Understanding and Designing Human Data Relations

Alex Bowyer

5 Case Study Two: The Human Experience of GDPR

In this chapter, I will describe the second major case study of this PhD, in which I took 11[1] participants through an longitudinal in-depth one-on-one process of three interviews with coaching and support in between, with the total engagement per participant lasting approximately 4 hours over a three month period. The purpose of the research was gain a deeper understanding of people's attitudes to the kinds of personal data held by companies in people's everyday lives and what they want from that data (in pursuit of RQ1) and specifically to examine the human experience of existing in a data-centric world (see 2.1), with each individual having a number of relationships with service providers that involve the use and holding of personal data; in

line with RQ2 the goal is to better understand the role of that data in those relationships. In particular, having gained an initial understand of attitudes, hopes and expectations, a further objective was to examine how those expectations might change during the journey of digital life mapping, data request making, receiving and examining of data, and scrutiny of responses, collectively forming a holistic understanding of "the human experience of accessing your data with GDPR."

In section 5.1, I will expand on chapter 2 to explain the context of using GDPR in research as a means to retrieve personal data. In 5.2, I will explain the stages of the interview process (including details of how participants were sensitised) as well as the preparatory and intermediate steps I undertook as researcher. In section 5.3, I will explain the model of personal data types developed for this study, and will present quantitative and summary data from the interviews, explaining how participants' GDPR access requests progressed, highlighting participants' shared hopes and goals, and examining in particular how their perceptions of power and trust were affected by the experience. In section 5.4, I will describe the three themes uncovered through thematic analysis: that organisations provided participants with insufficient transparency to meet participants' hopes and their legal obligations (5.4.1), that people struggle to find meaning and value in their data when they do manage to access it (5.4.2), and that providers' data practices (in particular their GDPR request handling) can be harmful to their users' trust, but that greater openness can have an opposite, positive impact (5.4.3). I will discuss the implications of these findings with reference to prior literature, from the perspective of policymakers (5.5.1), data-holding companies (5.5.2), and individuals (5.5.3). Finally in 5.6, I will summarise these insights in terms of how they advance our understanding of the research questions and their wider significance.

5.1 Context: Accessing Your Personal Data Using Your GDPR Rights

5.1.1 The Current Need for Data Access

As established in 2.1.2 and 2.2.4, people live digital lives, inevitably involving the use of myriad digital services that collect personal data, which is subsequently mined for value and exploited at scale, creating an imbalance of power between data holders and data subjects, and a exclusionary landscape around data use which is difficult for individuals to navigate: having acquired data about individuals, this becomes a focus for service providers' decision-making and customer relations become less important. This everyday context is the chosen research setting for this case study.

Section 2.1.4 established how unaware many people are of this imbalance around data, that there is a want[2] for effective access to data to restore individual agency. As described in section 2.1.3, policymakers have been attempting since the 1970s to introduce legislation to tilt the balance of power back towards individuals, most recently and most notably the European Union's General Data Protection Regulation, which legally endows at least 513 million individuals[3] with new rights to timely data access, explanation, erasure and correction (Information Commissioner's Office, 2018).

Data protection and misuse issues have grown in the public awareness since the Snowden revelations in 2013 (Gellman, 2013), and have become even more important following the Cambridge Analytica scandal in 2018 ('Facebook–Cambridge Analytica Data Scandal', 2014; Chang, 2018), which may have resulted in manipulation of voting outcomes through personal data use, and the COVID-19 pandemic in 2021 (O'Donnell, 2020; Hamon *et al.*, 2021). Since the GDPR's launch in May 2018, it has undoubtedly resulted in new data access offerings; many large consumer companies have developed 'privacy hubs' or improved privacy policies where individuals can learn how their personal data is handled or access data download portals to easily download copies of it ('Privacy - Apple (UK)', no date; 'Privacy & Terms – Google', no date; 'Privacy', no date; 'Facebook - Data Policy', no date). Almost all data controllers and processors have now updated their privacy policies to include clear processes for data subjects to request copies of their personal data per their GDPR access rights.

However, it is not known how effective these offerings and processes are for service users, and how individuals feel about them in light of this backdrop of public concern. No service providers make data access statistics publicly available, but anecdotal reports from industry insiders suggest GDPR access rights and data download dashboards are not well-known and hardly used. This presents an opportunity to take individuals who have not previously used these capabilities on a journey of discovery that might enable us to assess the impact of these processes over time and whether–by compelling data holders to create such offerings and respond to access requests–GDPR succeeds in its goals to 'enhance the data protection rights of individuals' (Council of the European Union, 2015) and to give people 'control over their personal data' (The European Parliament and the Council of the European Union, 2016).

5.1.2 Current GDPR Research and its Limitations

Since it came into effect in May 2018, the GDPR has opened up new possibilities for research (Comandè and Schneider, 2021); the ability to obtain one's data records from organisations provides the general public with a potential deeper view inside those organisations, much like the UK's Freedom of Information Act has provided a view into governmental and public sector organisations, enabling research and improving accountability (Savage and Hyde, 2014). Such legally-enforced transparency can also provide researchers with a window into organisations and their processes that was previously only available based on goodwill. Ausloos and Veale (Ausloos, 2019; Ausloos and Veale, 2020) provide an outline approach for using the GDPR in research as well as describing the many ethical and metholodogical considerations that should be made. GDPR research can however be as simple as inviting participants to exercise their rights of access and talking to them about the experience and any changes in their perspective, which is the approach this study uses, as detailed below.

The GDPR process itself has also been examined from many perspectives by researchers: to understand data holder's compliance

with legislation (Ausloos and Dewitte, 2018; Arfelt, Basin and Debois, 2019); to evaluate data portability (Wong and Henderson, 2018) and 'privacy by design' (Waldman, 2020); to compare its effectiveness in public/private sector contexts (Quinn, 2021) or in improving explainability (Hamon *et al.*, 2021), fairness (Kasirzadeh and Clifford, 2021), consent (Human and Cech, 2021), transparency (Spagnuelo, Ferreira and Lenzini, 2019) and the reduction of data breach risks (Gonscherowski and Bieker, 2018). Potential negative impacts have also been considered; the GDPR could be seen as a threat to privacy (Bufalieri *et al.*, 2020) or as an impediment to health research (Clarke *et al.*, 2019).

Clearly GDPR has spurred a broad variety of research, spanning legal, social and technology domains. Yet, there is scant research into the individual human experience of the GDPR. Alizadeh et al. conducted a study with 13 users of a German loyalty programme and interviewed them before, during and after they made GDPR data requests (Alizadeh *et al.*, 2019), finding better responses and GDPR education were needed. This is a good example of the sort of work that is needed to explore the human perspective on the GDPR journey, though this particular study was limited in breadth (only one service provider was targeted) and in depth (the data returned from companies was discussed largely at a high level of 'were your expectations met?' and potential to use the data for one's own benefits was not examined). The implications of the experience upon the participants' relationship with the provider were also not explored, it seems that impacts of data handling practice upon relationships is an under-researched area in general. Recent work (Bufalieri *et al.*, 2020; Glavic *et al.*, 2021; Zuckerman, 2021) has established that openness and transparency around data handling are key to services establishing individuals' trust; indeed an echo of this was seen in a public sector context in Case Study One (see Chapter 4). In a commercial context, such changes in trust can impact customer satisfaction and business success.

At a more fundamental level, there is a need to understand the *experience* people have when using the GDPR; companies' GDPR processes have been designed to comply with litigation rather than by

focusing on individual needs or desires (Abowd and Mynatt, 2000; McCarthy and Wright, 2004; Wright and McCarthy, 2008) (for more details on experience-centred design refer to section 3.2). It is highly likely that some of these will have been overlooked. Such experiential understanding could inform the design of improvements to companies' GDPR mechanisms, as well as identifying specific needs that might be best met through improvements to policy, including to the GDPR itself.

5.1.3 Human-Data Interaction: Towards a Human-centric Personal Data Ecosystem

Given the fact that data-centric services now span all aspects of our lives, and the amount of personal data about individuals has grown, it has become critical to think about the way people interact with data as a 'whole life' problem. This is one of the reasons this study focuses on the layman rather than a particular demographic, and 'everyday services' rather than a particular domain. Data has transcended the machine and now encodes facts about our lives, it exists across devices and across providers (Weiser, 1991; Mydex CIC, 2010; Abowd, 2012). This means that personal information management has become a sociotechnical problem (see section 2.3.3), that can no longer be solved as a filing-and-retrieval problem as per traditional PIM approaches (see 2.2.2), but only when considered as multi-party negotiation over representation, ownership, access and consent. It is important to evaluate the GDPR in this context. Up to now, individuals have not had the means to participate in or initiate such negotiations. On paper, it would seem that GDPR rights do convey this capability, but it is not known whether in practice, service providers' responses to GDPR can actually deliver data subjects the ability to take part in negotiations around data in a fully-informed way. While some research on relationships around data and data as a shared resource is now emerging (see 2.2.5), the relationship with data-holding service providers has not been examined in this way.

A roadmap for best practice in this space can be found in the emergence of the 'personal data ecosystem' concept (see 2.3.4). Researchers have identified that a human-centric approach to personal

data is needed, placing individuals at the centre, as controllers and overseers of their own personal data (Mydex CIC, 2010; Symons *et al.*, 2017). This is an emergent space of much activity and research ('Human Data Interaction Project at the Data to AI Lab, MIT', 2015; 'HDI Network Plus, University of Glasgow', 2018; 'HDI Lab, Heerlen', 2020; BBC R&D, 2017; MyData, 2017; Symons *et al.*, 2017; MyData.org, 2018) and provides a strong framing for us to evaluate the human experience of–and interaction with–the GDPR; given people's diminished agency and control over their data (Woolgar, 2014; Crabtree and Mortier, 2016), do the GDPR's access rights, as implemented by service providers, provide the effective access (Gurstein, 2011) people need? Does the GDPR help people to achieve legibility, agency and negotiability, the three tenets of Human-Data Interaction (see section 2.3.2 and (Mortier *et al.*, 2014)).

This case study aims to explore the research gap identified in 5.1.2 above, from this perspective of greater human-centric need in a sociotechnical multi-party data use context. It will do so by scrutinizing the experience of using one's GDPR rights to discover how well the process meets individuals' needs and expectations; in the process the object is to uncover problems and identify possible solutions that could address them.

5.2 Study Design and Configuration

To address these research objectives, 31 qualitative interviews were conducted, with a convenience sample of 11 individuals from a population of researchers and students at (or connected with) Newcastle University, aged 20-40 years; self-identifying as 5 females and 6 males. Participants were not data experts (only 1 had previously made a GDPR request), but were computer-literate, educated to degree level, and used to reflecting critically on their own behaviours and opinions. Participants were compensated for their time with Amazon vouchers worth £20.

