Understanding the Family Perspective on the Storage, Sharing and Handling of Family Civic Data

Alex Bowyer1, Kyle Montague1, Stuart Wheater2, Ruth McGovern3, Raghu Lingam3, Madeline Balaam4

Open Lab, Newcastle
upon Tyne, UK
{kyle.montague, a.bowyer2, stuart.wheater}@ncl.ac.uk
Arjuna Technologies,
Newcastle upon Tyne,
UK
@ncl.ac.uk
@arjuna.com
Institute of Health & Society, Newcastle
upon Tyne, UK
initial.lastname
@ncl.ac.uk
Media Technology and Interaction Design,
KTH Royal Institute of Technology,
Stockholm, Sweden
balaam@kth.se

ABSTRACT

Across social care, healthcare and public policy, enabled by the “big data” revolution (which has normalized large-scale data-based decision-making), there are moves to “join up” citizen databases to provide care workers with holistic views of families they support. In this context, questions of personal data privacy, security, access, control and (dis-)empowerment are critical considerations for system designers and policy makers alike.

To explore the family perspective on this landscape of what we call Family Civic Data, we carried out ethnographic interviews with four North-East families. Our design-game-based interviews were effective for engaging both adults and children to talk about the impact of this dry, technical topic on their lives. Our findings, delivered in the form of design guidelines, show support for dynamic consent: families would feel most empowered if involved in an ongoing co-operative relationship with state welfare and civic authorities through shared interaction with their data.

Author Keywords
User-centered design; personal data; social care; healthcare; big data; civic data; data privacy; ubicomp; data security; data sharing; dynamic consent; design games; family design games; boundary objects; family research; ethnographic interviews; family civic data; citizen-state relations; GDPR

ACM Classification Keywords
H.5.2. User Interfaces: User-centered design

INTRODUCTION

Government and civic authorities in the UK, as in other countries, are seeking ways to better serve the most disadvantaged sections of society. Issues including youth crime, poverty, unemployment, homelessness, truancy and domestic abuse, bring misery and upset to British families as well as significant cost to the state [18]. Civic authorities, health trusts and welfare organizations hope to create a better quality of life for vulnerable or at-risk families by using more “joined-up” data-handling practices to ensure that public sector workers who support families directly are able to see the complete picture of a family’s life [44–48]. However, the sharing of data about families holds the potential to significantly disempower them as citizens, by hiding individual facts or additional context and enabling group-based profiling [4,34]. Decisions made based upon shared but erroneous data could have profound knock-on impacts across all parts of a family’s life as they interact with different services, a risk that the EU aims to specifically address through legislation [49]. Families could also be exposed to identity fraud, social problems or discrimination if their private data is shared too widely. In this context, it is critical that the family perspective about how their data will potentially be stored, handled and shared is understood, and the primary aim of our research is to address this gap in understanding.

In this paper, we contribute a new understanding that families would like their Family Civic Data1 to be handled as a precious commodity that must be looked after in partnership with the state through ongoing conversations and a dynamic consent model. We describe three guidelines that can inform the design of future civic data-handling systems. Additionally, we share our methodology, Family Design Games, which we found highly effective at motivating intergenerational participation and contribution from both adults and children to a meaningful discussion of the topic of civic data, and explain the steps needed to adapt this technique to explore other tricky research topics.

1 We introduce the term Family Civic Data as a useful shorthand phrase. See Methodology for definition.
The Troubled Families Program (TFP)
In 2012, the UK government implemented the Troubled Families Program [18,23], having found that £9 billion per annum was being spent through public services upon just 120,000 families, due to their reduced life chances, long-term health problems and reduced productivity at school and at work. The TFP funded new initiatives to connect social care, health and welfare services, schools and police so that social workers could see the complete set of issues affecting a family in one place, with an objective of breaking repeated generational cycles of problems through better-informed and targeted interventions. At its outset, the TFP focused on direct family engagement, but soon evolved to harness large-scale data-sharing technologies. The sharing of civic data between data-holding authorities has been shown to be useful in giving citizens a better quality of service [12], and a thriving international data linkage network now exists, especially in healthcare [21]. However, projects reliant upon sharing of sensitive family data must obtain a family’s consent. The NHS’s failed care.data project, which closed following widespread public distrust of the data collection and sharing it required, showed that if these systems are to work they must reflect what families and citizens want and need [50]. A good example of this is gov.uk [35], an e-government system designed from a user-centric ‘citizen needs’ perspective, which has been shown to promote accountability [42] and increase trust and confidence in government [38].

