

Understanding and Improving Human Data Relations

Alex Bowyer

Contents

Abstract	1
Dedication	2
Acknowledgements	2
Lists of Tables and Figures	4
Tables in Chapter 3	4
Tables in Chapter 4	4
Tables in Chapter 5	5
Tables in Chapter 7	5
Tables in Additional Reference Information	5
Figures in Chapter 1	5
Figures in Chapter 2	5
Figures in Chapter 3	6
Figures in Chapter 5	6
Figures in Chapter 7	7
Tables in Additional Reference Information	8
Bibliography	9

Abstract

Technologies including PCs, smartphones, and cloud computing have transformed the world: In our daily lives we interact with many businesses and public services who (to reduce costs) increasingly seek to rely on data collection and processing rather than face-to-face user interactions to inform their decisions. This creates an *imbalance of power* between those who hold data and the individuals about whom data is stored, who cannot easily see their personal data or how it is used. This *Digital Civics* PhD research explores, from a pragmatic, individualist, constructivist perspective, the topic of ***Human Data***

Relations. Through two qualitative case studies across public and private sectors, it answers the question, ‘*What relationship do people need with their personal data?*’. Case Study One focuses on *Early Help* social care: Through four workshops with supported families, social workers and staff, a deep understanding of the individual perspective on civic personal data use is established. *Shared data interaction* is explored as a means to shift the balance of power towards the individual while maintaining an effective care relationship. Case Study Two is a three-month study exploring 10 participants’ experience of using *GDPR data access rights* to view their own data, resulting in insights into individual needs and the challenges of data-centric service relationships, and recommendations for improvement of policies and practices. With reference to literature from the fields of *Personal Information Management*, *Human Data Interaction* and *MyData* personal data ecosystems, these case studies contribute to a unified understanding of *six core needs* that people have in Human Data Relations. In the final chapter, the thesis discusses the *practical pursuit* of these goals, drawing on first-hand knowledge acquired from expert participation in industrial research projects at BBC R&D and Hestia.ai/SITRA, *mapping out the landscape for future research and innovation*.

Dedication

For my children Rosie, Joey, and Zach; my nephew Elliott; and my nieces Amy and Lyla. My wish is that that you and your generation might soon experience a future where technology can truly help people and empower them to thrive, and where personal data drives human flourishing more than corporate profit. I hope that this research can in some small way contribute to a better future for you all.

Acknowledgements

No-one really knows how hard a PhD is until they are already well beyond the point of no return. It is demanding, challenging and often thankless and lonely work. You spend countless hours striving to find meaning among masses of data or iterating endlessly to concisely express complex, nebulous and elusive ideas. This endeavour is made even harder when you are a mature student with financial and parental responsibilities. At times, especially during the final unfunded writing-up period, the impacts upon my life and those around me have been huge and unreasonable. For this reason, the greatest thanks of all go to my wife Joni Bowyer, who has stood by me throughout, picking up the slack again and again where I could not. She has endured the impacts of money, time, uncertainty and divided attention that this unforgiving work has thrust upon our family. She has also provided practical help on countless occasions with everything from poster layout to time management to grammatical advice. I love you forever, Joni. Thank you.

The next person I want to thank is Jack Holt, who dedicated many weeks and

months of his life to collaborate with me to analyse of mountains of participant data from Case Study Two and to co-write the paper (Bowyer *et al.*, 2022) with me. I have absolutely no doubt that without his dedication, it would have been impossible to complete and publish the GDPR study in any reasonable timeframe.

I would like to thank, in reverse chronological order, my supervisors and all the other faculty and staff who have supported me on this six-year journey:

- Dave Kirk, for both detailed and high-level advice in bringing my thesis to conclusion during the final year;
- Jan Smeddinck, for his calm, pragmatic encouragement and thorough feedback and on drafts and plans through the latter half of my research;
- Rob Wilson, for always reminding me to stay grounded in the data, and for his sage advice in matters sociotechnical and philosophical;
- Josephine Go Jefferies, for much-needed scrutiny and challenges to my writing and many detailed chapter draft reviews;
- Patrick Olivier, Pete Wright and Dave Kirk, for their continuing commitment to help me find ways to make the PhD financially viable through the finding and accommodating of peripheral paid work;
- Kyle Montague, for his valuable input on study design in the early stages of my PhD, and for advocating to protect my independence and integrity as a researcher during problematic negotiations with a partner organisation;
- Phil Lord, for assistance and advocacy with those same issues at a crucial time;
- Madeline Balaam, for inspiration on participatory methods right at the start, and for helping me develop a paper writing style;
- Rachel Pattinson, for being the best CDT manager a postgraduate could hope for;
- Alex, Fion, Glau, James, Sara, Paul, Nicola and all the other admin staff, for countless random assists;
- Rob Comber, Simon Bowen, Matt Wood and all the other lecturers who taught me valuable Digital Civics and research skills during the MRes, that helped shape me into the researcher I am.

I have been fortunate to take this journey with dozens of other researchers. I would like to especially thank:

- Tom Maskell, for his help with wrestling with concepts around data access and involvement, and companionship on many bus commutes in the early years;
- Sunil Rodger, for valuable moral support, writing camaraderie, and practical advice for the last two years; and
- Stuart Wheeler, for myriad data discussions and tactical discussions that helped me get through a difficult period and often went beyond project business.

I would also like to thank, in no particular order:

- Rebecca Nicholson, Sean Peacock, Jen Manuel, Rosie Bellini, Megan Venn-Wycherley, Kieran Cutting, James Hodge, Hazel Dixon, Seb Prost, Sara Armouch and all the other Digital Civics PhDs whose company and mutual support I enjoyed during dozens of writing sessions and discussions;
- (Matt) Marshall, for sharing his powerful script framework for markdown-based thesis writing and document generation (Marshall, 2020), and supporting me in adapting it; this made thesis development so much easier.
- Louis Goffe, Debbie Smart, Kat Jackson, Liam Spencer, Ruth McGovern and Kyle Montague for giving their time to help run workshops with participants in Case Study One;
- Paul-Olivier Dehayé, Mike Martin, Soheil Human, Jasmine Cox, Peter Wells, Ian Forrester, Rhianne Jones, Tim Broom, Suzanne Clarke, Hannes Ricklefs, Max Leonard, Chris Gameson, Iain Henderson, Dalia Al-Shahrabi, Sarah Knowles, and Michael Jelly, all of whom I have had the pleasure to work with, ruminate with or learn from;
- Kellie Morrissey, Aare Puusaar, Andy Dow, Zander Wilson, Raghdah Zahran, Michael Jelly, Jay Rainey, David Williams, Ben Wright, Paul Whittles, Jon Bowyer, and all the other lovely people who have encouraged me and supported me along the way on this journey; and
- the research participants themselves, who shall remain nameless but without which this work could not exist.

Finally, I would like to thank my parents, Jim and Rosi Bowyer. You have always been there for me, whenever I needed you. Thank you for your unquestioning love and support in an ever-changing world.

Lists of Tables and Figures

Lists of Tables by Chapter

Tables in Chapter 3

Table 3.1 - Context One (Civic Data & Early Help): Participants involved in Research Activities leading into Case Study One.

Table 3.2 - Context Two (Digital Life): Participants Involved in Digital Life Research Activities Leading into Case Study Two.

Tables in Chapter 4

Table 4.1 - Example Categories of Family Civic Data

Table 4.2 - Case Study One Group Design Workshops

Table 4.3 - Theme 1 - Meaningful Data Interaction for Families: Subthemes & Participant Quotes

Table 4.4 - Theme 2 - Giving a Voice to the Family: Subthemes & Participant Quotes

Table 4.5 - Theme 3 - Earning Families' Trust Through Transparency: Subthemes & Participant Quotes

Tables in Chapter 5

Table 5.1 - Types of Data Holding Organisation Targeted for GDPR Requests by Study Participants

Table 5.2 - Types of Personal Data Potentially Accessible from Data Holders via GDPR Rights

Table 5.3 - Presence and Quality Assessments of GDPR Responses by Data Type (as Percentages)

Table 5.5 - Participants' Hopes, Imagined Data Uses and Goals for GDPR, as well as Resultant Outcomes