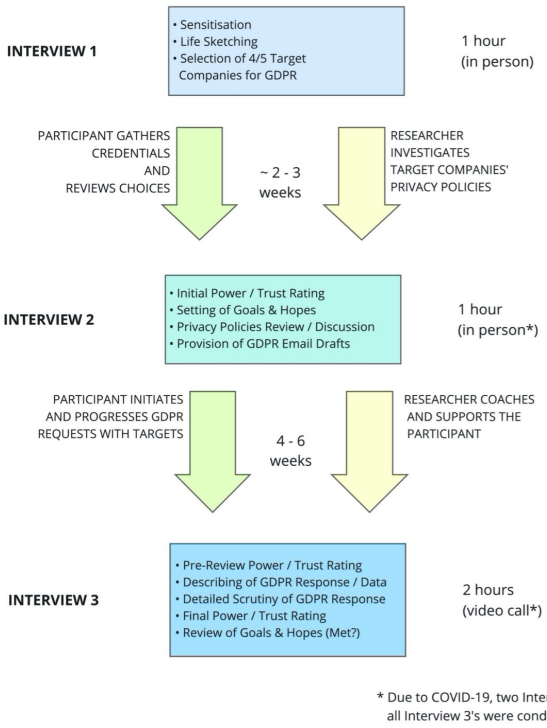

Figure 23: A Journey Map of Each Participant's Study Progression

Each participant's journey progressed at its own pace (see Figure 23) with participants invited to three separate 1-on-1 interviews between December 2019 and April 2020. The scope and purpose of each interview was as follows:

1. **Interview 1: Sensitisation, Life Exploration and Company Selection** [1 hour, in person]. Participants were sensitised to the research context using an interactive tour of a poster display on the topics of GDPR rights, potential data-holding organisations, potential types of data and potential uses for GDPR-obtained data. Baseline data was collected on participants' hopes and motivations, their current understanding of personal data, data access, data control, and power as it relates to data. Using a sketch interviewing (Hwang, 2021) technique, participants mapped out their 'data lives' (e.g. Figure 24), annotating key organisations that they have relationships with, types of data those companies might hold, and feelings about such data use and storage by each holder. Each participant selected 3-5 candidate companies to explore with GDPR requests.

2. **Interview 2: Privacy Policy Reviewing, Goal Setting and GDPR Request** Initiation [1 hour, in person]. To stimulate reflective thinking and measure impacts, participants were asked to discuss and score their initial feelings of trust and power with each company. Participants then viewed key sections of privacy policies on a screen with the researcher, to identify each company's statements on collection and use of personal data. Participants then initiated an email GDPR request for each company, which had been prepared using a tried-and-tested template generated by personaldata.io (Wiki.personaldata.io, no date). Interview 2 took place in person, except for P10 & P11 whose interviews took place over Zoom due to the COVID-19 pandemic.

3. **Interview 3: Detailed GDPR Response Review** [2 hours, online video call] Having allowed sufficient time for GDPR requests to conclude (there is a legal duty to reply within 30 days), a deep dive into the specifics of each GDPR experience took place. Participants' personal data was not collected by the research team, only described verbally; screen sharing was used to show excerpts to the researcher where the participant wished to do so. Participants were asked a structured set of questions about the completeness and value of any data returned, as well as new evaluations of trust and power, whether their hopes had been met, and any general feelings about the experience. Answers were recorded in a screen-shared spreadsheet, which was also used to structure the discussion (for a sample see Ap).

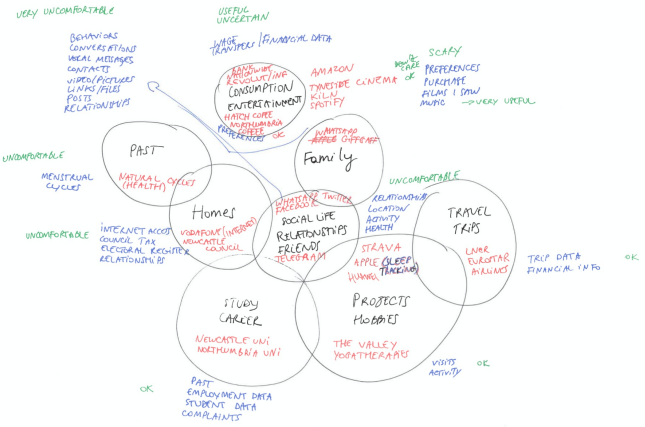

Figure 24: An Example Life Sketch from Interview 1, with Data

Interviews were audio and video recorded, then auto-transcribed using Google Recorder/Zoom, producing a 370,000-word corpus. Transcripts were split up by topic and analysed through reductive coding cycles to produce thematic findings (see 5.4). Quantitative data from interview spreadsheets was summarised and analysed (see 5.5). Sketches, recordings, screenshots and field notes aided interpretation of the transcripts.

5.3 GDPR Request Outcomes

5.3.1 Interview 1: GDPR Target Selection

Initially eight participants chose 5 target companies and three chose 4 to request data from. One participant (P9) withdrew from the study due to COVID-19 after Interview 1. Five participants withdrew a chosen company upon further consideration. Reasons for withdrawing chosen targets included having one's personal data mixed with other household members (Netflix), the account being in someone else's name (Morrisons), not wishing to impact active customer support matters (LNER), and inability to contact the provider by email (ifun.tv, see below). One participant selected Newcastle University, which was vetoed by the research team to avoid conflicts of interest. Hence, 41 out of a possible 52 GDPR subject access requests were made (to 28 distinct data holders) as shown in Table 8:

Table: Table 8. Types of Data Holding Organisation Targeted for GDPR Requests by Study Participants[a]

Type of Company	Company Names[a]
Major Internet Companies	Apple (3), Amazon, Facebook (4), Google (5)
Hardware Companies	Apple (3), Huawei, Google(5), Philips Hue (smart lightbulb manufacturer)
Online Platforms/Websites	Airbnb, Bumble (dating site), Check My File, Credit Karma, Direct Line, last.fm, LinkedIn
Social Networks & Dating	Facebook (4), Instagram, LinkedIn, Bumble (dating site)
Software/App Manufacturers	Freeprints, Niantic (creators of Pokémon Go), Natural Cycles (a menstrual tracker), Revolut, Spotify
Transport Companies	Tyne Tunnels, Nexus (Tyne & Wear Metro), LNER
Retailers & Loyalty Schemes	Amazon, Tesco, Sainsbury's, Nectar
Telcos	Virgin Media, Three
Sports Clubs	Sunderland AFC

[a] Where a company was chosen by more than one participant, the number of participants choosing that company is shown in brackets.

To ensure fairness and consistency, the aim was that all GDPR requests be sent by e-mail to the identified Data Protection Officer,

requesting both a subject access request (Information Commissioner's Office, 2021a) and a data portability request (Information Commissioner's Office, 2021b) be initiated, and specifically enumerating and asking for those datapoints that the company stated in its privacy policy, as well as those which the GDPR entitles individuals to obtain. To identify these datapoints, company privacy policies were analysed and the necessary information was compiled in personaldata.io's semantic wiki ('List of target companies for GDPR requests', no date) which has a feature to generate bespoke GDPR request emails, which we adapted and used (INSERT APPENDIX REF). Facebook, Apple, Huawei and Philips Hue do not offer a contact e-mail address, so the email text (shortened where length restrictions applied) was pasted into a contact form. In one case, entertainment website ifun.tv, the only available means of contact was via WeChat, resulting in the participant (a Chinese citizen) choosing not to contact ifun.tv due to fear of Chinese government surveillance. From our analysis of companies' privacy policies and with reference to GDPR rights, we constructed a taxonomy of the types of personal data that could be returned, using terms from those privacy policies and GDPR legislation: there are five types of personal data, as shown in Table 9.

Table: Table 9. Types of Personal Data Potentially Accessible from Data Holders via GDPR Rights

Type of Personal Data	Description	Examples
Volunteered Data	Data that the data subject has directly provided to the company through upload, contact or form completion.	Personally Identifiable Information (PII), contact details, user-generated content, photos, files, profiles, settings, communication history, financial information, security credentials, surveys/forms.
Observed Data	Data that has been indirectly or automatically collected about the data subject through product/service use or customer/staff interaction.	App usage information, behaviour on website, search/browse history, location tracking/tags, activity/health tracking, technical/device information, network/telco/ISP information, cookies & pixel trackers, staff observations, customer interaction notes.
Derived Data	Inferred data or profiles that have been created through algorithmic or human analysis of volunteered, observed or acquired data.	Interest profiles, advertising demographics, market segmentation, customer categorization, product/service recommendations, internal customer codes.
Acquired Data	Data that has been obtained or purchased from external sources such as civic records, reference agencies, advertisers or third parties.	Public records and information from internet searchs, reports or reviews from individuals, electoral roll data, credit checks, fraud checks, criminal record checks, e-mail/interest lists from advertisers, information shared between affiliates, sister companies or partner organisations.
Metadata	Information about how the other four categories of data have been handled, including storage, processing, uses, decision-making and external sharing.	Names of third parties data has been shared with, details of where data is stored and when/where it has exited the EU, explanations of how data has been used in automated or human decision making, legal bases for storage and processing.

5.3.2 Interview 2: Privacy Policy Review and Goal Setting

Participants reviewed and discussed privacy policies for their chosen target companies and were asked to define hopes and expectations for each GDPR request (see Table 12). These most commonly related to seeing the breadth and depth of data collection by companies, understanding what was being inferred and how personal data was used, and to use such information to better assess trustworthiness of those companies. Other motivators included the desire to reflect on one's own past data to gain self-insight, and to take control of or delete held data. Minor motivators included learning, creativity, fun, nostalgia, curiosity and the desire to shed light on specific incidents or answer specific questions.

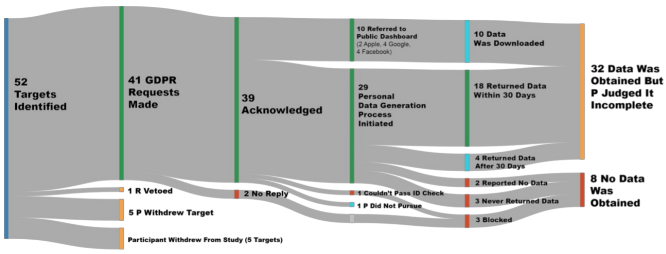

Figure 25: A Sankey diagram giving an overview of the GDPR requests undertaken by our participants (P)

At the conclusion of interview 2, participants were provided with the emails and instructions to start their GDPR requests, which progressed as illustrated in Figure 25. Eight requests resulted in no data being obtained, due to either data holder non-responsiveness, inability to access the right account or satisfy ID requirements, or confirmation being received that there was no data to supply. 32 requests (80%) resulted in at least some data being returned; 10 of these directed the participant to use a publically-available download dashboard such as Google Takeout, and the rest resulted in data being made individually available. Of these, one was mailed as printouts, another was mailed on CD-R, and the rest were delivered by e-mail (sometimes involving a secured online website to download). While 22 companies supplied bespoke data packages, 4 did not return it within the 30 days the legislation specifies (note: requests took place within the context of a global pandemic so response rates may not be typical). Following discussion, participants judged that all 32 requests receiving data had failed to return all requested data (across all five of the categories in

Table 9).

5.3.3 Interview 3: Reviewing the GDPR Response

Once each participant's GPDR requests had reached a conclusion point (as described above), they were invited to discuss the GDPR response in detail. Participants were asked to describe (and optionally show) the data they had received, then to evaluate the data holder's response for each data type, according to multiple metrics designed to assess the perceived quality of the GDPR request handling and the subjective value of any returned data. All questions were posed from the perspective of (a) the data that providers said they collect and process in their privacy policies, and (b) the rights that the GDPR specifies, to ensure discovery of missing data or unfulfilled rights would be considered objectively. Participant responses were considered quantitatively (see Table 10) and qualitatively (see section 5.4).