Data Sharing, Data Handling and Consent
Consent to share data has been most thoroughly examined in the medical profession, where it offers clear, uncontroversial benefits that are easy for families to see [5]. Informed consent (where a patient is informed about intended data usage then given an “opt-in” or “out-out” choice), has become the norm and increases trust [36]. Patients are happy to share their data for research use [41]. However, patient data is routinely used for purposes beyond those for which consent was given, suggesting that a simple Yes/No view of consent is insufficient to cope with the rapidly changing informational needs of health and social care professionals and patients in the big data era [27,43].

New and competing models for consent have emerged. In open consent, patients provide broad, unrestricted consent to sharing data for unspecified purposes [20]. Here, the sharing of truths is deemed important for the good of public health, more so than individual privacy. A competing model is that of dynamic consent, which recognizes the need to provide protection for individual interests in the complex and evolving landscape of data sharing [14,15,40]. Here, consent is viewed as an ongoing relationship, where the patient is kept informed of how their data is being used and is free to express different preferences at any time.

Informed consent has been widely adopted in research and in software design, and is often considered adequate to satisfy the legal obligations upon data-holders [11]. In HCI, research has shown that software designs which ask users to give informed consent to standard end-user-license agreements (EULAs) during installation are ineffective, with 74% of people rarely or never browsing them [13]. Patrick and Kenny suggest four design principles for a more privacy-aware interface design: comprehension, consciousness, control and consent [29]. Consent is paramount because it is “the primary means for individuals to exercise their autonomy and protect their privacy” [8]. Morrison et. al. argue that we need to increase end-user understanding of the consent question, and propose a delayed intervention approach, where the user is asked for consent in the context of their task using a visualized, personalized representation of relevant data to make the question more meaningful [24]. This “just-in-time” approach has been endorsed in the US Federal Trade Commission’s best practice guidelines for mobile app developers [9], and Facebook also follows this model.

Such approaches may be more effective, but still do not address a key failing of informed consent – that users must be able to meaningfully and knowingly revoke consent as easily as they give it. Consent routinely becomes a ‘point of severance’, beyond which users surrender all agency and control over their personal data [19]. The field of Ubiquitous Computing (ubicomp) has paid particular attention to consent in recent years, due to the complexities which arise when citizens use multiple devices in many contexts; a consent given in one context may no longer apply in another. Luger and Rodden suggest harnessing Friedman [10]’s idea of ‘actively embedding human values’ in the systems we design and recasting consent as a social process [19], where communication replaces ‘informing’ and users are given regular and comprehensible feedback to support an ongoing relationship with their data – a design that exemplifies the dynamic model of consent.

Family Civic Data
We introduce the term Family Civic Data to refer to the data about families held by any local or national authority - any and all data that might be stored about a family. By ‘civic’ we refer to the administration of society and citizen relations as a whole, not just of the local region, as a preferable term to ‘social’, which carries online and leisure connotations. Our working definition of ‘data’ is ‘information that has been stored’ (typically on a computer). It is important to consider the meaning of ‘family’. Using any single factor such as biology, co-residence, or parenting, is insufficient for defining family [26]. We took a pragmatic approach of including any self-declared family group, including single parents. Cornford et. al. [6] have identified that this ambiguity around the concept of family causes great difficulty for the state, especially when attempting to construct a holistic view of a family, as it is difficult to identify which individuals to include and whose records to link together. They conclude that a simple database is insufficient for encapsulating a family’s details, and that any system handling Family Civic
Data must be agile and configurable by both professionals and citizens, and should support aggregation, negotiation, personal development, planning and more complex capabilities.