Table 5.6 - Theme 1 - Insufficient Transparency: Subthemes & Participant Quotes

Table 5.7 - Theme 2 - Confusing & Unuseable Data: Subthemes & Participant Quotes

Table 5.8 - Theme 3 - Fragile Relationships: Subthemes & Participant Quotes

Tables in Chapter 7

Table 7.1 - Eight Lenses on Personal Data

Tables in Additional Reference Information

Table ARI.1 - Best and Worst Data Holders for GDPR, according to Participants' Judgements

Lists of Figures by Chapter

Figures in Chapter 1

Figure 1.1 - Poster Presentation of Case Study One

Figure 1.2 - The Structure of This Thesis

Figures in Chapter 2

Figure 2.1 - The Wisdom Curve: Making Data into Meaningful Information

Figure 2.2 - Li *et al.*'s Stage-based Model of Personal Informatics Systems

Figures in Chapter 3

Figure 3.1 - My Action Research Approach

Figure 3.2 - “Family Facts” — What is Data?

Figure 3.3 - Walls of Data — Sensitising Participants to the World of Commercially-held Data and GDPR

Figure 3.4 - Sentence Ranking — Bringing Support Workers and Families to a Shared Problem Space

Figure 3.5 - Family Civic Data Cards — Things to Think With

Figure 3.6 - Personal Data Examples — Making Data Relatable

Figure 3.7 - Home Interviewing: Card Sorting with a Family in Their Living Room

Figure 3.8 - Ideation Decks — Combining Random Design Ingredients to Generate New Ideas

Figure 3.9 - Group Poster Design — A Participant-designed Poster to Advertise Features of Imagined Data Interface Products

Figure 3.10 - Storyboarding Cards — A Collaboratively-constructed Narrative Created through Discussion from a Palette of Possible Parent and Staff Actions

Figure 3.11 - Thematic Analysis of Qualitative Data using Quirkos for Case Study One

Figure 3.12 - Spreadsheet-based Quantitative Analysis of Interview Data for Case Study Two

Figure 3.13 - Pilot Study Recruitment Poster

Figure 3.14 - How the Case Studies and Peripheral Activities Contribute to This Thesis

Figures in Chapter 4

Figure 4.1 - Participants’ Shared Values Deduced from Sentence Rankings Data

Figure 4.2 - Current Model of Data Interaction, and Proposed Model of Shared Data Interaction

Figures in Chapter 5

Figure 5.1 - A Journey Map of Each Participant’s Study Progression

Figure 5.2 - An Example Life Sketch from Interview 1, with Data Handling Companies in Red, Data Types in Blue, and Feelings in Green

Figure 5.3 - Sankey Overview of Participants’ GDPR Requests

Figure 5.4 - Longitudinal Distribution of Net Changes in Participants' Perceived Power and Trust Scores

Figure 5.6 - Participants' Perceived Trust in Provider at Different Stages of the GDPR/Study Process

Figures in Chapter 7

Figure 7.1 - The Two Motivations for HDR: Controlling Your Personal Data Ecosystem and Utilising Your Information About Your Life

Figure 7.2 - Mapping the Six Wants into Objectives for the HDR Opportunity Landscape

Figure 7.3 - Obstacles and Resulting Insights in the HDR Opportunity Landscape

Figure 7.4 - Life Concept Modelling

Figure 7.5 - Mock-up of a Unified TV Viewing History Interface

Figure 7.6 - SubsCrab: An Example Application for Ecosystem Detection and Visualisation

Figure 7.7 - Some of the Many Aspects of Metadata that Might Exist About a Datapoint or Dataset

