1573912. URL https://www.tandfonline.com/doi/full/10.1080/1369118X.2019.1573912.

[137] Anna Lauren Hoffmann. Even When You Are a Solution You Are a Problem: An Uncomfortable Reflection on Feminist Data Ethics. *Global Perspectives*, 2(1):21335, March 2021. ISSN 2575-7350. doi:10.1525/gp.2021.21335. URL https://online.ucpress.edu/gp/article/2/1/21335/116649/Even-When-You-Are-a-Solution-You-Are-a-Problem-An.

[138] Anna Lauren Hoffmann. Terms of inclusion: Data, discourse, violence. *New Media & Society*, 23(12):3539–3556, December 2021. ISSN 1461-4448, 1461-7315. doi:10.1177/1461444820958725. URL http://journals.sagepub.com/doi/10.1177/1461444820958725.

[139] Sean Hollister. Google contractors reportedly targeted homeless people for Pixel 4 facial recognition, October 2019. URL https://www.theverge.com/2019/10/2/20896181/google-contractor-reportedly-targeted-homeless-people-for-pixel-4-facial-recognition.

[140] Leona M Holloway, Cagatay Goncu, Alon Ilsar, Matthew Butler, and Kim Marriott. Infosonics: Accessible Infographics for People who are Blind using Sonification and Voice. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–13, New York, NY, USA, April 2022. Association for Computing Machinery. ISBN 978-1-4503-9157-3. doi:10.1145/3491102.3517465. URL https://doi.org/10.1145/3491102.3517465.

[141] Allison M. Horst, Alison Presmanes Hill, and Kristen B. Gorman. Palmerpenguins: Palmer Archipelago (Antarctica) Penguin Data, 2020. URL https://github.com/allisonhorst/palmerpenguins.

[142] Dorothy Howard and Lilly Irani. Ways of Knowing When Research Subjects Care. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 1–16, New York, NY, USA, May 2019. Association for Computing Machinery. ISBN 978-1-4503-5970-2. doi:10.1145/3290605.3300327. URL http://doi.org/10.1145/3290605.3300327.

[143] James M. Hudson and Amy Bruckman. "Go Away": Participant Objections to Being Studied and the Ethics of Chatroom Research. *The Information Society*, 20(2):127–139, April 2004. ISSN 0197-2243, 1087-6537. doi:10.1080/01972240490423030. URL http://www.tandfonline.com/doi/abs/10.1080/01972240490423030.

[144] Macartan Humphreys. Reflections on the Ethics of Social Experimentation. *Journal of Globalization and Development*, 6(1):87–112, June 2015. ISSN 2194-6353, 1948-1837. doi:10.1515/jgd-2014-0016. URL https://www.degruyter.com/view/journals/jgd/6/1/article-p87.xml. Publisher: De Gruyter Section: Journal of Globalization and Development.

[145] Amy Hurst and Jasmine Tobias. Empowering individuals with do-it-yourself assistive technology. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*, ASSETS '11, pages 11–18, New York, NY, USA, October 2011. Association for Computing Machinery. ISBN 978-1-4503-0920-2. doi:10.1145/2049536.2049541. URL https://doi.org/10.1145/2049536.2049541.

[146] Edwin L Hutchins, James D Hollan, and Donald A Norman. Direct manipulation interfaces. *Human-Computer Interaction*, 1:311–338, 1985.

[147] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, and Björn Eiderbäck. Technology Probes: Inspiring Design for and with Families. In *ACM Conference on Human Factors in Computing Systems (CHI)*, pages 17–24, New York, NY, USA, 2003. ISBN 978-1-58113-630-2. doi:10.1145/642611.642616. URL http://doi.acm.org/10.1145/642611.642616.

[148] Liesbeth Huybrechts, Henric Benesch, and Jon Geib. Institutioning: Participatory Design, Co-Design and the public realm. *CoDesign*, 13(3):148–159, July 2017. ISSN 1571-0882. doi:10.1080/15710882.2017.1355006. URL https://doi.org/10.1080/15710882.2017.1355006. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/15710882.2017.1355006.

[149] James Hyatt. The Birth of the Ethics Industry. *Business Ethics: The Magazine of Corporate Responsibility*, 19(2):20–27, May 2005. doi:10.5840/bemag200519229. URL https://www.pdcnet.org/pdc/bvdb.nsf/purchase?openform&fp=bemag&id=bemag_2005_0019_0002_0020_0027.

[150] Jane Im, Jill Dimond, Melody Berton, Una Lee, Katherine Mustelier, Mark S. Ackerman, and Eric Gilbert. Yes: Affirmative Consent as a Theoretical Framework for Understanding and Imagining Social Platforms. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pages 1–18, New York, NY, USA, May 2021. Association for Computing Machinery. ISBN 978-1-4503-8096-6. doi:10.1145/3411764.3445778. URL http://doi.org/10.1145/3411764.3445778.

[151] Lilly C. Irani and M. Six Silberman. Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 611–620, Paris France, April 2013. ACM. ISBN 978-1-4503-1899-0. doi:10.1145/2470654.2470742. URL https://dl.acm.org/doi/10.1145/2470654.2470742.

[152] Jack Henderson and Matt Prewitt. Data Coalitions and Escrow Agents. Technical report, RadicalxChange Foundation, June 2023. URL https://www.radicalxchange.org/media/documents/data-coalitions-and-escrow-agents.pdf.

[153] Liz Jackson. Disability dongle: A well-intended, elegant, yet useless solution to a problem we never knew we had. Disability dongles are most often conceived of and created in design schools and at IDEO., 2019. URL https://twitter.com/elizejackson/status/1110629818234818570.

[154] Marie-Andrée Jacob. Form-Made Persons: Consent Forms as Consent's Blind Spot. *PoLAR: Political and Legal Anthropology Review*, 30(2):249–268, 2007. Publisher: Wiley Online Library.