As part of our grey-literature review of the field, we collected examples of different types of Family Civic Data, in order to develop a taxonomy of terms and data types, as shown in Table 1.

<table>
<thead>
<tr>
<th>S</th>
<th>Type</th>
<th>Examples / Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Personal details</td>
<td>Date of birth, address, telephone number.</td>
</tr>
<tr>
<td>2</td>
<td>Relationships</td>
<td>Marital status, ex’s, step-parents.</td>
</tr>
<tr>
<td>3</td>
<td>Children</td>
<td>Parentage, adoption, fostering, childcare.</td>
</tr>
<tr>
<td>4</td>
<td>School Records</td>
<td>Attendance (truancy), special needs.</td>
</tr>
<tr>
<td>5</td>
<td>Academic Results</td>
<td>SATs, reports, exam failures, training courses</td>
</tr>
<tr>
<td>6</td>
<td>Social Support</td>
<td>Social worker visits &amp; notes, details of family crises, interventions, allegations.</td>
</tr>
<tr>
<td>7</td>
<td>Welfare Benefits</td>
<td>Jobseeker’s Allowance, child support, Disability Living Allowance, tax credits</td>
</tr>
<tr>
<td>8</td>
<td>Family Finances</td>
<td>Salary, savings, credit cards, spending, debt</td>
</tr>
<tr>
<td>9</td>
<td>Employment</td>
<td>Job history, periods of unemployment, performance at work, NI, PAYE, pensions</td>
</tr>
<tr>
<td>10</td>
<td>Housing</td>
<td>Council house provision, eligibility criteria.</td>
</tr>
<tr>
<td>11</td>
<td>Legal documents</td>
<td>Birth/marriage/death certificates, citizenship/immigration status, work permits</td>
</tr>
<tr>
<td>12</td>
<td>Criminal records</td>
<td>Arrests, cautions, offenders’ registers, prison time, speeding tickets, spent convictions</td>
</tr>
<tr>
<td>13</td>
<td>Court orders</td>
<td>Restraining orders, lawsuits, custody, ASBOs</td>
</tr>
<tr>
<td>14</td>
<td>Domestic Violence</td>
<td>Allegations made, medical records, social/legal interventions, victim support</td>
</tr>
<tr>
<td>15</td>
<td>GP records</td>
<td>GP’s notes, prescriptions, tests, referrals</td>
</tr>
<tr>
<td>16</td>
<td>Hospital records</td>
<td>Operations, hospital stays, emergency care</td>
</tr>
<tr>
<td>17</td>
<td>Medical conditions</td>
<td>Diagnoses, diseases, allergies, blood type</td>
</tr>
<tr>
<td>18</td>
<td>Mental health</td>
<td>PTSD, breakdowns, depression, sectioning</td>
</tr>
<tr>
<td>19</td>
<td>Addictions</td>
<td>Substance abuse, gambling, rehab, crime</td>
</tr>
<tr>
<td>20</td>
<td>Library Usage</td>
<td>Books/CDs borrowed, computer access</td>
</tr>
<tr>
<td>21</td>
<td>Sports &amp; Health</td>
<td>Gym usage, class attendance</td>
</tr>
<tr>
<td>22</td>
<td>Shopping Habits †</td>
<td>Loyalty cards, store &amp; online purchases</td>
</tr>
<tr>
<td>23</td>
<td>Transport Data †</td>
<td>Buses used, ANPR tracking, walking patterns</td>
</tr>
</tbody>
</table>

Table 1. Example Categories of Family Civic Data.

Super-categories (S) are: family (1), education (2), welfare (3), money/work data (4), civil (5), crime (6), medical (7), leisure (8).

In this paper, we augment existing research, which frames the storage, sharing and combining of Family Civic Data as a problem for the state best solved through technology [44–48], with a meaningful understanding of the family perspective. In doing so, we present a contrary perspective that families require more, not less, human interaction. We show that while it is possible to design new, holistic systems for presenting and interacting with Family Civic Data, these will only be effective at helping families if such systems are designed to meet both citizen and practitioner needs and facilitate use by both parties in close collaboration as part of a human interaction.