Figure 7.8 - The Panopticon Structure of the Illinois State Penitentiary

Figure 7.9 - Human Values, as Identified in BBC R&D Research Funded by Nesta

Figure 7.10 - A Contact-and-Calendar-centric PDS Approach

Figure 7.11 - The Scattered Data Relating to a Vacation

Figure 7.12 - Mock-up of a Unified Interface for a Vacation

Figure 7.13 - Annotating Data with Semantic Context

Figure 7.14 - Theory of Change [ToC]: The Four Dimensions of Change

Figure 7.15 - HDR Approach 1: Discovery-Driven Activism

Figure 7.16 - HDR Approach 2: Building the Human-centric Future

Figure 7.17 - Conceptual Model for a Personal Data Store System

Figure 7.18 - High Level Data Types

Figure 7.19 - Life Information Modelled as Happenings

Figure 7.20 - A Simple PDS Life Information Presentation Model

Figure 7.21 - Mock-up of Life Information Presented in a PDS Interface

Figure 7.22 - Life Partitioning Analogy using a Cluedo™ board

Figure 7.23 - Mock-up: Browsing by Areas of Life

Figure 7.24 - Identifying Entity Associations in Data

Figure 7.25 - Facebook's World2vec Model, Semantically Modelling Human Information from Social Media Posts on Facebook

Figure 7.26 - Identifying the Attributes of Data

Figure 7.27 - Determining the Nature of a Piece of Data

Figure 7.28 - Attributes of Data

Figure 7.29 - Actions One Might Perform on Life Information

Figure 7.30 - Questions One Might Ask of Life Information

Figure 7.31 - Example Taxonomies for Life Information Navigation

Figure 7.32 - HDR Approach 3: Defending User Autonomy and Hacking the Information Landscape

Figure 7.33 - The Modern 'Black Box' View of Technology

Figure 7.34 - HDR Approach 4: Winning Hearts and Minds: Teaching, Championing and Selling the Vision

Figure 7.35 - SILVER Health Data Viewing Interface

Figure 7.36 - Summary of Generalised Change Strategies for Pursuing Better HDR, Using the ToC Model

Tables in Additional Reference Information



Figure 1: Figure ARI.1: Extract of Sample Scenario Storyboarding Exercise walkthrough

Figure ARI.x - Screenshot from Quirkos During Coding Process

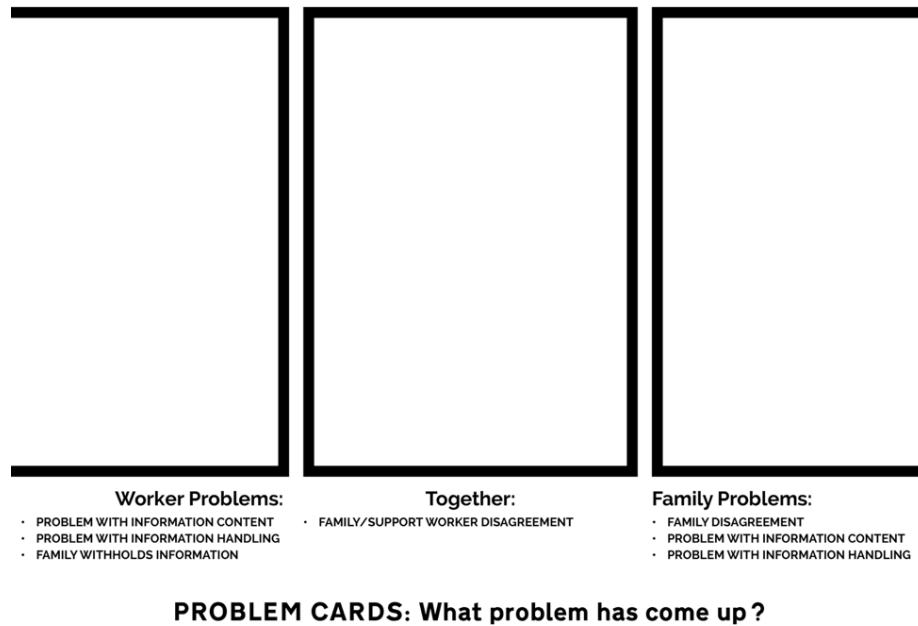


Figure 2: Figure ARI.2: Example Backing Mat for Storyboard Decks

Figure ARI.x - Screenshot from Quirkos at End of Coding Process

Figure ARI.c - Screenshot from Workflowy During Theme Construction

Figure ARI.x - Private Data Viewing Monitor with Viewing Glasses

Bibliography

Bowyer, A. *et al.* (2022) ‘Human-GDPR interaction : Practical experiences of accessing personal data’, *CHI '22*.

Marshall, M. (2020) ‘Markdown thesis’. Available at: <https://gitlab.com/mrshll1001/markdown-thesis>.