Table: Table 10. Presence and quality assessments of GDPR responses by data type (as percentages[a])

Type	Valued? [b]	Returned?	Complete?	Accurate?	Understandable?	Meaningful?	Usable?	Useful?
Derived	82%	39%	10% (dk:13%) [c]	20% (dk:20%) [c]	40% (p:40%) [d]	40%	0%	20%
Acquired	81	49	16 (dk:16)	50 (dk:25)	75 (p:0)	50	25	17
Metadata	73	4	0 (dk:7)	0 (dk:0)	0 (p:100)	0	0	0
Volunteered	57	53	55 (dk:0)	92 (dk:0)	72 (p:20)	72	52	58
Observed	48	33	18 (dk:12)	57 (dk:30)	61 (p:20)	57	52	61

[a] Percentages represent the proportion of "Yes" answers to each question, per data subtype, from all those *where a judgement was given*.
[b] Participants were asked whether this category of data from each provider would be valuable *if they were to receive it*.
[b] dk = don't know (percentage of cases where participants felt unable to assess data accuracy or completeness).
[c] p = partially (percentage of cases where data was judged *partially* understandable).

Table 10 shows quality assessments for each data type, with rows descending by subjective value. Notably, the kinds of data participants value most (derived, acquired and metadata) were less frequently returned, especially metadata (returned in 4% of cases). Where data was returned in these categories, it suffered from poor data quality, often judged as incomplete, inaccurate, unusable and not useful (although acquired data was largely understandable). At 53%, even the most returned category, volunteered data, was lacking. Where it was returned, accuracy (92%), meaningfulness (72%) and understandability (72%) were high. Observed data was least valued and also rarely returned or complete (yet judged to be of moderate

quality). Across all data types, data was only judged to be complete in 22% of cases, and in 62% of cases personal data specified in privacy policies to be collected was not returned, despite the legal obligation.

The above quality and coverage datapoints also allowed us to extract some information about which service providers were strongest or weakest in each category, and overall. This was done by tallying the "Yes" responses for each category and overall, then dividing by the number of times that provider was selected, to avoid inflating scores for popular companies. The outcome of this analysis is shown in Table 11. The companies that fared worst overall were those that did not return any data at all in response to a GDPR request (Sainsbury's, Freeprints, Tyne Tunnels, LinkedIn, Huawei, Bumble, LNER). As a caveat, it should be noted that Sainsbury's and Huawei *did* respond, claiming to hold no data for the requesting participant. The other named companies here did not respond at all, despite at least two follow-up emails being sent to them, and despite in some cases having initially acknowledged and promised to satisfy the request.

Companies producing responses with good coverage and good quality included Niantic, Nectar and Sunderland AFC as well as to a lesser extent Natural Cycles, Revolut, Spotify, Tesco and Amazon. Facebook and Google fared well for the breadth of data returned (due in part to their download dashboards), though the quality of Google's data was found lacking across multiple categories. Last.fm (owned by CBS) fared poorly overall due to poor category coverage, despite the data that it did return being of high quality.

Table: Table 11. Best and Worst Data Holders in Different Categories, According to Participants' Judgements[a]

Category / Metric	Best Companies	Worst Companies
Availablity of Data / Breadth of Data Returned	Nectar, Niantic, XYZ Sports Club, Natural Cycles, *Facebook, Google, Spotify, Revolut*	Sainsbury's, Freeprints, XYZ Tollways, LinkedIn, Huawei, Bumble, XYZ Trains, XYZ Transit, *Three, Philips Hue (Signify), Check My File*
Completeness of Returned Data	Niantic, *Nectar, XYZ Sports Club*	Sainsbury's, XYZ Tollways, Freeprints, XYZ Transit, LinkedIn, Huawei, Revolut, Bumble, LNER, last.fm (CBS), *Google, Tesco*
Accuracy of Returned Data	XYZ Sports Club, Niantic, *Tesco, Nectar, Amazon, Natural Cycles*	Direct Line, last.fm (CBS), *Google*
Understandability of Returned Data	Nectar, Spotify, XYZ Sports Club, Niantic, Apple, *last.fm (CBS)*	AirBNB, Virgin Media, *Google, Instagram, Tesco*
Meaningfulness of Returned Data	Niantic, *Spotify, XYZ Sports Club, Natural Cycles, last.fm (CBS)*	AirBNB, Credit Karma, Philips Hue (Signify), Direct Line
Usability of Returned Data	Amazon, last.fm (CBS), *Facebook*	AirBNB, Credit Karma, Virgin Media, XYZ Sports Club, Huawei, Three, *Google*
Usefulness of Returned Data	Amazon, Facebook, Virgin Media, Spotify, Revolut, Niantic, last.fm (CBS)	AirBNB, Credit Karma, Nectar, Direct Line, Three, *Google*
OVERALL[a]	**Niantic, *XYZ Sports Club, Facebook, Spotify***	**Sainsbury's, Freeprints, XYZ Tollways, LinkedIn, Huawei, Bumble, XYZ Trains, *last.fm (CBS), Philips Hue (Signify), XYZ Transit***

[a] Companies were ranked according to total number of all responses in that category for this company that were "Yes" .
[b] Company names in normal text are best/worst; names in italics are second best/second worst.

At the conclusion of the final interview, participants were reminded of the specific hopes and anticipated data uses they had expressed at the start of their journey and asked about how well each goal had been met. These answers were recorded and combined to produce percentage values showing in how many cases goals were fully met, partially met, or not met at all, as shown in Table 12.

Participants felt their goals were not fully met in 78% of cases, and 54% were not met at all. Specific shared problem areas included (1) the desire to understand what providers infer from held data (7 participants), which was unmet in 73% of cases and only fully met in 7% of cases; and (2) the desire to delete one's data, which was a stated goal in 10 cases but was only met in one of them. Four wholly unmet hopes were to investigate specific incidents (GDPR responses were often delivered as a one-off package without any kind of backchannel or opportunity to ask questions), to secure data, to check accuracy, and to move data to another service.

Table: Table 12. Participants' hopes, imagined data uses and goals for GDPR, as well as resultant outcomes

Hope or Goal	Distinct instances of this goal	Distinct participants	Specific companies in mind for this goal, if any	Was this hope met? Unmet?	Partially met?	Fully met?
GOALS RELATING TO ACCOUNTABILITY AND CONTROL:						
Understand the breadth and depth of what data is collected	24	7	Amazon, Apple, CheckMyFile, Credit Karma, Facebook, Google, LNER, Nectar, Philips Hue, Spotify, Tesco, Three, Virgin Media	42%	17%	42%
Understand what is inferred about you from your data	15	7	Amazon, Apple, Direct Line, Google, Instagram, last.fm, LNER, Spotify, Tesco, Three	73	20	7
Assess provider trustworthiness	12	6	Apple, Credit Karma, Direct Line, Facebook, Freeprints, Nectar, Niantic, Sunderland AFC, Tesco, Three	42	42	17
Remove your data & control/limit its use	10	3	Bumble, ifun.tv, Instagram	90	0	10
See inside 'black box' algorithms & processes	9	4	Amazon, Facebook, Google, Tesco	56	11	33
Understand how and why your data is used	6	5	Direct Line, Google	50	33	17
Investigate specific questions or incidents	4	4	AirBNB, Three, Credit Karma, Instagram	100	0	0
Learn about data use and how to be safer online; educate others	3	2		0	33	67
Secure data about you and identify risks and leaks	2	2	Apple, Facebook	100	0	0
Check accuracy of data about you	1	1	CheckMyFile	100	0	0
Move your data to another service	1	1		100	0	0
Test your data rights	1	1		0	100	0
GOALS RELATING TO USING DATA FOR PERSONAL BENEFIT:						
Reflect on past activities & gain insights	14	5	AirBNB, Apple, Google, last.fm, LNER, Tesco, Virgin Media	57%	36%	7%
Find patterns/habits & track goals	6	5	last.fm, Nectar, Spotify, Tesco	17	50	33
Combine data from many sources for deeper insights	3	2	Philips Hue, Google	33	67	0
Play with, create, hack & remix your data	3	3	Google	67	0	33
Nostalgia, fun & inspiration	3	3	Spotify, Niantic	33	33	33
Keep your own data archive	2	2	last.fm	0	50	50
OVERALL	**18 goal types**	**10 people**	–	**54%**	**24%**	**22%**

5.3.4 Perceived Power and Trust

The research team examined how participants' feelings towards the data holders changed throughout the process. Participants were asked to assess trust from 0 (total distrust) to 10 (total trust). They were asked to assess their perceived power on a scale of -5 (total provider power) through 0 (balanced power) to +5 (total individual power). Participants were invited to explain their reasoning for initial ratings and for any changes. By repeating the same question at different times, longitudinal comparisons could be made. Changes in attitude were observed; these changes are summarized in Figure 26. Many participants' attitudes did change as a result of the experience, for both perceived power (45% of cases) and trust (66% of cases). For those with changed attitudes, the change was often negative: in 63% of cases where participants perceived a change in individual power, that change was a loss in individual power, and in the majority (52%) of cases, participants felt more distrustful of GDPR targeted companies after completing the process (constituting 79% of cases where a change in trust was perceived). However, it is important to note that in some cases GDPR had a positive impact; in 17% of cases participants

felt their perceived power had increased, and in 14% of cases participants felt more trusting of providers after GDPR.

Figure 26: Distribution of Net Changes in Participant's Perceived Power and Trust Scores over the Study's Duration

Looking deeper into these datapoints, changes in attitude could be attributed both to the impact of reviewing the privacy policy as well as to the experience of the GDPR process and the discursive review of GDPR responses. Figures 27 and 28 show snapshots of power and trust ratings at different points in the process which illustrate these impacts. Looking to explain these changes qualitatively, it was found that privacy policies often contradicted participants' expectations, resulting in discomfort. In two cases (Philips Hue and last.fm) privacy policy review revealed that the service relationship was with a completely different company than the participant thought, which was disturbing to them. LinkedIn's privacy policy was noteworthy as being exceptionally clear, reassuring and trust-enhancing to the participant, largely due to its 'easy read' text sidebars but also good use of examples. Google's privacy hub with its video explainers was considered easy to understand but necessarily broad (given their breadth of services) and thus over-simplified, raising uncertainty about generalisations made, and in some cases increasing distrust.