[155] Javier Creus, Andrea Barbiero, Joan Guanyabens, Gabriela Masfarré, Mara Balestrini, Georgina Campins, Valeria Righi, and Daniela Arguedas. Salus.Coop: Towards citizen governance and management of health data. Technical report, December 2016. URL https://static1.squarespace.com/static/5f232e29457c080a19e70bf5/t/5f8a1e40ce7cf718b1d0ebb0/1602887251669/SalusCoop_Informe.pdf.

[156] Jean Baker Miller. Women and Power. In Thomas Wartenberg, editor, *Rethinking Power*. SUNY Press, Albany, NY, 1992.

[157] Jonathan Zong and Beth Semel. Form, Content, Data, Bodies: Jonathan Zong and Beth Semel on Biometric Sans, February 2021. URL https://somatosphere.com/2021/form-content-data-bodies.html/.

[158] Shuli Jones, Isabella Pedraza Pineros, Daniel Hajas, Jonathan Zong, and Arvind Satyanarayan. "Customization is Key": Four Characteristics of Textual Affordances for Accessible Data Visualization, July 2023. URL http://arxiv.org/abs/2307.08773. arXiv:2307.08773 [cs].

[159] Crescentia Jung, Shubham Mehta, Atharva Kulkarni, Yuhang Zhao, and Yea-Seul Kim. Communicating Visualizations without Visuals: Investigation of Visualization Alternative Text for People with Visual Impairments. In *IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE VIS)*, volume PP,

2021. doi:10.1109/TVCG.2021.3114846. URL https://ieeexplore.ieee.org/abstract/document/9552938.

[160] Kade Crockford and Carl Takei. When We Fight, We Win: Victories in the Fight Against Face Surveillance Keep Piling Up, May 2021. URL https://www.aclu.org/news/privacy-technology/when-we-fight-we-win-victories-in-the-fight-against-face-surveillance-keep-piling-up/.

[161] Alison Kafer. *Feminist, Queer, Crip*. Indiana University Press, 2013. ISBN 978-0-253-00922-7. URL https://www.jstor.org/stable/j.ctt16gz79x.

[162] H.G. Kaper, E. Wiebel, and S. Tipei. Data Sonification And Sound Visualization. *Computing in Science Engineering*, 1(4):48–58, 1999. ISSN 1558-366X. doi:10.1109/5992.774840.

[163] Elena Karahanna and Detmar W. Straub. The Psychological Origins of Perceived Usefulness and Ease-of-use. *Information and Management*, 35(4):237–250, 1999. ISSN 0378-7206. doi:10.1016/S0378-7206(98)00096-2. URL https://doi.org/10.1016/S0378-7206(98)00096-2.

[164] Brian Keegan and J. Nathan Matias. Actually, It's About Ethics in Computational Social Science: A Multi-party Risk-Benefit Framework for Online Community Research. *AAAI Spring Symp. on Observational Studies through Social Media and Other Human-Generated Content*, 2016.

[165] Os Keyes, Nikki Stevens, and Jacqueline Wernimont. The Government Is Using the Most Vulnerable People to Test Facial Recognition Software. *Slate*, March 2019. ISSN 1091-2339. URL https://slate.com/technology/2019/03/facial-recognition-nist-verification-testing-data-sets-children-immigrants-consent.html.

[166] Hyeoneui Kim, Elizabeth Bell, Jihoon Kim, Amy Sitapati, Joe Ramsdell, Claudiu Farcas, Dexter Friedman, Stephanie Feudjio Feupe, and Lucila Ohno-Machado. iCONCUR: informed consent for clinical data and bio-sample use for research. *Journal of the American Medical Informatics Association*, 24(2):380–387, 2017. Publisher: Oxford University Press.

[167] N. W. Kim, S. C. Joyner, A. Riegelhuth, and Y. Kim. Accessible Visualization: Design Space, Opportunities, and Challenges. *Computer Graphics Forum (EuroVis)*, 40(3):173–188, 2021. ISSN 0167-7055, 1467-8659. doi:10.1111/cgf.14298. URL https://onlinelibrary.wiley.com/doi/10.1111/cgf.14298.

[168] Wolf Günther Koch. State of the Art of Tactile Maps for Visually Impaired People. In Manfred Buchroithner, editor, *True-3D in Cartography: Autostereoscopic and Solid Visualisation of Geodata*, Lecture Notes in Geoinformation and Cartography, pages 137–151. Springer, Berlin, Heidelberg, 2012. ISBN 978-3-642-12272-9. doi:10.1007/978-3-642-12272-9_9. URL https://doi.org/10.1007/978-3-642-12272-9_9.

[169] Barbara A. Koenig. Have We Asked Too Much of Consent? *The Hastings Center report*, 44(4):33–34, 2014. ISSN 0093-0334. doi:10.1002/hast.329. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4249719/.

[170] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790, June 2014. ISSN 0027-8424, 1091-6490. doi:10.1073/pnas.1320040111. URL https://www.pnas.org/content/111/24/8788. Publisher: National Academy of Sciences Section: Social Sciences.

[171] Robert Kraut, Judith Olson, Mahzarin Banaji, Amy Bruckman, Jeffrey Cohen, and Mick Couper. Psychological Research Online: Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American Psychologist*, 59 (2):105–117, 2004. ISSN 1935-990X(Electronic),0003-066X(Print). doi:10.1037/0003-066X.59.2.105. Place: US Publisher: American Psychological Association.

[172] Neha Kumar and Naveena Karusala. Braving Citational Justice in Human-Computer Interaction. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI EA '21, pages 1–9, New York, NY, USA, May 2021. Association for Computing Machinery. ISBN 978-1-4503-8095-9. doi:10.1145/3411763.3450389. URL https://dl.acm.org/doi/10.1145/3411763.3450389.