**METHODOLOGY**

In order to address our main research aim of understanding how families would prefer their data to be handled, shared and stored, we set ourselves two objectives. The first was to design methods of interacting that would allow families to meaningfully converse about the dry, abstract topic of civic data. Our second was to motivate and engage participation from the whole family, including children. Our method, *Family Design Games*, which was ethically approved by our University’s IRB, consists of a home-based ethnographic interview [22], structured as a series of games designed to make the topic interesting and to encourage intergenerational conversation about Family Civic Data. Our approach draws from the *design games* of Brandt and Messeter [2,3], but also from Sarah Pink’s sensory ethnography techniques [30]. Interviews-in-place can offer deeper insights than a traditional interview; as Dawn Mannay observes, *life happens around you* [22]. Home visits enabled us to interact with the whole family at once, grounding our design understanding in their lived experience. Projects such as VOME [51] have used mixed methods to research attitudes to privacy and consent, but much human-data interaction research has focused upon the individual, so our focus on engaging the whole family together through design games in the family home brings a fresh perspective to this under-researched field.

Our interview schedule, showing the questions posed by each activity, is included in the supplemental materials. Each session began by inviting families to consider who might store information about a family. This established a shared understanding of Family Civic Data, while also encouraging broader thinking about data. Then, in our *Family Facts* game, we presented participants with a series of six *Petsals* (writable paper tags to be hooked over poles) containing fictional ‘facts’ about a family, for example ‘Mum Sarah (54) buys three packets of cigarettes every week.’ These scenarios were designed to provoke consideration of information ownership. The family was invited to discuss each Fact and categorize it as “Not Data”, “Our Data” (meaning, this belongs to the family) or “Their Data” (meaning, this belongs to the state or data-holding authority) by putting it over the appropriate pole (see Figure 2). Participants then created their own Facts, drawing inspiration from their own lives, and discussed and categorized them in the same way. The objective was to
endow an appreciation of the inherent conflicts of interest between family and state on the storage and sharing of Family Civic Data, and to prime participants to be able to consider who should control it.

**Card Sorting**

Inspired by Brandt and Messeter’s observation that game pieces can be used to create common ground and as “things-to-think-with” [3,28], we created a set of Data Cards, that would serve as a visual and tangible representation of Family Civic Data, and as boundary objects [1,17] to bring researcher and participants’ worlds closer together. The development of the cards and associated lexicon made accessible the topic of data, which, as we discovered during recruitment, many families feel unqualified to talk about. A Data Card was created for each category in Table 1, including a summary and meaningful examples, so that the cards would be easy to digest, yet still contain sufficient detail to stimulate thinking. Keeping child-friendliness in mind, bright colors were a key element of the design. The cards (see Figure 1) were printed on high-quality, thick card with a glossy finish using a business card printing service.

For our main exercise, families were invited to perform a card sorting task using these Data Cards. We labelled the extremities of a corkboard so as to give meaning to the space on the board, creating two notional graph axes – a ‘control’ axis, and a ‘risk’ axis (see Figure 3). A family member was invited to pick a Data Card from the deck, read out the card’s data type, summary and example data, and then discuss with their family where to pin the card on the board, considering:

i. how much control they think the family (‘we’) should have of that data versus the authority holding the data (‘they’),
ii. how risky it would be if this type of data was lost, stored with mistakes in it, or shared with the wrong people, and
iii. how it should be placed relative to other cards already pinned on the board.
This was repeated around a dozen times, with a different family member picking the card each time. Figure 4 shows a family carrying out this activity.

We then conducted a follow-on exercise (designed to be fun for all ages) to consider how each data type should be handled. A set of custom stickers were produced, containing statements that we thought a family member might say about a specific type of data, for example, “this makes me uncomfortable” or “I don’t care about that.” Plain white stickers were also available so that participants could write their own where no suitable sticker was provided for the idea they wished to express.