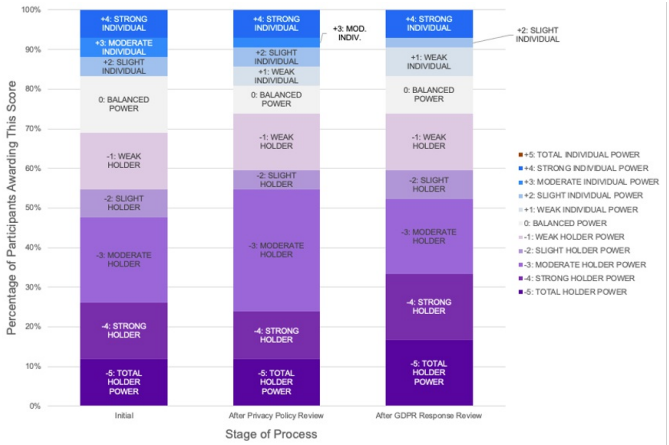

Figure 27: Perceived Power Balance Between Individual and Data Holder at Different Stages of the GDPR/Study Process

Considering the process as a whole, participants' attitudes were impacted particularly by the "hassle" (P11) they experienced in getting through the data access process, and from the realization that what seemed at first glance to be a thorough response, when examined more deeply in Interview 3 and viewed through the lens of the privacy policy promises and one's GDPR rights to the five categories of data, was in fact quite poor.

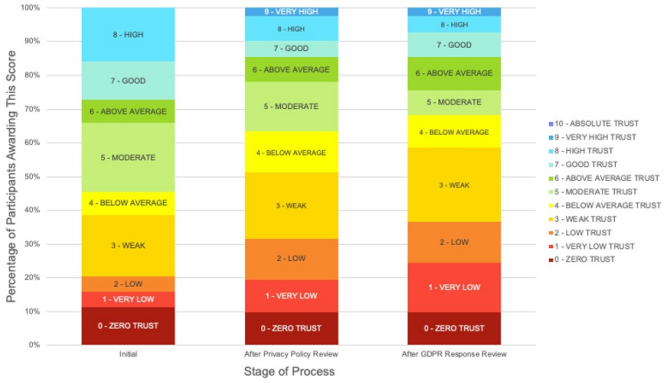

Figure 28: Participants' Perceived Trust in Provider at Different Stages of the GDPR/Study Process

5.4 Thematic Findings

[description like in c4 plus include some of following] Here I present outcomes from a deeper analysis of the participant experiences summarized above, derived through over 200 person-hours of iterative data analysis [76] of the interview transcripts. Our three key thematic findings are:

5.4.1 Themes & Subthemes

[convert this to prose]

1. **Insufficient Transparency** (Theme 1): Organisations appear evasive over data when responding to GDPR, leaving people "in the dark" even after making GDPR requests.
2. **Confusing Data** (Theme 2): When presented with their data, people struggle to understand it and relate it to their lives and are not able to make use of it.
3. **Fragile Relationships** (Theme 3): Companies' data practices, and in particular their privacy policies and GDPR response handling, can be impactful to customer relationships, carrying a risk of damaging trust but also the potential to improve relations. These themes are detailed in 5.1, 5.2 and 5.3 respectively.

[insert Table for each of the three themes]

Table 12. **Theme 1 - Insufficient Transparency**. Subthemes & Participant Quotes.

Subtheme	Description	Quote
Non-Compliance	Some providers failed to provide data on time or at all. Many participants found the data that was returned incomplete, and viewed this as non-compliance. Data holders' freedom to return incomplete data was seen as an expression of their power.	*"I am surprised at Google's unwillingness to provide me with all of the data … they haven't provided me with all of my data. And that's not legal."* [P7]

Subtheme	Description	Quote
A Desire for Understanding	Participants want to see, know and understand the data held about them. There was particular interest to see data collected or inferred about them without their involvement, and to understand how data is used and shared. In the face of providers making decisions based on data and processes that they could not observe, participants felt powerless.	*"[Companies have more power] because they're making decisions about things and you don't know how they're making those decisions."* [P5]
Inadequate Data Responses	Participants judged returned data to be lacking both in quantity, coverage and in the required context that would give it meaning. They felt that they did not receive good quality information, and that many questions remained unanswered. There was disappointment that, due to inadequate responses, GDPR did not confer any power to the individual, but recognition that it could still provide some insight into companies' approaches to data.	*"It's kind of disappointing because I would have hoped that this process would have levelled the user power versus the organisation power in a way that holds them accountable and [it doesn't] seem to be doing that."* [P1]

Table 13. **Theme 2 - Confusing Data**. Subthemes & Participant Quotes.

Subtheme	Description	Quote
Data Formats and Usability	Participants anticipated receiving data in formats they could explore, visualise, mashup and play with, but in fact often received data that lacked explanations. Data was often arranged in ways that were more reflective of internal systems than being optimised for use or understanding. In some cases	*"They did give me the data, but not how it fitted together. It's like being given the bricks to a house, and then they're like 'Here's your house'. It doesn't really mean anything when it's just bricks, if you don't know how to put it together."* [P5]
The Search for Meaning and Value in Data	Participants found the large volumes of data that were sometimes returned overwhelming, and wanted summaries and breakdowns to understand it, as well as tools to help them make sense of and explore the (often technically formatted) data. Data that spanned a period of time was judged particularly meaningful as it could serve as a window into past memories and would allow for trends and changes over time to be observed.	*"[It's] almost too much […] for a normal person to be able to process and understand […] It could do with a document detailing, like, 'this is what is in here'."* [P1]

Subtheme	Description	Quote
The Practicality of Using- or Deleting– Your Data	Participants wanted to use returned data to better understand themselves, but given that returned data lacked visualisations and interpretations, they were unable to practically use data in this way. There was also a strong desire to delete held data, or restrict its use, though participants did not see a clear path to achieving this.	*"[Companies did not] tell me what they are doing with [my data].. And sometimes I think my willingness to give a company data might be quite intrinsically linked with what they're gonna do with it."* [P7]

Table 14. **Theme 3 - Fragile Relationships**. Subthemes & Participant Quotes.

Subtheme	Description	Quote

Subtheme	Description	Quote
Power and Enforced Trust Through Data Holding	Participants feel that the sacrifice of (or the giving of permission to collect) personal data is a necessary cost in order to get the valued benefits of the services they want to use, something they are pressured to do and have no choice about. Such sacrifice is seen as the giving up of power, as participants lack access and control to that data. This amassing of data was sometimes seen as surveillance, and some saw great potential for misuse and abuse of it.	*"For me to have power over my data, I think is a fair and normal thing. But for a company to have power over [my] data means that it's basically a proxy to have power over me."* [P8]
Accountability and Perceptions of Data Holders	Participants entered the study with varying impressions of providers, and wanted to assess data practices in order to hold them to account. Participants' various observations reveal a strong link between their perceptions of providers' data handling practices and the trust they hold in those same providers.	*"When I like the company already, I'm more willing to give them my data."* [P2]

Subtheme	Description	Quote
Changed Perspectives Through Scrutiny	In general, the more that participants found out about data-centric practices through the process of scrutinizing privacy policies and making data access requests, the more they distrusted providers. Failure to explain or provide complete data was harmful to trust. Conversely, where providers were more transparent or participants did obtain interesting data insights, trust was increased.	*"If someone's not completely open with you, then you're like, 'What are you hiding?', which means you trust them less."* [P4]

5.4.2 Theme 1: Many Companies are Evasive and People are "Still in the Dark"

5.4.2.1 Non-Compliance

Responses to GDPR requests were broadly unsatisfactory (see section 4). Cases where data was not returned on time, or at all, were clear breaches of the legislation. Participants saw the incompleteness of returned data as a further failure to comply. At the beginning of the study, participants were reminded of their GDPR legal rights (most were already aware to some degree, as expected for such participants [90]), and several participants referred to these rights when reflecting on their experiences. e.g., P5 used legal rights to gauge the quality of one of her responses: "I feel more concerned now, […] what they've given me seemed reasonable. But then comparing against what we asked them for, what I'm legally [entitled to], it's a fraction." –P5 The extent to which participants expected their rights to be honoured

varied, with some participants expressing scepticism from the outset. To them, poor responses were consistent with their expectations; P6 reflected that his response from Facebook was “alarmingly unsurprising”. For others, failures to comply fully with the legislation did come as a surprise: “I am surprised at Google’s unwillingness to provide me with all of the data … they haven’t provided me with all of my data. And that’s not legal.”—P7 Failures to comply led to reflections that “there needs to be more enforcement” (P11), or that data holders are not under sufficient pressure to meet requirements. P6 amended his power evaluation of Facebook to reflect that he perceived them to have total power with respect to his data because the review of the data response had “made it clear which [data] they are prepared to share and which they aren’t”. Likewise, P11 characterised a selective response as an expression of power: “It seems like there’s a lot of derived data about things like purchases and stuff [that I would expect] that just isn’t there. So they’re free to not give me the data. That, to me, suggests they retain an awful lot of power.” —P11

5.4.2.2 A Desire for Understanding

As seen in Table 4, participants shared a common desire to see, know and understand the data that is held about them. At the most basic level, participants hoped to learn whether the data collected and held about them by the organisations exceeded what was required for the running of their services. For example, P11 sought to learn what data is being collected on him when he purchases train tickets: “I’d be interested to understand what data they have […] Is it just the patterns of my spending on trains, or is it a bunch of other stuff that they’re using for advertising to me?” - P11 In particular, participants were interested in data that they would not have been actively aware of. Participants felt that they were aware of the volunteered data (see Table 2) that they had deliberately provided; they sought awareness of data that had been collected or derived about them without their knowledge. “The bit that concerns me is where I don’t know what data is being taken by companies. If I’m registering for a library or something, I know [what] data I’m giving to them, but what I don’t know is all the other stuff that they’re recording” – P9 Participants sought to

understand the data in detail and whether it was used to infer further information or to affect decision-making of the data holder. P4 speculated whether data gathered by his smart home lighting might reveal the times at which he typically slept or watched television. P7 reported feeling “weird” about targeted advertisements relating to pregnancy that had most likely been delivered to her based on demographics. This lack of awareness translates to the data holder having greater power “because they’re making decisions about things and you don’t know how they’re making those decisions” (P5). Participants were also curious about the handling of their data and potential sharing between organisations. P4 chose to request data from a company that collated data from credit agencies in order to gain “a picture of what other companies can currently expose”. Others wanted to examine the nature of retention and deletion of data, such as P10, talking about dating app Bumble: “Since I found my partner […] I deleted my account and I’ve been wondering, ‘Are they still keeping my data at the back?’” –P10