[173] Michael Kwet. Digital colonialism: US empire and the new imperialism in the Global South. *Race & Class*, 60(4):3–26, April 2019. ISSN 0306-3968. doi:10.1177/0306396818823172. URL https://doi.org/10.1177/0306396818823172. Publisher: SAGE Publications Ltd.

[174] Jingyi Li, Son Kim, Joshua A. Miele, Maneesh Agrawala, and Sean Follmer. Editing Spatial Layouts through Tactile Templates for People with Visual Impairments. In *ACM Conference on Human Factors in Computing Systems (CHI)*, pages 1–11, Glasgow, Scotland Uk, 2019. ISBN 978-1-4503-5970-2. doi:10.1145/3290605.3300436. URL http://dl.acm.org/citation.cfm?doid=3290605.3300436.

[175] Calvin A. Liang, Sean A. Munson, and Julie A. Kientz. Embracing Four Tensions in Human-Computer Interaction Research with Marginalized People. *ACM Transactions on Computer-Human Interaction*, 28(2):14:1–14:47, April 2021. ISSN 1073-0516. doi:10.1145/3443686. URL https://dl.acm.org/doi/10.1145/3443686.

[176] Liat Racin. Research in the Community Lightning Talk: The Community IRB Project, Liat Racin (The Engagement Lab at Emerson College), February 2016. URL https://vimeo.com/154358905.

[177] Zhicheng Liu, Biye Jiang, and Jeffrey Heer. imMens: Real-time Visual Querying of Big Data. *Computer Graphics Forum*, 32(3pt4):421–430, June 2013. ISSN 01677055. doi:10.1111/cgf.12129. URL https://onlinelibrary.wiley.com/doi/10.1111/cgf.12129.

[178] Alan Lundgard and Arvind Satyanarayan. Accessible Visualization via Natural Language Descriptions: A Four-Level Model of Semantic Content. In *IEEE Transactions on Visualization & Computer Graphics (Proc. IEEE VIS)*, page 11, 2021. doi:10.1109/TVCG.2021.3114770. URL http://vis.csail.mit.edu/pubs/vis-text-model.

[179] Alan Lundgard, Crystal Lee, and Arvind Satyanarayan. Sociotechnical Considerations for Accessible Visualization Design. In *2019 IEEE Visualization Conference (VIS)*, pages 16–20, October 2019. doi:10.1109/VISUAL.2019.8933762.

[180] Noni E MacDonald, Richard Stanwick, and Andrew Lynk. Canada's shameful history of nutrition research on residential school children: The need for strong medical ethics in Aboriginal health research. *Paediatrics & Child Health*, 19(2):64, February 2014. ISSN 1205-7088. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3941673/.

[181] Catriona Mackenzie and Natalie Stoljar. Introduction: Autonomy Refigured. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, pages 3–31. Oxford University Press, 2000. ISBN 978-0-19-512333-3.

[182] Marcel Mauss. The Gift: the form and reason of exchange in archaic societies. *L'Année Sociologique*, 1925.

[183] Marcus Abreu de Magalhaes. Data protection regulation: a comparative law approach. *International Journal of Digital Law*, 2(2):33–53, November 2021. doi:10.47975/IJDL.magalhaes.v.2.n.2. URL https://journal.nuped.com.br/index.php/revista/article/view/143.

[184] Rubén Alcaraz Martínez, Mireia Ribera Turró, and Toni Granollers Saltiveri. Accessible Statistical Charts For People With Low Vision And Colour Vision Deficiency. In *ACM International Conference on Human Computer Interaction*, Interacción '19, pages 1–2, New York, NY, USA, June 2019. ISBN 978-1-4503-7176-6. doi:10.1145/3335595.3335618. URL https://doi.org/10.1145/3335595.3335618.

[185] Alice E Marwick, Lindsay Blackwell, and Katherine Lo. Best Practices for Conducting Risky Research and Protecting Yourself from Online Harassment. Technical report, Data & Society, October 2016.

[186] Adrienne L. Massanari. Rethinking Research Ethics, Power, and the Risk of Visibility in the Era of the "Alt-Right" Gaze. *Social Media + Society*, 4(2):2056305118768302, April 2018. ISSN 2056-3051. doi:10.1177/2056305118768302. URL https://doi.org/10.1177/2056305118768302. Publisher: SAGE Publications Ltd.

[187] J. Nathan Matias. Going dark: Social factors in collective action against platform operators in the Reddit blackout. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 1138–1151, 2016.

[188] J. Nathan Matias and Merry Mou. CivilServant: Community-Led Experiments in Platform Governance. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pages 1–13, Montreal QC, Canada, 2018. ACM

Press. ISBN 978-1-4503-5620-6. doi:10.1145/3173574.3173583. URL http://dl.acm.org/citation.cfm?doid=3173574.3173583.

[189] J. Nathan Matias, Merry Ember Mou, Jonathon Penney, and Maximilian Klein. Do Automated Legal Threats Reduce Freedom of Expression Online? Preliminary Results from a Natural Experiment. September 2020. doi:DOI 10.17605/OSF.IO/NC7E2. URL https://osf.io/nc7e2/. Publisher: OSF.

[190] Doug McAdam. Recruitment to High-Risk Activism: The Case of Freedom Summer. *American Journal of Sociology*, 92(1):64–90, 1986. ISSN 0002-9602. URL https://www.jstor.org/stable/2779717. Publisher: University of Chicago Press.

[191] Carole McGranahan. Refusal and the Gift of Citizenship. *Cultural Anthropology*, 31(3):334–341, August 2016. ISSN 1548-1360. doi:10.14506/ca31.3.03. URL https://journal.culanth.org/index.php/ca/article/view/ca31.3.03. Number: 3.