Reviewing and Reflecting on a Mock-up Interface

We then interviewed families about the requirements for an interface to allow access to and control of a family’s data. We contextualised this discussion by providing printouts from spreadsheets of (fictional) data of arrests, benefits and domestic violence cases (which is how Family Civic Data is currently shared between agencies), as well as some news articles illustrating what can go wrong when citizens’ data is leaked, or contains errors or omissions. Then, we presented two single-page printed mock-up interfaces for TFP social workers (see Figure 5), explaining how a social worker might view and use a family’s data. Finally, we invited comment on the design and asked participants what they considered the most important features or attributes of a computer system that handles Family Civic Data.

Family 1 was an academic couple in their twenties (with an infant, not present), who took up the option of meeting at our University rather than in their home; the remaining sessions were all home visits. Family 2 was a twentysomething mother along with her 5- and 12-year-old children and her fiftysomething mother. Family 3 was a couple in their thirties, whose children were not present. Family 4 was a single mother in her forties, and her four youngest children, aged between 6 and 15.

FINDINGS

The primary purpose of our activities was to produce a good conversation which we would then be able to analyze. We photographed the physical outputs, during and after each session, but their sole purpose was to add contextual understanding to conversation transcripts. Transcripts and field notes were coded and organized through inductive data display cycles, as per the Miles and Huberman framework [32]. The themes uncovered by our analysis are presented below under three main questions: what does this data mean to the family; when should it be in control of the authorities; and how should it be handled? In addition to these main qualitative findings, we use a quantitative analysis to show how families collectively think different types of data should be handled. Finally, ethnographic and methodological observations are recorded separately.

What does Family Civic Data mean to the family?

Data (and information) is private

In the Family Facts exercise, families talked about everyday information about their lives as being no-one else’s business — “just personal [private] stuff that no-one has written down” (F2) — that only becomes data when it is reported or someone wants to know about it. Families expressed a desire to keep control of their “delicate”, “sensitive” (F1) information, which “should be kept private until we say otherwise.” (F4). Relationship data regularly came up as an example of “inherently personal” (F1) data that “no-one has a right to know” (F2).

Data handling requires trust

Because information can “reveal a lot about a person” (F1), it entrusts a great deal of power to the data-holder. Families desire to keep control of their data was often motivated by fear; they expressed a desire to keep control when they perceived risks of it being shared or misused, fearing the consequences of mishandling. Often implicit in this was a distrust of authorities. For example, (F3), who we discovered had been heavily involved with the criminal and welfare system as a result of addiction-related issues, felt that authorities often shared, withheld or misrepresented their data against the family interests, and thus wanted full control of all the data about their situation.

Participants

Working with a local community center as well as through word-of-mouth, we recruited four families from the region. Two hours was allowed for each family visit, and families were rewarded with a gift voucher for a family day out. Families were reassured that they would only need to talk about families in general, and that sharing specifics from their own lives would be entirely optional. All sessions were audio recorded and field notes were written later the same day after each session.

2 In this report, we use the shorthand (F1) to refer to ‘a member of Family 1,’ or in some cases to the whole family.
Data is risky
All Family Civic Data was perceived as risky. More severe perceived consequences of data mishandling led to families wanting more control. Families perceived a variety of risks – criminal, medical, welfare, social and psychological:

Hacking and identity fraud was often mentioned; (F2) noted that “when you think of the forms you fill in to apply for a course, we give them everything you need to clone us.” (F1) were afraid to complain to police about a neighbor for fear of retribution if their anonymity were compromised, while (F2) feared that knowledge of their salary information would make them a target for burglary. Information about parentage was identified as carrying risk of significant psychological damage, “He thinks that is his dad and he doesn’t know any different. But obviously with him having social and emotional problems, if he found out, it would mess with him.” (F3) Families were also concerned by the potential for social stigmatization if information became public, especially in the form of gossip from neighbors. (F2) recalled how she and her sister “wouldn’t have any of the free school meals, because you had to line up for a dinner ticket, then everyone knew you were different.” However, families could see the benefits of certain risky data being stored and appropriately shared, for example the medical risks if their medical histories were not made available to doctors and hospitals. (F3) recounted an incident where a probation worker’s refusal to disclose his whereabouts to his wife put his own welfare at risk.