5.4.2.3 Inadequate Data Responses

While the extent to which hopes were met varied across individuals and data holders (see Table 3), in many cases the desire for greater awareness and understanding was not satisfied. Volunteered data (e.g. basic personal information or user-generated content) was most often reported as complete. Participants already expected data holders to have this information and often found it mundane and uninteresting. P6 remarked that Facebook “give you that kind of descriptive boring data which is mainly all publicly available anyway” and that they had omitted “the stuff that I would consider valuable to them”. Frequently, participants commented that they still did not know what data was held, how it was used, or have answers to the questions that motivated their data requests. At the end of the study, when asked his feelings on his data being stored by organisations in general, P11 concluded that: “I still am quite concerned about how much data organisations have, particularly how they link that other data and how data is bought and sold, and I haven’t really got any answers on that.” – P11 In some cases, participants found that returned data was extensive, but

excluded context that would give it meaning. P5 received data from a car insurance company that utilises a mobile app for the purpose of generating a driving score and reflected: “I could see the data – it was the score that was weird for me. Like, it doesn’t tell you how it’s calculated.” In addition to the poor quality of data returned, participants often found the process of requesting and accessing the data to be unnecessarily difficult and time-consuming. Four participants (P4, P5, P7 & P11) independently described needing to jump “through hoops” in order to access their data. P10, on hitting a processual barrier, remarked that “I feel like they give you a response that [makes it so] you cannot proceed intentionally”. Participants identified that the painful and tedious processes they had experienced would be off-putting to many. P1 suggested that, without our automated generation of access request wording, it would be “a lot harder to get meaningful data out”. P7 also attributed successful requests to the guidance through the process, and that “even though I did the process correctly, I still didn’t get that much back”. Asked about what she found surprising about one of her requests, P5 responded “how difficult it was just to get my data, and the fact that I had to ask them about six different times”. However, not all of the data requests involved such painful processes: “Some companies make it dead easy to get, but then the data is not massively useful. Other companies make it easy to get, and it’s quite useful. Other companies make it a pain in the neck to get it.” - P11 Although positive experiences did occur, common issues with the process and the responses contributed to frustration and disappointment with the overall effectiveness of the data requests. P10 felt that “frankly, [GDPR] doesn’t have as much influence as I expected”. Similarly, P1 commented that: “It’s kind of disappointing because I would have hoped that this process would have levelled the user power versus the organisation power in a way that holds them accountable and [it doesn’t] seem to be doing that” – P1 Despite misgivings and feeling “in the dark” (P4), several participants found the process thought-provoking and report feeling more aware about their data sharing practices and settings. In some cases, this led to intentions to continue with further data requests, such as P4, who found that the process “got me thinking about, like what other things

could I try, and what other sources of personal data are there". P8 reflected that "it's a skill and a kind of knowledge that I think everyone should [have]. I don't think it's normal that I felt so clueless". Others commented more directly on the value of understanding GDPR itself, such as P7, who reported gaining "insights into how big companies are actually handling these requests". In the case of P2, this new perspective on GDPR is framed primarily as a lowered expectation of what it can achieve: "[I] think the exercise was useful in the sense that I understand what a GDPR request can do and what it cannot do. And there's a lot it cannot do. And I think it might seem that it gives you a lot of power, but really, it doesn't." – P2

5.4.3 Theme 2: People Struggle to Understand, Use and Relate to Their Data

5.4.3.1 Data Formats and Usability

Many of the other hopes, uses and plans identified by participants prior to the data requests related to the value that might be found within the data itself (see Table 5). Consistent with personal informatics literature [70], participants anticipated reflecting on and seeking insights in personal data, as well as putting data to practical uses such as budgeting, using it as an aid for remembering and archiving, or using the data for creative or fun purposes. In some cases, potential uses were difficult to predict, as there was some uncertainty about what data would be returned and what level of detail there would be. For example, P5 imagined creatively combining different datasets, but that this would depend on the data format returned. P4 was optimistic about building visualisations of Google data in this way, but uncertain on the detail: "I think … you could do some interesting mashups, but I don't really know what with until I've got the data. It depends on the data; I'm sure there could be some cool uses of it" – P4 Once data was received, participants struggled to interpret and understand it to a sufficient extent to be able to identify the useful data or meaningful information they had hoped for. Returned data formats and response structure were extremely varied, as was the degree to which the data was accompanied by explanations, keys or summaries. The

inconsistency in responses was noticed by P1, who expressed that “it would be nice if these companies had a standardised model of how this information is presented to people, so it can be easily understood.” However, different recipients with different goals and technical proficiency would need data in different structures and formats: “They have provided it in formats where I can see that, if I were a developer, I could do things with it, […] but if I was not that sort of person, it might be quite difficult to understand” – P7 P10, who received a response in JSON format, was critical of this, because “for normal people who don’t understand programming, I feel it’s just, there’s no use at all”. A further barrier was the inclusion of information that is meaningful to the data holder but requires further explanation, such as the inclusion of a screenshot of an internal system in one of P11’s responses that was “completely non-understandable”. He went on to question what it can and should mean for data to be usable, as “for me, as a technical person, having a CSV of data is quite useful, potentially, but actually what can I do with that if it’s Tesco’s internal systems data?”. Similarly, P4 reflected that some of the received data “could be valuable if you knew what the hell [was] in there”. Participants also commented on the absence of machine-readable data, as in the case of P4, whose response from his Internet Service Provider included a Microsoft Word document with pasted images from an internal portal. In evaluating its usability and usefulness, he identified that “it depends on what you want to get out of it, really. If you want to view the data they have about you, it’s quite usable. If you want to do something automated, then it’s not.” In some cases, data was considered to be usable, but the lack of explanation or guidance required in order to identify and access the value in their data was problematic. In the words of P5: “They did give me the data, but not how it fitted together. It’s like being given the bricks to a house, and then they’re like ‘Here’s your house’. It doesn’t really mean anything when it’s just bricks, if you don’t know how to put it together.” – P5

5.4.3.2 The Search for Meaning and Value in Data

Problems understanding and extracting value from data were exacerbated when large quantities of data were delivered. P1

compared the variation in quantities across responses as “either like death by thirst or death by drowning – in this case it would be better to drown, but still not ideal”. Large quantities of data were harder to digest when presented using obscure formatting and proprietary codes. For P4, when examining data from Philips Hue, there was “just so much of it that it’s impossible to know … you’d have to spend a few hours going through this and being like, ‘OK, what does that line mean, and that symbol, and that code?’”. Participants identified a need for summaries of their data and breakdowns of the data that had been returned. P1, for example, found that one data response was “almost too much […] for a normal person to be able to process and understand what’s in there. It could do with a document detailing, like, ‘this is what is in here’.” Some participants argued that the data received was often not understandable or useable without tools that are designed to simplify or visualise that data “for a non-techie person” (P11). For example, P7 found some of her JSON data more understandable once it had been interpreted via jsonlint.com, an online formatter suggested by the researcher. P2 reasoned that data holders are using certain tools to understand and analyse data “and they’re not just looking at a JSON file, so I would like to have the same visualisation [as them]”. Some participants did identify parts of their data that were meaningful, useful or otherwise significant. Data that spanned a period of time was considered particularly meaningful. Such historical data was particularly sought after as a means of remembering, with data potentially serving as a “window into your past” (P11). P5 saw value in perusing music-listening data “just because it’s cool to look back on stuff that you’ve done and you don’t necessarily distinctly remember it”. P6 reflected on some of his data as “a kind of personal history that has been quantified and sort of datafied”. For him, the value around such data is in the small details that form part of the context of certain life events: “I would like to […] build a picture, not just like, ‘I remember going to Reykjavik’, but if there’s other data around that time [I could] sort of paint a biography of myself” – P6 The length of time covered by a dataset also affords individuals with additional capabilities, such as the ability to capture trends and changes over time. The increased value of data covering a larger proportion of one’s lifetime was

recognised by P11 when selecting organisations to request data from: “I would actually be interested in last.fm, partly because the data goes back to 2008 … Spotify only goes back about four or five years and not everything I listen to is on Spotify” – P11

5.4.3.3 The Practicality of Using–or Deleting–Your Data

Several participants intended to use data collected over time to better understand themselves and their habits. For example, P2 hoped that his data response would contain, or allow the production of, breakdowns and charts that would help him to learn about his food shopping habits. P10 was able to access details of the accumulation of her spending on micro-transactions on the mobile game Pokémon Go that had not been available to her through the interface of the app itself. P11 had hoped to be able to use train-ticket purchase data to see what he could derive about his journeys in terms of geography, cost, journey length, etc. Participants also considered the increased value of historical data from the perspective of the data holder. For example, P11 saw mundane data collected over a long period as a liability: “10 years of worth of shopping records … how much would that be worth to a health insurance company, and would they succumb to the temptation to sell that on?”. P10, a Chinese citizen, preferred that data not be retained for long periods, as “in China, [there is a trend] that as soon as someone becomes famous, people begin digging [through] all their past experiences.” Preference for short data retention appeared as a recurring notion. With data holders collecting and retaining data for specified purposes relating to the delivery of services, participants questioned data keeping practices. For example P11 identified that “the thing that kind of concerns me about that is that I haven’t used Tesco online for years, like at least four or five years I think, so why are they hanging on to my IP address from five years ago?”. Most participants described the ability to delete or enforce the deletion of their data as having control over it. In some cases, participants indicated their existing intention to have data deleted, with GDPR expected to play a role in the enforcement or verification of that deletion – particularly in the case of P10, who wished to ensure the permanent deletion of her Bumble and Instagram accounts. In others,

deletion was seen as an important option, especially when the holding of sensitive data was considered a liability that was tolerated only in return for an actively-used service: “I now use a different one, but I used, for about a year, an app to track my menstrual cycle. [It was my] main contraception method, so that’s things that this company probably has. Now that I’m not using it any more, I don’t know if they delete the things or not” – P8 Participants foresaw potential uses of returned data to protect their data interests, including checking the accuracy, security and breadth of collected data to flag unforeseen concerns. Many participants hoped to make changes in data habits, privacy settings and choice of service provider following their anticipated learning from making a GDPR request, that might lead to an increased sense of individual safety and data control: “I want to understand how much they’re keeping. And what they’re doing with it. I’m hoping that by knowing that, I might change my behaviour about all the data I accidentally create.” – P7 However, without better data legibility and explanations, or clear pathways to deleting data, the ability to make such choices was hindered. E.g., P7 remarked: “I guess that’s one of … my criticisms of GDPR in general - that although I can understand what data a company holds about me, there’s no obligation for them to tell me what they are doing with it.. And sometimes I think my willingness to give a company data might be quite intrinsically linked with what they’re gonna do with it.” – P7

5.4.4 Theme 3: Poor GDPR Handling Can Damage the Fragile Trust Relationship

5.4.4.1 Power and Enforced Trust Through Data Holding

A significant proportion of our discussions explored participants’ views on their relationships with providers. Participants were found to be uncomfortable about data collection, especially because of a sense (see section 5.1) of being in the dark about current data practices: “I’m curious… I wonder what they’ve got on me. […] If it’s anything other than the barest minimum that is necessary for them to do their job […] then I get creeped out by that.” – P11 All 11 participants expressed the idea that the sacrifice of data is something that they have grown to

tolerate in exchange for some benefit. P6 tolerates data collection by travel agents because “they might help me pick a better deal next year.” P11 said he was happy for Tesco to collect data in order to “profile me to try to sell me more cheese, fine, whatever,” though expressed caution that he doesn’t “know what else they’re doing with it,” and more generally was “deeply concerned” about unseen data trading. The benefit can be convenience too; P10 had logged into Pokémon Go with her Facebook account [implying data collection by Facebook] “because it’s much easier”. This uneasy trade-off surfaced most often in the context of recommendations; generally, participants valued data-based suggestions provided they were “relevant” (P1, P8) and not too “intrusive” (P1, P6). P8 said that relevant music recommendations were “very useful” but found Amazon shopping recommendations “very scary” because “I don’t want to see that I’m predictable” and felt that “if someone out there knows [what I want] before you [it’s] like taking agency away from me.” Participants felt most uneasy about the amount of “intimate” (P1,P2) data that providers collect: “I don’t know how comfortable I am with Facebook having as much information as they do about my social circles.” – P1 P2 said he feels “quite vulnerable” that his Google search terms “say pretty much everything you have done… the most intimate things you were thinking about”. It is clear that data sacrifice is only tolerable up to certain limits: P10 said of Niantic that “as long as they don’t sell where I live or my daily routine, I’m fine” and this motivated her to want to check how her data was being used. However, she was critical of their use of in-game benefits as leverage for continued access to users’ location data, as “they pressure you into that [and] you don’t want to lose out”. Multiple participants spoke of such pressure to share data in order to access services or features, and the sense of having no choice. P11 singled out ISPs as having the potential to track everything their customers look at online, noting that “I don’t think you’ve got much choice about that.” While data holders typically acquire permission for their collection and processing of data, P2 felt that giving permission for data collection is “not granular enough”, and in P11’s view “it’s not a negotiation at all, it’s all or nothing.” Accordingly, participants reported feeling resigned about personal data collection: “I feel like it’s inevitable