[192] Carole McGranahan. Theorizing Refusal: An Introduction. *Cultural Anthropology*, 31(3):319–325, August 2016. ISSN 1548-1360. doi:10.14506/ca31.3.01. URL https://journal.culanth.org/index.php/ca/article/view/ca31.3.01. Number: 3.

[193] MDN Contributors. ARIA - Accessibility, November 2021. URL https://developer.mozilla.org/en-US/docs/Web/Accessibility/ARIA.

[194] Herman Melville. Bartleby, the Scrivener: A Story of Wall Street. *Putnam's Monthly. A Magazine of Literature, Science, and Art*, 1853.

[195] Amanda Morris. For Blind Internet Users, the Fix Can Be Worse Than the Flaws. *The New York Times*, July 2022. ISSN 0362-4331. URL https://www.nytimes.com/2022/07/13/technology/ai-web-accessibility.html.

[196] Meredith Ringel Morris, Jazette Johnson, Cynthia L. Bennett, and Edward Cutrell. Rich Representations of Visual Content for Screen Reader Users. In *ACM Conference on Human Factors in Computing Systems (CHI)*, pages 1–11, Montreal QC Canada, 2018. ISBN 978-1-4503-5620-6. doi:10.1145/3173574.3173633. URL https://dl.acm.org/doi/10.1145/3173574.3173633.

[197] Gavin Mueller. *Breaking things at work: The Luddites are right about why you hate your job.* Verso Books, 2021.

[198] Kevin Munger. Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment. *Political Behavior*, 39(3):629–649, September 2017. ISSN 0190-9320, 1573-6687. doi:10.1007/s11109-016-9373-5. URL http://link.springer.com/10.1007/s11109-016-9373-5.

[199] Tamara Munzner. A Nested Model for Visualization Design and Validation. *IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE VIS)*, 15(6):921–928, 2009. ISSN 1941-0506. doi:10.1109/TVCG.2009.111. URL https://ieeexplore.ieee.org/abstract/document/5290695. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[200] Tamara Munzner. *Visualization Analysis and Design.* A K Peters/CRC Press, New York, October 2014. ISBN 978-0-429-08890-2. doi:10.1201/b17511.

[201] Helen Nissenbaum. Privacy as Contextual Integrity. *Washington Law Review*, 79:39, 2004.

[202] Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, November 2009. ISBN 978-0-8047-7289-1.

[203] Safiya Umoja Noble. *Algorithms of oppression: How search engines reinforce racism*. New York University Press, New York, NY, US, 2018. ISBN 978-1-4798-3724-3 (Paperback); 978-1-4798-4994-9 (Hardcover). Pages: xv, 229.

[204] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 1–13, New York, NY, USA, April 2020. Association for Computing Machinery. ISBN 978-1-4503-6708-0. doi:10.1145/3313831.3376321. URL http://doi.org/10.1145/3313831.3376321.

[205] Olivia Solon. Facial recognition's 'dirty little secret': Social media photos used without consent, March 2019. URL https://www.nbcnews.com/tech/internet/facial-recognition-s-dirty-little-secret-millions-online-photos-scraped-n981921.

[206] Dan O'Sullivan and Tom Igoe. *Physical computing: sensing and controlling the physical world with computers*. Course Technology Press, Boston, MA, USA, 2004. ISBN 978-1-59200-346-4.

[207] Alex Pasternack. Here are the data brokers quietly buying and selling your personal information. *Fast Company*, March 2019. URL https://www.fastcompany.com/90310803/here-are-the-data-brokers-quietly-buying-and-selling-your-personal-information.

[208] Paul Ramsey, Margaret Farley, Albert R. Jonsen, and Marcia R. Wood. *The Patient as Person: Explorations in Medical Ethics*. Yale University Press, 2 edition, 2002.

[209] Kenny Peng, Arunesh Mathur, and Arvind Narayanan. Mitigating Dataset Harms Requires Stewardship: Lessons from 1000 Papers. *arXiv:2108.02922 [cs]*, November 2021. URL http://arxiv.org/abs/2108.02922. arXiv: 2108.02922.

[210] Gabriel Pereira. Towards Refusing as a Critical Technical Practice: Struggling With Hegemonic Computer Vision. *A Peer-Reviewed Journal About*, 10(1):30–43, August 2021. ISSN 2245-7755. doi:10.7146/aprja.v10i1.128185. URL https://aprja.net//article/view/128185. Number: 1.

[211] Jay Peters. IBM will no longer offer, develop, or research facial recognition technology, June 2020. URL https://www.theverge.com/2020/6/8/21284683/ibm-no-longer-general-purpose-facial-recognition-analysis-software.

[212] Peter Pirolli and Stuart Card. Information foraging. *Psychological Review*, 106 (4):643–675, 1999. ISSN 1939-1471(Electronic),0033-295X(Print). doi:10.1037/0033-295X.106.4.643. Place: US Publisher: American Psychological Association.

[213] Venkatesh Potluri, Tadashi E Grindeland, Jon E. Froehlich, and Jennifer Mankoff. Examining Visual Semantic Understanding in Blind and Low-Vision Technology Users. In *ACM Conference on Human Factors in Computing Systems (CHI)*, 2021. doi:10.1145/3411764.3445040. URL https://dl.acm.org/doi/abs/10.1145/3411764.3445040.

[214] Venkatesh Potluri, John Thompson, James Devine, Bongshin Lee, Nora Morsi, Peli De Halleux, Steve Hodges, and Jennifer Mankoff. PSST: Enabling Blind or Visually Impaired Developers to Author Sonifications of Streaming Sensor Data. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, UIST '22, pages 1–13, New York, NY, USA, October 2022. Association for Computing Machinery. ISBN 978-1-4503-9320-1. doi:10.1145/3526113.3545700. URL https://dl.acm.org/doi/10.1145/3526113.3545700.

[215] Puneet Chawla Sahota. Research Regulation in American Indian/Alaska Native Communities: Policy and Practice Considerations, 2009. URL http://www.ncai.org/policy-research-center/initiatives/Research_Regulation_in_AI_AN_Communities_-_Policy_and_Practice.pdf.

[216] Katyanna Quach. While Google agonizes over military AI, IBM is happy to pick up the slack, even for the Chinese military, 2019. URL https://www.theregister.co.uk/2019/04/12/ibm_ai_database/.

[217] Rayna Rapp. *Testing women, testing the fetus: The social impact of amniocentesis in America*. Routledge, 2004.

[218] John Rawls. *A Theory of Justice*. Harvard University Press, 1971. ISBN 978-0-674-00077-3. URL http://www.jstor.org/stable/j.ctvkjb25m.

[219] Donghao Ren, Bongshin Lee, Matthew Brehmer, and Nathalie Henry Riche. Reflecting on the Evaluation of Visualization Authoring Systems : Position Paper. In *2018 IEEE Evaluation and Beyond - Methodological Approaches for Visualization (BELIV)*, pages 86–92, October 2018. doi:10.1109/BELIV.2018.8634297.

[220] G. Robertson, R. Fernandez, D. Fisher, B. Lee, and J. Stasko. Effectiveness of Animation in Trend Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1325–1332, November 2008. ISSN 1077-2626. doi:10.1109/TVCG.2008.125. URL http://ieeexplore.ieee.org/document/4658146/.

[221] Hans Rosling. The best stats you've ever seen, 2006. URL https://www.ted.com/talks/hans_rosling_the_best_stats_you_ve_ever_seen. https://www.ted.com/talks/hans_rosling_the_best_stats_you_ve_ever_seen.

[222] David J. Rothman. *Strangers at the bedside: A history of how law and bioethics transformed medical decision making.* Routledge, 2017.

[223] Sally Wyatt. Non-users also matter: The construction of users and non-users of the Internet. In Nelly Oudshoorn and Trevor Pinch, editors, *How users matter: The co-construction of users and technology*, pages 67–79. MIT Press, Cambridge, MA, 2005. ISBN 978-0-262-65109-7.

[224] Salomé Viljoen. Data as Property?, October 2020. URL https://www.phenomenalworld.org/analysis/data-as-property/.

[225] Salomé Viljoen. Data Relations. *Logic Magazine*, (13), May 2021. URL https://logicmag.io/distribution/data-relations/.

[226] Sara Ahmed. No, June 2017. URL https://feministkilljoys.com/2017/06/30/no/.

[227] Sarah L. Fossheim. How (not) to make accessible data visualizations, illustrated by the US presidential election., 2020. URL https://fossheim.io/writing/posts/accessible-dataviz-us-elections/.

[228] Sarah Lucia Hoagland. *Lesbian Ethics: Toward a New Value, Palo Alto: Institute of Lesbian Studies.* Institute of Lesbian Studies, Palo Alto, 1988.

[229] Sasha Costanza-Chock. Mapping the Repertoire of Electronic Contention. In Andrew Opel and Donnalyn Pompper, editors, *Representing Resistance: Media, Civil Disobedience and the Global Justice Movement.* Greenwood, NJ, 2003.

[230] Sasha Costanza-Chock. *Design Justice: Community-Led Practices to Build the Worlds We Need.* MIT Press, March 2020. ISBN 978-0-262-04345-8. URL https://mitpress.mit.edu/books/design-justice.

[231] Christine Satchell and Paul Dourish. Beyond the user: use and non-use in HCI. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group on Design: Open 24/7 - OZCHI '09*, page 9, Melbourne, Australia, 2009. ACM Press. ISBN 978-1-60558-854-4. doi:10.1145/1738826.1738829. URL http://portal.acm.org/citation.cfm?doid=1738826.1738829.

[232] Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. Vega-Lite: A Grammar of Interactive Graphics. In *IEEE Transactions on Visualization & Computer Graphics (Proc. IEEE VIS)*, 2017. doi:10.1109/TVCG.2016.2599030. URL https://ieeexplore.ieee.org/abstract/document/7539624.

[233] Benjamin Saunders, Jenny Kitzinger, and Celia Kitzinger. Anonymising Interview Data: Challenges And Compromise In Practice. *Qualitative Research*, 15(5):616–632, 2015. ISSN 1468-7941. doi:10.1177/1468794114550439. URL https://doi.org/10.1177/1468794114550439.

[234] Lasse Scherffig. There is no interface (without a user). A cybernetic perspective on interaction. *Interface Critique*, (1):58–80, June 2018. ISSN 2570-0480. doi:10.11588/ic.2018.1.44739. URL https://journals.ub.uni-heidelberg.de/index.php/ic/article/view/44739. Number: 1.

[235] Freedom Scientific. JAWS Web Verbosity, 2021. URL https://www.freedomscientific.com/SurfsUp/Web_Verbosity.htm.

[236] Elaine Sedenberg and Anna Lauren Hoffmann. Recovering the History of Informed Consent for Data Science and Internet Industry Research Ethics. SSRN Scholarly Paper ID 2837585, Social Science Research Network, Rochester, NY, September 2016. URL https://papers.ssrn.com/abstract=2837585.

[237] Beth Semel. The Body Audible: From Vocal Biomarkers to a Phrenology of the Throat, September 2020. URL https://somatosphere.com/2020/the-body-audible.html/.

[238] Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye. Reflective Design. In *ACM Conference on Critical Computing: Between Sense and Sensibility*, CC '05, pages 49–58, New York, NY, USA, 2005. ISBN 978-1-59593-203-7. doi:10.1145/1094562.1094569. URL http://doi.acm.org/10.1145/1094562.1094569.