that if you want to access their services at all, in any normal kind of way, that you automatically have to give them your data." – P7
Participants widely equated the holding of data as a source of power: P7 felt that to have control you need choice and said that "when I think about other people having my data […] the control isn't sitting with me." The ability of data holders to limit access to one's data is also viewed as an indicator of their power: P1 said that "If you're not getting what you perceive to be yours back in completion then you're not in control of your own data and you have fairly little power over it." When asked to define power in the context of data, P8's comments echoed prior research describing data as a proxy for direct involvement [18]: "For me to have power over my data, I think is a fair and normal thing. But for a company to have power over [my] data means that it's basically a proxy to have power over me." The notion of power through data was reflected in participants' evaluations of power balance: in 69% of relationships participants felt that the data holder had more power than them (rising to 74% after GDPR), whereas in only 17% of cases (unchanged by GDPR) did participants feel they had more power. Participants identified a range of concerns relating to an association of mass data collection with power. P1 noted that companies that "know a lot about everyone will inherently be able to have power either through persuasion or manipulation". In P6's view, Facebook used their knowledge of their users' friendships and relationships to "hook your attention" and prevent users deactivating accounts in a manner that was "disingenuous". Participants also felt that some data holders held so much data that it had begun to resemble surveillance, such as in the case of P1, who used "an absurd amount of [Google's] services" and reflected that "if I'm driving somewhere, I've got Google Maps open, so they know exactly where I'm going, they know how fast I'm going, they know what I'm listening to while I'm driving". Participants feared this kind of deep knowledge of individuals could be "used against" them (P2). P11 felt that Apple had enough data to "screw me over", and P5 considered her car insurer Direct Line to be able to use her data to "judge" her, and that "it's not like I can contest the data and say 'Actually, no, I disagree'." In a more extreme illustration, P10 shared her fears that data collected by

WeChat and Weibu (the Chinese equivalents of Facebook Messenger and Twitter) would be at risk of abuse by the Chinese government. Ultimately, participants felt that their data was “revealing” (P2,P3,P11) a lot of information about them, and so their only real option to maintain their privacy was to prevent data collection in the first place by not using that service at all (P1,P2,P3,P7,P10,P11).

5.4.4.2 Accountability and Perceptions of Data Holders

Participants said they sought knowledge on how their data is being collected, handled and used (as described in section 5.1) in order to better inform their choice of providers. In learning about how different companies operated with respect to their data, participants sought accountability and whether or not they should make more effort to better manage privacy settings or change behaviours in order to limit data collection. At the beginning of the study, participant opinions about data practices were often based on general factors such as reputation, size or business model. For example, participants who chose Apple as one of the organisations to request data from (P2, P10, P11) each reported firm pre-existing expectations. P2 described being “more at ease” with Apple, on the basis that their business model was focussed on hardware, than with Google, who were perceived as “making money through data”. He also noted that they were “positioning themselves as a defender of privacy rights”, a point echoed by P11, who was “curious to find out if their marketing claims match their reality around privacy” and wondered if the GDPR request might make him “reassess certain choices”. P10 reported an extreme decrease in trust, not due to the GDPR request, but to a documentary that she had watched between interviews that had caused her to become suspicious of their control over her hardware. Here, GDPR represented an opportunity to compare data expectations with reality. While Apple seemed mainly to benefit from existing trust in relation to their data practices, participants had concerns about other organisations. P6 found that Facebook had “in every shape or form, shown themselves not to be trusted”, an opinion that he supported by referencing “high profile news stories where they have done unscrupulous things and are very willing to just hand over data”.

Similarly, P9 reported feeling “slightly dubious” about Amazon due to things that she “had read in the press and about their ethics that may or may not be true, and just the size of them … and just the level of data, as well”. Participants were suspicious of businesses where there was a lack of clarity with respect to how they made money, while those companies that offered a paid service were considered more trustworthy. For example, P8 reported trusting Natural Cycles, stating that “one of the main things was there [are] no ads. It’s a paid service, so there’s no, like, ‘you don’t have to pay but we use your data … to make money’”. Participants also expected good data practices from companies that had made a positive impression on them in unrelated ways. P2 outlined that “when I like the company already, I’m more willing to give them my data”. This was often influenced by the perceived quality of services and software. E.g. P1 found that “in the same way that Amazon is quite janky, Google feels fairly polished and so I trust them more.” Google appeared to benefit from the same effect with other participants, who felt that trust was earned through the provision of valuable services, as with P4, who summarised that “the amount I trust [Google] is in line with the utility I get from them,”. P8, feeling comforted by Natural Cycles’ payment model, also felt encouraged by a sense of shared values: “This is woman-empowerment-orientated, so in that sense I think I do put my trust there as well.” –P8

5.4.4.3 Changed Perspectives Through Scrutiny

Participants’ evaluations of trust in data holders showed a tendency to diminish over the course of the data request process, with some distrust arising following reviews of privacy policy and some following reviews of the data responses (see Figures 5 and 6). In some cases, this could be attributed to an increased awareness of data collection practices, as with P5’s decrease of trust in Spotify after examining their privacy policy “because they shouldn’t need to know that much about me, they should just give me music”. However, it is notable that there does not appear to have been a corresponding reduction in perceived power: “They’ve not given me everything back that I thought they’d be collecting, which makes me trust them less. So power-wise, I don’t

think it's changed, but trust, I think it has." – P1 As identified by P1, the reasons associated with his lowering of trust scores related not to the data itself, but to the perceived partial or full non-compliance (see section 5.1). After receiving her data response from Spotify, P5 lowered her evaluation of trust further still "because they didn't say anything about what they're doing with my data or where it's going". P2 reduced his trust score for AirBNB "because of the way they've handled [the data request], and the way they've made it hard for me to read the data". Similarly, P7 downgraded her score for LinkedIn "because I feel like they have my data and [they've] not bothered to find my data, and that makes me feel like I shouldn't trust them quite as much". P8 lowered her trust score for Natural Cycles "because I think it's hard to get any sensitive data, and it's not really made clear what they're using it for". It appears that some companies have lost users' trust through a lack of transparency when responding to data requests. In the words of P4, "If someone's not completely open with you, then you're like, well 'What are you hiding?', which means you trust them less." "I think the lack of transparency in a lot of these processes has not helped, you know, if Tesco had […] plain English processes for getting the data and you've got the data in a plain English way, that would do a lot to bolster trust." —P11 Positive impacts of transparency for data holders occurred in those cases where trust remained the same or even slightly increased over the process of our study. For example, P5 reflected that her initial view on Instagram may have been "a little harsh" and that she "actually really liked what they sent … in comparison to the three others, I was genuinely, I opened Instagram's one and I was like 'this is really cool.'" P10 was very impressed with the response from Niantic and indicated that she trusted them very highly "because they replied really fast, the data provided is very detailed, and their attitude towards this whole issue is very positive," concluding that they are "a really nice company" and even indicating an increased willingness to spend money on their product. P6 trusted Sunderland AFC because "they were really kind of upfront and … I got the data from them first, […] no messing about, the format they gave me just made sense." In these comments, we can see an indication that, although the data requests often did not live up

to the hopes of the participants, positively engaging with the process was influential and did affect the outlook of our participants. In particular, close attention was paid to the willingness of companies to be transparent and forthcoming, with GDPR representing an opportunity to test organisations on their data practices and assess their integrity and competence as holders of their data.

5.5 Discussion

This study examined the GDPR's effectiveness in improving individuals' access and control over their personal data. The participants' experiences support the established power imbalance (see section 2.1) and suggest GDPR largely fails to empower individuals: both objectively (to the extent possible by this limited sample), in that most companies do not comply fully (either by returning insufficient and inadequate data, or by failing to return data on time or at all), and subjectively, in that returned data was often difficult to understand, impractical for use, and raised new questions and concerns. The findings also indicate that swift, transparent, and easy-to-use GDPR procedures can positively impact an individual's perception of an organisation. In light of these findings, this discussion offers insights on how the personal data landscape might be redesigned through policy (5.5.1) and business practice (5.5.2), and how individual action can have important impact too (5.5.3) – all in pursuit of the human-centric empowerment goals described in 5.1 as well as 2.2 and 2.3):

5.5.1 Implications for Policymakers: Compliance, Quality and Ongoing Access

Despite significant and obvious investment in dashboards, processes and bespoke data package production, the findings (while limited by the small number of participants) indicate that inadequate compliance with the GDPR is common. The findings are consistent with literature too: the participants' issues with completeness and compliance echo those first reported within the GDPR's first year [REF 9], suggesting completeness and compliance have not improved over this period.

However, the focus was on the effectiveness and experience of engaging with GDPR procedures from the individual's perspective. Participants' experiences were overwhelmingly of disappointment and frustration, with their hopes rarely met. They found that data holders often did not engage meaningfully with the process, and that the responses typically excluded or obscured data that could have provided them with the insights into their data privacy and the organisation's data practices that they sought. Evaluations of perceived power compared to data holders largely remained the same or worsened after accessing data through GDPR, and participants were not confident in the capabilities of the legislation to shift the balance of power. The process was perceived by some as a "box-ticking exercise" that was both frustrating and time-consuming and did not ultimately help them. Even though in 7% of cases participants did feel empowered by the GDPR, *all* participants receiving data were in practice left with additional time-consuming and sometimes technically-skilled work to take advantage of or interpret their returned data. This suggests that to improve the situation, policymakers need to make changes towards:

1) Better Compliance Through Enforcement of Complaints. At present, enforcement of the GDPR is uneven; each country has its own DPA (for example in the UK, this is the Information Commissioner's Office or ICO) and complaints are rarely pursued for individual cases. Instead, cases are processed by specific DPAs in a form similar to a class action lawsuit. This means that individuals have little impact when they do raise a complaint, and many GDPR complaints "become lost or resulted in lengthy delays" [REF 21], or may even be erroneously dropped [REF 72]. Until individuals have a clear and effective means to issue complaints [REF 11] that result in enforcement action (or a clear threat of it), it is likely that individuals will continue to have little recourse other than to repeat the request and hope similarly dissatisfied individuals will act on their behalf. Data holders must be held to account when they do not deliver the full set of data that they report possessing, or when they fail to do so within the legally obligated time frame.