[239] Ather Sharif, Sanjana Shivani Chintalapati, Jacob O. Wobbrock, and Katharina Reinecke. Understanding Screen-Reader Users' Experiences with Online Data Visualizations. In *ACM Conference on Computers and Accessibility (SIGACCESS)*, ASSETS '21, pages 1–16, New York, NY, USA, 2021. ISBN 978-1-4503-8306-6. doi:10.1145/3441852.3471202. URL http://doi.org/10.1145/3441852.3471202.

[240] Ather Sharif, Olivia H. Wang, Alida T. Muongchan, Katharina Reinecke, and Jacob O. Wobbrock. VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–19, New York, NY, USA, April 2022. Association for Computing Machinery. ISBN 978-1-4503-9157-3. doi:10.1145/3491102.3517431. URL https://dl.acm.org/doi/10.1145/3491102.3517431.

[241] Linda Sheppard and Frances K. Aldrich. Tactile Graphics In School Education: Perspectives From Pupils. *British Journal of Visual Impairment*, 19(3):93–97, 2001. ISSN 0264-6196. doi:10.1177/026461960101900303. URL https://doi.org/10.1177/026461960101900303.

[242] Ashley Shew. Ableism, Technoableism, and Future AI. *IEEE Technology and Society Magazine*, 39(1):40–85, March 2020. ISSN 1937-416X. doi:10.1109/MTS.2020.2967492.

[243] Katie Shilton. Values and Ethics in Human-Computer Interaction. *Foundations and Trends® in Human–Computer Interaction*, 12(2):107–171, July 2018. ISSN 1551-3955, 1551-3963. doi:10.1561/1100000073. URL https://www.nowpublishers.com/article/Details/HCI-073. Publisher: Now Publishers, Inc.

[244] Ben Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In BENJAMIN B. Bederson and BEN Shneiderman, editors, *The Craft of Information Visualization*, Interactive Technologies, pages 364–371. Morgan Kaufmann, San Francisco, 2003. ISBN 978-1-55860-915-0. doi:10.1016/B978-155860915-0/50046-9. URL https://www.sciencedirect.com/science/article/pii/B9781558609150500469.

[245] Ben Shneiderman. Technology-Mediated Social Participation: The Next 25 Years of HCI Challenges. In Julie A. Jacko, editor, *Human-Computer Interaction. Design and Development Approaches*, pages 3–14, Berlin, Heidelberg, 2011. Springer. ISBN 978-3-642-21602-2. doi:10.1007/978-3-642-21602-2_1.

[246] Herbert Simon. *The Sciences of the Artificial*. MIT Press, Cambridge, MA, USA, 1969. ISBN 978-0-262-19051-0.

[247] Audra Simpson. On Ethnographic Refusal: Indigeneity, 'Voice' and Colonial Citizenship. *Junctures: The Journal for Thematic Dialogue*, (9), 2007. ISSN 1179-8912. URL https://junctures.org/index.php/junctures/article/view/66. Number: 9.

[248] Audra Simpson. *Mohawk Interruptus: Political Life Across the Borders of Settler States*. Duke University Press, Durham, NC, 2014. ISBN 978-0-8223-5655-4.

[249] Mario Luis Small. 'How many cases do I need?': On science and the logic of case selection in field-based research. *Ethnography*, 10(1):5–38, March 2009. ISSN 1466-1381. doi:10.1177/1466138108099586. URL https://doi.org/10.1177/1466138108099586. Publisher: SAGE Publications.

[250] C Matthew Snipp. What does data sovereignty imply: what does it look like? In Tahu Kukutai and John Taylor, editors, *Indigenous Data Sovereignty*. ANU Press, 1st edition, November 2016. ISBN 978-1-76046-030-3. doi:10.22459/CAEPR38.11.2016.03. URL http://press-files.anu.edu.au/downloads/press/n2140/pdf/ch03.pdf.

[251] Daniel J. Solove. Privacy self-management and the consent dilemma. *Harv. L. Rev.*, 126:1880, 2012.

[252] Daniel J. Solove. The Myth of the Privacy Paradox. SSRN Scholarly Paper ID 3536265, Social Science Research Network, Rochester, NY, February 2020. URL https://papers.ssrn.com/abstract=3536265.

[253] Jan-Philipp Stein and Markus Appel. How to deal with researcher harassment in the social sciences. *Nature Human Behaviour*, 5(2):178–180, February 2021. ISSN 2397-3374. doi:10.1038/s41562-020-01011-6. URL http://www.nature.com/articles/s41562-020-01011-6.

[254] Daniel Stoller. IBM Hit With Lawsuit Claiming Image Use for Facial Recognition. *Bloomberg Law*, January 2020. URL https://news.bloomberglaw.com/privacy-and-data-security/ibm-hit-with-lawsuit-claiming-image-use-for-facial-recognition.

[255] Chris Stolte, Diane Tang, and Pat Hanrahan. Polaris: A System for Query, Analysis, and Visualization of Multidimensional Relational Databases. *IEEE Transactions on Visualization and Computer Graphics*, 8(1), 2002.

[256] Stop LAPD Spying Coalition. Before the Bullet Hits the Body: Dismantling Predictive Policing in Los Angeles. Technical report, May 2018. URL https://stoplapdspying.org/wp-content/uploads/2018/05/Before-the-Bullet-Hits-the-Body-May-8-2018.pdf.

[257] Alan R. Tait, Terri Voepel-Lewis, Shobha Malviya, and Sandra J. Philipson. Improving the readability and processability of a pediatric informed consent document: effects on parents' understanding. *Archives of pediatrics & adolescent medicine*, 159(4):347–352, 2005. Publisher: American Medical Association.

[258] Kim TallBear. *Native American DNA: Tribal Belonging and the False Promise of Genetic Science*. University of Minnesota Press, 2013. ISBN 978-0-8166-6585-3. doi:10.5749/minnesota/9780816665853.001.0001. URL https://minnesota.universitypressscholarship.com/10.5749/minnesota/9780816665853.001.0001/upso-9780816665853.