2) Policies to Enforce Better Quality Responses. Many participants received data in frustrating formats, including screenshots, printouts or files that were too technical or littered with acronyms. Data was provided in formats too technical to understand, or not technical enough to be usable (see 5.4.3.1), showing a demand for both human-readable information summaries and machine-readable data files, where most providers typically provide only one or the other. Policymakers could provide suggested data formats or even propose new standards; this would help data portability, improve effectiveness [REF 44] and legibility [REF 80], can reduce costs through common tooling and catalyse the building of tools to interpret and understand data. Such standards are emerging [REF 78] as they are a technological necessity for data unification, but lack adoption.

3) Policies to Enforce Data Access as Ongoing Support, not One-Time Delivery. A radical redesign of policy is needed to give people the practical outcomes they desire and, according to the GDPR itself, deserve. Data access needs to be seen as more than "the delivery of data files". People need understanding of their data and of its handling, and this is the measure by which compliance should be assessed. The explanations GDPR mandates are not forthcoming; of the 119 hopes expressed by participants (see Table 5), 70 (59%) related to acquiring greater understanding of data practices. 38 (54%) of these were unmet, and a further 15 (21%) were only partially met. By mandating data holders to support individuals with not just the delivery of data, but assistance to understand that data, policies could become more impactful, not least because such understanding is critical to inform judgements around consent, loyalty and compliance.

5.5.2 Implications for Data Holders: Earn Trust by Opening Up Data and Enabling Users

While this study, and the GDPR itself, might seem adversarial to data holders given the goal to reduce their power by imposing new procedures, the findings emphasise the role of personal data in consumer relations. Data holders are likely aware of the paramount role of personal data in decision-making, but may not be aware of

individuals' perceptions about this. The findings suggest that failure to satisfy users who are concerned about the collection and usage of their personal data risks harms to consumer trust and confidence, at least for those users, and perhaps for others they might influence. In turn, however, this presents opportunities to use the mechanisms of the GDPR for customer loyalty and building better relations. In 52% of cases, following the process of examining privacy policies and engaging in GDPR data requests resulted in a decrease in reported trust in the data holder. While such impacts may for now be minimal, as only a small proportion of users read privacy policies [REF 92] and–one can assume–an even smaller number conduct GDPR requests, this is likely to change as issues around data privacy and trust continue to take centre stage in global geopolitics [REF 98,REF 107]. Furthermore, the growing number of businesses focused on "getting your data" or "taking control" [REF 25,REF 40,REF 97,REF 116–118] suggest demand for data access is growing. From the findings, there are three three positive takeaways for data holders:

1) Data transparency is an opportunity to increase customer loyalty and trust. GDPR's basic rights provide a starting point for delivering practical data transparency that will allow organisations to demonstrate that they are deserving of trust. By responding clearly and engaging openly and helpfully with GDPR data requests, organisations can demonstrate consistency between their privacy policy and their actions and demystify to their users the role that data holds in their business model. Research has shown that explanations can "ease humans' interactions with technology […], help individuals understand a system's function, justify system results, and increase their trust" [REF 42]. In 14% of cases, our participants felt more trusting of the service brand as a result of their GDPR experience (sometimes even displacing prior apprehensiveness or distrust), citing reasons such as speedy, hassle-free responses, clear and understandable data, providers being upfront and open with data, and staff who exhibited a positive attitude to the request.

2) Data transparency is an opportunity for improved and re-imagined customer relations around data. Beyond the opportunity

to improve trust, the mechanisms of data transparency suggested by the GDPR provide individuals with new capabilities for data curation and involvement. By offering individuals the ability to engage in empowering data interactions, data holders have the opportunity to improve engagement with their organisation and their services. If organisations view personal data as a shared resource to be curated and co-owned by the individuals that contributed it, there may be correspondingly shared benefits: for the individual, a sense of agency, influence and negotiability [80]; and for the service provider, an incentive for individuals to generate and share more data, an increased likelihood of individuals correcting inaccurate data, and more reliable and human-centric forms of ongoing consent closer to dynamic consent [61] than today's ineffective models of informed consent [73].

3) New customer demands indicate untapped business opportunities. As the 500-member-strong MyData Global organization [REF 82] shows, there is growing demand for personal data empowerment. People's personal data is splintered and trapped [REF 1,REF 16], and they cannot correlate data from different sources in order to reflect upon it, gain insights, and set goals [REF 70]. Due to commercial motivations, service providers generally deliver capabilities within a closed silo, not at the level of one's wider environment [REF 2]. To be better empowered the individual could be the point of integration, the centre of their own Personal Data Ecosystem (PDE) [REF 81]. Life-level capabilities [REF 17] and the opportunities that well-designed and well-regulated GDPR-type regulations promise in this regard have not yet been exploited. Thorough, complete and timely data access in standard formats, as mentioned above, will be critical to enabling this vision. Growing companies such as CitizenMe [REF 119], Digi.Me [REF 38], Mydex [REF 125], ethi [REF 58], HestiaLabs [REF 34], udaptor [REF 97] and exist.io [REF 120] as well as larger organisations like BBC R&D [REF 13] and Microsoft [REF 75] are already starting to innovate in this space.

5.5.3 Implications for Individuals: Becoming Aware of the Value and Power of Data, and Demanding More

While participants experienced disappointment and frustration in their GDPR journeys, all participants gained new understandings; if not always of their data itself, at least of their target companies' approach to data access requests. This new knowledge was sufficient to re-affirm or challenge existing attitudes or inform judgements–P1, for example, left Facebook after the study. Even an attempt to access data can be educational, and even a cursory look at a provider's 'What data do we collect' privacy policy section can provide pause for thought. Today, individuals remain largely in the dark about the collection, use and sharing of their data through a combination of perceived complexity and effort combined with a lack of clear benefits. Table 12, alongside the increased control and insight promised by the PDE movement and platforms linked in 5.5.1 and 5.5.2 above, provide a glimpse of what the future may hold: a world where individuals take more control of their data and gain actionable self-insights. Three key messages for individuals can be inferred:

1) Your data is used to represent you and define your user experience. We hand over our data in exchange for access to services, but providers then use it (usually in aggregate) e.g. to inform product design or decide what content you see. This 'innocent' handover of data is in fact giving providers the means by which we are treated and – at times – controlled. Recognizing the crucial role of data (and our limited influence over it) is the first step to pursuing greater agency and control.

2) Your data contains meaningful and valuable data about your life. Data, as our participants found, is dry and technical, but they all sought meaning and value within it (see 5.2.2). Within data lies potentially rich information about one's life and past activity – some of which can even be inaccessible through any other means. This highlights both a risk (that others might gain this insight) and a potential benefit (that we could access this insight ourselves). In this context, data deletion without keeping a copy may be inadvisable. To access the value in data, individuals will need to demand data standards, better access and control mechanisms and insight tools.

3) Self-education and awareness enable accountability and informed choices. The findings highlight a lack of knowledge. Transparency is critical to judging ‘to what extent the bargain is fair’ [REF 66]. It is not always delivered, but GDPR makes it your right; a right that cannot be fully refused. Through challenging poor GDPR responses and demanding better information, individuals can have impact. Providers are ultimately motivated by public demand—one of the reasons download dashboards exist. Through the public pressure of negative attention, companies can be motivated to improve data access [REF 33]. With patience, GDPR rights can be exploited to force small changes.

5.6 Summation

Through a longitudinal study of 10 participants lasting three months, this case study has qualitatively, and to a lesser extent quantitatively, evaluated the human experience of using one’s GDPR access rights and of living with data-centric service provider relationships.

The findings, while not statistically representative, suggest that people currently lack awareness of held data and its uses by service providers. By guiding participants on a journey of discovery and careful scrutiny, encouraging them to draw their own conclusions about service providers on the basis of companies’ own promises, individuals’ legal rights, and our participants own hopes (see Table 12), we have shown that such a journey can be educational and enlightening with regard to increasing awareness, but also can seriously damage brand loyalty and trust in providers if comprehensive and well-explained data is not returned in a supportive and open manner (see 5.4.4).

The experience of GDPR seems to be an unsatisfactory one for individuals; participants were generally still ‘in the dark’. We have highlighted serious problems with compliance (see 5.4.2.1): Participants received data that was incomplete, impractical for use, and they failed to acquire desired explanations. By its own aim to enhance individuals’ rights and control, the GDPR does not succeed. Participants continued to feel a lack of agency and choice, were largely

unable to pursue goals such as data checking, correction or deletion, and their perceived sense of power within the provider relationship was largely unchanged by the experience. Nor does the GDPR allow individuals to adequately pursue their own goals related to accountability, self-reflection or creative data exploration (see 5.4.3.3). Individuals cannot be given power over their data through designing better Human-Data Interaction interfaces alone, but only through redesigned policies and business strategies that take into account the sociotechnical context [REF 12, REF 17].

In order to bring the human-centric 'personal data ecosystem' concept closer to reality, action must be taken to improve both compliance and quality of GDPR responses. Considering our findings, there is cause for radical policy reform, to move away from 'data access as package delivery' and to provide individuals a more effective and ongoing two-way window into their data (see 5.5.1), providing ongoing awareness, accountability, and negotiability. Data needs to be expressed to individuals in ways they can understand, as little to no practical impact is currently being achieved by delivery of a one-time snapshot of some technical files.

For providers, the risk of reputational damage uncovered by this study should motivate them to engage meaningfully with data access requests; but such risk can be averted by redesigning both interfaces and processes to approach data access experiences as an opportunity to educate, and to build trust and loyalty, perhaps even through establishing progressive co-operative data stewardship relationships that truly *involve* the service user (see 5.5.2). While the GDPR experience is often disappointing and frustrating, it can provide insights that help individuals to challenge their assumptions, re-evaluate choices, and in some rare cases, feel empowered to act upon their data. Wider assertion of GDPR rights could demonstrate a desire for data holders to be transparent; without such visible demand, little may change (see 5.5.3).

Considering RQ1 (the pursuit of a a deeper understanding of people's attitudes to everyday data holding and people's wants from that data),

this study suggests that people struggle to develop the meaningful relationship with their data that they desire, because of the difficulties faced in seeing, accessing and understanding it. They are aware that within data is the potential for value to themselves, but cannot access that value, which in turn causes feelings of resignation, concern, distrust or suspicion towards data holders. What they seek most are two things: sufficient understanding to better judge the value exchange they have signed up for with providers (see goals in top half of Table 12), and good quality insights from data that would allow them to understand themselves better, learn from the past, set personal goals, and harness personal data for their individual benefit (see goals in lower half of Table 12). This duality of needs around data interaction is expanded upon in (Bowyer, 2021).

With respect to RQ2 (the pursuit of a better understanding of the role of that data in everyday service relationships), the findings suggest that personal data, held by providers, serves as a proxy for direct user involvement, and is treated as such. Once users have sacrificed their data, or given permission for its collection, they are rarely consulted and most services exclude them from seeing how that data will travel through the organisation and be used in decision making; this is consistent with the 'point of severance' concept observed by Luger and Rodden (Luger and Rodden, 2013). As a result, the trust relationship between service provider and service user is extremely fragile, highly susceptible to subjective impressions of service brands, and as the findings show, discovery of poor data practices or a lack of transparency around data is sufficient to harm that relationship and in some cases even motivate individuals to change provider. As discomfort grows and scrutiny occurs, individuals can be expected to lose trust and loyalty. At the same time, this same data could play a central role in a re-invigorated relationship between a provider and a user, one based upon earned trust. It appears that providing easy, clear, data access and showing a willingness to respond to questions and explain data usage to users could be sufficient to allay concerns and instil strong customer loyalty. Of course, this assumes that the openness offered reveals practices the user finds agreeable, so

perhaps this in some way explains why some companies that have more commercially-motivated approaches to personal data use, such as Facebook and Google, are apparently less willing to engage in transparency and user empoerment around data.