[259] Marli Terblanche and Lesley Burgess. Examining the readability of patient-informed consent forms. *Open access journal of clinical trials*, 2:157–162, 2010. Publisher: Dove Press.

[260] The CITI Program. Social-Behavioral-Educational (SBE) Comprehensive, 2018. URL https://about.citiprogram.org/en/course/human-subjects-research-2/.

[261] Bart Thomee, David A. Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. YFCC100M: The New Data in Multimedia Research. *Communications of the ACM*, 59(2):64–73, January 2016. ISSN 00010782. doi:10.1145/2812802. URL http://arxiv.org/abs/1503.01817. arXiv: 1503.01817.

[262] John R Thompson, Jesse J Martinez, Alper Sarikaya, Edward Cutrell, and Bongshin Lee. Chart Reader: Accessible Visualization Experiences Designed with Screen Reader Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pages 1–18, New York, NY, USA, April 2023. Association for Computing Machinery. ISBN 978-1-4503-9421-5. doi:10.1145/3544548.3581186. URL https://dl.acm.org/doi/10.1145/3544548.3581186.

[263] Dari Trendafilov, Roderick Murray-Smith, and Daniel Polani. Empowerment as a metric for Optimization in HCI. In *CHI 2015 Workshop on Principles, Techniques and Perspectives on Optimization and HCI*, 2015.

[264] Krystal S. Tsosie. Models of Data Governance and Advancing Indigenous Genomic Data Sovereignty. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 3592–3592, 2020.

[265] Krystal S. Tsosie, Joseph M. Yracheta, and Donna Dickenson. Overvaluing individual consent ignores risks to tribal participants. *Nature Reviews Genetics*, 20(9):

497–498, September 2019. ISSN 1471-0064. doi:10.1038/s41576-019-0161-z. URL http://www.nature.com/articles/s41576-019-0161-z. Number: 9 Publisher: Nature Publishing Group.

[266] Krystal S. Tsosie, Joseph M. Yracheta, Jessica A. Kolopenuk, and Janis Geary. We have "gifted" enough: indigenous genomic data sovereignty in precision medicine. *The American Journal of Bioethics*, 21(4):72–75, 2021.

[267] Eve Tuck and K. Wayne Yang. Unbecoming Claims: Pedagogies of Refusal in Qualitative Research. *Qualitative Inquiry*, 20(6):811–818, 2014. doi:10.1177/1077800414530265. URL https://doi.org/10.1177/1077800414530265. _eprint: https://doi.org/10.1177/1077800414530265.

[268] Jakob von Uexküll. *A foray into the worlds of animals and humans: with a theory of meaning*. Posthumanities. University of Minnesota Press, Minneapolis, 2010. ISBN 978-0-8166-5900-5 0-8166-5900-1.

[269] Una Lee. Announcing the release of the Consentful Tech curriculum!, November 2021. URL https://www.consentfultech.io/announcing-the-release-of-the-consentful-tech-curriculum.

[270] United States, National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report : Ethical Principles and Guidelines for the Protection of Human Subjects of Research, 1979. URL https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html. Last Modified: 2018-01-15 00:00:00.

[271] US Department of Health and Human Services. Revised Common Rule, January 2017. URL https://www.hhs.gov/ohrp/regulations-and-policy/regulations/finalized-revisions-common-rule/index.html. Last Modified: 2017-01-19 00:00:00.

[272] Janet Vertesi. Internet Privacy and What Happens When You Try to Opt Out, May 2014. URL https://time.com/83200/privacy-internet-big-data-opt-out/.

[273] Nicholas Vincent, Hanlin Li, Nicole Tilly, Stevie Chancellor, and Brent Hecht. Data Leverage: A Framework for Empowering the Public in its Relationship with Technology Companies. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, pages 215–227, New York, NY, USA, March 2021. Association for Computing Machinery. ISBN 978-1-4503-8309-7. doi:10.1145/3442188.3445885. URL http://doi.org/10.1145/3442188.3445885.

[274] Jessica Vitak, Katie Shilton, and Zahra Ashktorab. Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 941–953, 2016.

[275] Jessica Vitak, Nicholas Proferes, Katie Shilton, and Zahra Ashktorab. Ethics Regulation in Social Computing Research: Examining the Role of Institutional Review Boards - Jessica Vitak, Nicholas Proferes, Katie Shilton, Zahra Ashktorab, 2017. *Journal of Empirical Research on Human Research Ethics*, August 2017. URL http://journals.sagepub.com/doi/10.1177/1556264617725200.

[276] Jennie Vo, Dylan H. Hewitt, and Yingchen He. Web Accessibility: A Revisit of U.S. State and Territory COVID-19 Websites After Two Years. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 67(1):1084–1089, September 2023. ISSN 1071-1813. doi:10.1177/21695067231192275. URL https://doi.org/10.1177/21695067231192275. Publisher: SAGE Publications Inc.

[277] W3C. WAI-ARIA Graphics Module, 2018. URL https://www.w3.org/TR/graphics-aria-1.0/.

[278] W3C. Web Accessibility Laws & Policies, 2018. URL https://www.w3.org/WAI/policies/.

[279] W3C. WAI Web Accessibility Tutorials: Complex Images, 2019. URL https://www.w3.org/WAI/tutorials/images/complex/.

[280] WebAIM. Screen Reader User Survey #9 Results, 2021. URL https://webaim.org/projects/screenreadersurvey9/.

[281] Markus Weninger, Gerald Ortner, Tobias Hahn, Olaf Druemmer, and Klaus Miesenberger. ASVG Accessible Scalable Vector Graphics: Intention Trees To Make Charts More Accessible And Usable. *Journal of Assistive Technologies*, 9:239–246, 2015. doi:10.1108/JAT-10-2015-0124.