The general principles of earning trust through transparency, and rethinking data access as a means to involve users in decision-making, could be applied in a wide range of service endeavours that are currently very data-centric.

Bibliography

Abowd, G. D. (2012) 'What next, ubicomp?: celebrating an intellectual disappearing act', in *Proceedings of the 2012 acm conference on ubiquitous computing*. New York, New York, USA: ACM Press, pp. 31–40. doi: http://dx.doi.org/10.1145/2370216.2370222.

Abowd, G. D. and Mynatt, E. D. (2000) 'Charting Past, Present, and Future Research in Ubiquitous Computing', *ACM Transactions on Computer-Human Interaction*, 7(1), pp. 29–58. doi: 10.1145/344949.344988.

Alizadeh, F. *et al.* (2019) 'GDPR-reality check on the right to access data', in *ACM international conference proceeding series*. New York, New York, USA: ACM Press, pp. 811–814. doi: 10.1145/3340764.3344913.

Arfelt, E., Basin, D. and Debois, S. (2019) 'Monitoring the GDPR', in *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, pp. 681–699. doi: 10.1007/978-3-030-29959-0_33.

Ausloos, J. (2019) ‘GDPR Transparency as a Research Method’, *SSRN Electronic Journal*, (May), pp. 1–23. doi: 10.2139/ssrn.3465680.

Ausloos, J. and Dewitte, P. (2018) *Shattering one-way mirrors-data subject access rights in practice*. Available at: www.irissproject.eu https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3106632.

Ausloos, J. and Veale, M. (2020) ‘Researching with Data Rights’, *Technology and Regulation*, pp. 136–157.

BBC R&D (2017) ‘Human Data Interaction - BBC R&D’. Available at: https://www.bbc.co.uk/rd/projects/human-data-interaction.

Bowyer, A. (2021) ‘Human-Data Interaction has two purposes: Personal Data Control and Life Information Exploration’. Available at: https://eprints.ncl.ac.uk/273832#.

Bufalieri, L. *et al.* (2020) ‘GDPR: When the right to access personal data becomes a threat’. doi: 10.1109/icws49710.2020.00017.

Chang, A. (2018) ‘The Facebook and Cambridge Analytica scandal, explained with a simple diagram - Vox’. Available at: https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram.

Clarke, N. *et al.* (2019) ‘GDPR: an impediment to research?’, *Irish Journal of Medical Science (1971-)*. Springer, 188(4), pp. 1129–1135.

Comandè, G. and Schneider, G. (2021) ‘Can the GDPR make data flow for research easier? Yes it can, by differentiating! A careful reading of the GDPR shows how EU data protection law leaves open some significant flexibilities for data protection-sound research activities’, *Computer Law & Security Review*. Elsevier, 41, p. 105539.

Council of the European Union (2015) ‘Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)’.

Brussels. Available at: http://data.consilium.europa.eu/doc/document/ST-9565-2015-INIT/en/pdf.

Crabtree, A. and Mortier, R. (2016) 'Personal Data, Privacy and the Internet of Things: The Shifting Locus of Agency and Control', *SSRN Electronic Journal*, pp. 1–20. doi: 10.2139/ssrn.2874312.

'Facebook–Cambridge Analytica Data Scandal' (2014). Available at: https://en.wikipedia.org/wiki/Facebook–Cambridge_Analytica_data_scandal.

'Facebook - Data Policy' (no date). Available at: https://www.facebook.com/about/privacy (Accessed: 9 August 2021).

Gellman, B. (2013) 'Edward Snowden, after months of NSA revelations, says his mission's accomplished', *The Washington Post*, 23. Available at: http://www.washingtonpost.com/world/national-security/edward-snowden-after-months-of-nsa-revelations-says-his-missions-accomplished/2013/12/23/49fc36de-6c1c-11e3-a523-fe73f0ff6b8d_story.html%5Cnhttp://www.washingtonpost.com/world/national-security/edward-.

Glavic, B. *et al.* (2021) 'Trends in Explanations: Understanding and Debugging Data-driven Systems', *Foundations and Trendsin Databases*. Now Publishers, Inc., 11(3), pp. 226–318. doi: 10.1561/XXXXXXXXX.Boris.

Gonscherowski, S. and Bieker, F. (2018) 'Who You Gonna Call When There's Something Wrong in Your Processing? Risk Assessment and Data Breach Notifications in Practice', in *IFIP international summer school on privacy and identity management*. Springer, pp. 35–50.

Gurstein, M. B. (2011) 'Open data: Empowering the empowered or effective data use for everyone?', *First Monday*. First Monday, 16(2). doi: 10.5210/fm.v16i2.3316.

Hamon, R. *et al.* (2021) 'Impossible Explanations? Beyond explainable AI in the GDPR from a COVID-19 use case scenario', in *Proceedings*

of the 2021 acm conference on fairness, accountability, and transparency, pp. 549–559.

'HDI Lab, Heerlen' (2020). Available at: https://hdilab.com/.

'HDI Network Plus, University of Glasgow' (2018). Available at: https://hdi-network.org/.

Human, S. and Cech, F. (2021) 'A human-centric perspective on digital consenting: The case of GAFAM', *Smart Innovation, Systems and Technologies*, 189, pp. 139–159. doi: 10.1007/978-981-15-5784-2_12.

'Human Data Interaction Project at the Data to AI Lab, MIT' (2015). Available at: https://hdi-dai.lids.mit.edu/.

Hwang, E. (2021) 'Sketching Dialogue : Incorporating Sketching in Emphatic Semi-structured Interviews for HCI'.

Information Commissioner's Office (2018) 'Your data matters - Your rights'. Available at: https://ico.org.uk/your-data-matters/.

Information Commissioner's Office (2021a) 'Your right of access'. Available at: https://ico.org.uk/your-data-matters/your-right-to-get-copies-of-your-data/ (Accessed: 23 August 2021).

Information Commissioner's Office (2021b) 'Your right to data portability'.

Kasirzadeh, A. and Clifford, D. (2021) *Fairness and Data Protection Impact Assessments*. Association for Computing Machinery (1), pp. 146–153. doi: 10.1145/3461702.3462528.

'List of target companies for GDPR requests' (no date). Available at: https://wiki.personaldata.io/wiki/Item:Q2369 (Accessed: 22 September 2021).

Luger, E. and Rodden, T. (2013) 'An informed view on consent for ubicomp', in *UbiComp 2013 - proceedings of the 2013 acm international joint conference on pervasive and ubiquitous computing*. New York, New York, USA: ACM Press, pp. 529–538. doi:

10.1145/2493432.2493446.

McCarthy, J. and Wright, P. (2004) 'Technology as experience', *Interactions*, 11(5), pp. 42–43. doi: 10.1145/1015530.1015549.

Mortier, R. *et al.* (2014) 'Human-data interaction: The human face of the data-driven society', *Available at SSRN 2508051*. doi: 10.2139/ssrn.2508051.

MyData (2017) 'Declaration - MyData.org'. Available at: https://mydata.org/declaration/ (Accessed: 8 November 2019).

MyData.org (2018) 'MyData - Who we are'. Available at: https://mydata.org/about/.

Mydex CIC (2010) 'The Case for Personal Information Empowerment : The rise of the personal data store', *World*, pp. 1–44.

O'Donnell, B. (2020) 'Zoom, the office and the future: What will work look like after coronavirus?' Available at: https://eu.usatoday.com/story/tech/columnist/2020/09/07/zoom-work-from-home-future-office-after-coronavirus/5680284002/.

'Privacy' (no date). Available at: https://privacy.linkedin.com/ (Accessed: 9 August 2021).

'Privacy - Apple (UK)' (no date). Available at: https://www.apple.com/uk/privacy/ (Accessed: 9 August 2021).

'Privacy & Terms – Google' (no date). Available at: https://policies.google.com/ (Accessed: 9 August 2021).

Quinn, P. (2021) 'Research under the GDPR–a level playing field for public and private sector research?', *Life Sciences, Society and Policy*. Springer, 17(1), pp. 1–33.

Savage, A. and Hyde, R. (2014) 'Using freedom of information requests to facilitate research', *International Journal of Social Research Methodology*. Routledge, 17(3), pp. 303–317. doi: 10.1080/13645579.2012.742280.

Spagnuelo, D., Ferreira, A. and Lenzini, G. (2019) ‘Accomplishing Transparency within the General Data Protection Regulation.’, in *ICISSP*, pp. 114–125.

Symons, T. *et al.* (2017) ‘Me, my data and I: The future of the personal data economy’, *DECODE (DEecentralised Citizen Owned Data Ecosystems) Report*, (732546), p. 88. Available at: https://media.nesta.org.uk/documents/decode-02.pdf.

The European Parliament and the Council of the European Union (2016) ‘Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data’. Available at: https://eur-lex.europa.eu/eli/reg/2016/679/oj https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=ES.

Waldman, A. E. (2020) ‘Data Protection by Design ? A Critique of Article 25 of the GDPR’, 1239(2019), pp. 147–168.

Weiser, M. (1991) ‘The computer for the 21st century’, *Scientific American*, 265(3), pp. 94–105. doi: 10.1145/329124.329126.

Wiki.personaldata.io (no date) ‘Subject Access Request Template’. Available at: https://wiki.personaldata.io/wiki/Template:Access (Accessed: 21 September 2021).

Wong, J. and Henderson, T. (2018) ‘How Portable is Portable ? Exercising the GDPR ’ s Right to Data Portability’, *Acm*, pp. 911–920.

Woolgar, S. (2014) ‘Configuring the User: The Case of Usability Trials’, *The Sociological Review*, 38, pp. 58–99. doi: 10.1111/j.1467-954x.1990.tb03349.x.

Wright, P. and McCarthy, J. (2008) ‘Empathy and experience in HCI’, *Conference on Human Factors in Computing Systems - Proceedings*, pp. 637–646. doi: 10.1145/1357054.1357156.

Zuckerman, E. (2021) *Mistrust: Why Losing Faith In Institutions*

Provides The Tools To Transform Them. New York, NY, USA: W. W. Norton & Company, pp. 1–3. doi: 10.1017/ipo.2021.30.

---

1. 11 participants started the study but one dropped out after the first interview due to COVID-19. 31 interviews were conducted in all.↩

2. In this study and throughout this thesis, my usage of the word 'want' in the context of data capabilities deliberately includes both meanings of the word: the need or desire of the individual, but also that which they **lack**.↩

3. At the time of writing (autumn 2021) the GPDR legally applies in both the European Union and the United Kingdom, which have a total population of 513 million individuals [37]. GDPR rights are also conferred to any individual who is a customer of businesses with registered offices in EU or UK countries, meaning that these rights are in effect globally available for non-EU, non-UK users of many multi-national digital service providers.↩