[282] Cedric Deslandes Whitney, Teresa Naval, Elizabeth Quepons, Simrandeep Singh, Steven R Rick, and Lilly Irani. HCI Tactics for Politics from Below: Meeting the Challenges of Smart Cities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pages 1–15, New York, NY, USA, May 2021. Association for Computing Machinery. ISBN 978-1-4503-8096-6. doi:10.1145/3411764.3445314. URL https://dl.acm.org/doi/10.1145/3411764.3445314.

[283] Joseph W. Wiedel and Paul A. Groves. Tactual Mapping: Design, Reproduction, Reading and Interpretation. Technical Report D-2557-S 1969, Department of Health, Education, and Welfare, Washington, D.C., 1969. URL https://archive.org/details/tactualmappingde00jose_0.

[284] John Wilbanks. Design issues in e-consent. *The Journal of Law, Medicine & Ethics*, 46(1):110–118, 2018. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

[285] John T. Wilbanks. Electronic Informed Consent in Mobile Applications Research. *The Journal of Law, Medicine & Ethics*, 48(1_suppl):147–153, 2020. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

[286] Langdon Winner. Do artifacts have politics? *Daedalus*, 109(1):121–136, 1980. Publisher: JSTOR.

[287] Richmond Y. Wong. Tactics of Soft Resistance in User Experience Professionals' Values Work. *Proceedings of the ACM Human-Computer Interaction*, 5(CSCW2), October 2021. doi:10.1145/3479499. URL https://escholarship.org/uc/item/4dv5s8pm.

[288] David Wright. Redesigning informed consent tools for specific research. *Technical Communication Quarterly*, 21(2):145–167, 2012. Publisher: Taylor & Francis.

[289] Keke Wu, Emma Petersen, Tahmina Ahmad, David Burlinson, Shea Tanis, and Danielle Albers Szafir. Understanding Data Accessibility for People with Intellectual and Developmental Disabilities. In *ACM Conference on Human Factors in Computing Systems (CHI)*, 2021. doi:10.1145/3411764.3445743. URL https://dl.acm.org/doi/abs/10.1145/3411764.3445743.

[290] Yeshimabeit Milner. Abolish Big Data, July 2019. URL https://medium.com/@YESHICAN/abolish-big-data-ad0871579a41.

[291] Ji Soo Yi, Youn ah Kang, John Stasko, and J.A. Jacko. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE VIS)*, 13(6):1224–1231, 2007. ISSN 1941-0506. doi:10.1109/TVCG.2007.70515. URL https://ieeexplore.ieee.org/abstract/document/4376144. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[292] Anon Ymous, Katta Spiel, Os Keyes, Rua M. Williams, Judith Good, Eva Hornecker, and Cynthia L. Bennett. "I am just terrified of my future" Epistemic Violence in Disability Related Technology Research. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–16, Honolulu HI USA, April 2020. ACM. ISBN 978-1-4503-6819-3. doi:10.1145/3334480.3381828. URL https://dl.acm.org/doi/10.1145/3334480.3381828.

[293] Wai Yu, Ramesh Ramloll, and Stephen Brewster. Haptic Graphs For Blind Computer Users. *Lecture Notes in Computer Science*, 2058:41–51, 2001. doi:10.1007/3-540-44589-7_5. URL http://link.springer.com/10.1007/3-540-44589-7_5.

[294] J. J. Zhang. Research ethics and ethical research: some observations from the Global South. *Journal of Geography in Higher Education*, 41(1):147–154, January 2017. ISSN 0309-8265. doi:10.1080/03098265.2016.1241985. URL https://doi.org/10.1080/03098265.2016.1241985.

[295] Jonathan Zong. From individual consent to collective refusal: changing attitudes toward (mis)use of personal data. *XRDS: Crossroads, The ACM Magazine for Students*, 27(2):26–29, December 2020. ISSN 1528-4972. doi:10.1145/3433140. URL http://doi.org/10.1145/3433140.

[296] Jonathan Zong and J. Nathan Matias. Automated Debriefing: Interface for Large-Scale Research Ethics. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '18*, pages 21–24, Jersey City, NJ, USA, 2018. ACM Press. ISBN 978-1-4503-6018-0. doi:10.1145/3272973.3272992. URL http://dl.acm.org/citation.cfm?doid=3272973.3272992.

[297] Jonathan Zong and J. Nathan Matias. Bartleby: Procedural and Substantive Ethics in the Design of Research Ethics Systems. *Social Media + Society*, 8(1):205630512210770, January 2022. ISSN 2056-3051, 2056-3051. doi:10.1177/20563051221077021. URL http://journals.sagepub.com/doi/10.1177/20563051221077021.

[298] Jonathan Zong and J. Nathan Matias. Data Refusal from Below: A Framework for Understanding, Evaluating, and Envisioning Refusal as Design. *ACM Journal on Responsible Computing*, 1(1):10:1–10:23, March 2024. doi:10.1145/3630107. URL https://dl.acm.org/doi/10.1145/3630107.

[299] Jonathan Zong, Crystal Lee, Alan Lundgard, JiWoong Jang, Daniel Hajas, and Arvind Satyanarayan. Rich Screen Reader Experiences for Accessible Data Visualization. *Computer Graphics Forum*, 2022. ISSN 1467-8659. doi:10.1111/cgf.14519. URL https://vis.csail.mit.edu/pubs/rich-screen-reader-vis-experiences/.

[300] Jonathan Zong, Isabella Pedraza Pineros, Mengzhu Katie Chen, Daniel Hajas, and Arvind Satyanarayan. Umwelt: Accessible Structured Editing of Multimodal Data Representations, March 2024. URL http://arxiv.org/abs/2403.00106. Forthcoming in ACM CHI 2024.