Data is misleading
A common problem the families observed about data stored about them, is that there is no absolute truth; a lot of the data is recorded by partisan observers meaning it is at best subjective and at worst purely an opinion. While looking at the mock-up interfaces that invited social workers to provide numerical scores for troubled families’ progress, (F3) highlighted that these are just opinions, adding “It’s like they’re giving you marks... it shouldn’t be like that”. (F2) talked about the subjectivity that schools and playgroups would exert when deciding whether to (i) record and (ii) share details of an injury that a child had sustained while in their care. (F3) said that their social worker recorded details of the couple’s arguments, and threatened to report these to social services. (F2) talked about being wrongly classified as a troubled family because of their postcode. Ultimately, the families’ observations show that data can disempower a family, especially when they are labelled with categories such as ‘domestic violence’ or ‘troubled family’ and then judged solely by that label. In response to the Family Facts activity, all families felt that the scenario regarding a mother who had had a breakdown ten years earlier would lead to unfair judgement. The families noted that anyone could fall foul of such unfair judgements, even if they have done nothing wrong – “False allegations can be very damaging.” (F1) and “anyone can be harmed by data sharing.” (F3)

When does Family Civic Data belong to the state?

When “they” need to know
Families acknowledged that data is routinely generated and stored about them through interactions with state services, and that sometimes police, schools, medical professionals, care workers, and employers need that data to operate and perform their job effectively. The question of “Do they need to know?” is the major determinant in deciding whether data should be shared, controlled or accessed by public sector service providers. As (F2) put it “If they need the information, that’s a reason for them to have control [of it].” During our card sorting activity (F2) said they didn’t need the “not for police eyes” sticker because “the police should be allowed to see everything.” (F4) said they wouldn’t mind care workers having access to family data, but “Just them. Me and the people that really need it. Say, like, doctors, social services. Not just any randomers.” (F3) noted that sometimes organizations want, rather than need, your information – “When you’re getting Sky, when you’re getting a house, everyone wants to know if you’re working or on benefits.”

When “they” are responsible
In some cases, families felt that authorities having control of data would be beneficial. For example, F2, following an observation that she’d lost her old exam certificates, said “Most things I’d prefer other people to have control of […] then I know it’s somewhere where it won’t get lost.” Keeping data is a responsibility, and not having that responsibility removes the burden of worry. (F2) observed a similar ‘hassle-removing’ benefit from medical data sharing, saying she would opt-in to such sharing because “it’s a right nightmare when I’ve got to go back and forwards to explain everything.” Other examples where the data-holder was deemed to have a clear responsibility to control the data were medical prescriptions, adoption data (F2), and sensitive data about addictions (F3). Coupled with responsibility – and in line with literature – was a sense of trust, with both (F3) and (F4) noting the trust they put in their social workers and doctors not to share their data.

When the data can serve the public interest
Families observed that in certain circumstances, data should be shared with relevant authorities or, sometimes, with the public at large. (F2) said that criminal records should be shared if there is a danger to the public, and (F1) talked about the importance of sharing information that will protect a child from potential abuse. (F3), from his experience of working in a substance abuse charity, observed that part of the social worker’s responsibility is to interpret data and spot patterns, and in this context, it is good to have access to as much data as possible, giving, as he put it, “many hands, many eyes, many ears, many indicators.” He felt that not enough data sharing happens between public sector agencies, which (F1) also observed from her voluntary experience with Sure Start, a governmental child support initiative. (F4) said people should be able to know who lived in their house previously:
“it could have been a dodgy person and then [...] you could be getting people coming in, and getting their drama in your life.” Despite having viewed tracking people’s journeys as “scary and stalkery”. (F2) did say that there is a wider public benefit to large scale data collection: “If they’re tracking where people are going and there’s not a bus going to that place, they know to put a bus on.”

When the family has nothing to lose
In several cases families were happy for the authority to control their data when they felt there was no risk. (F2) were unconcerned about shopping data and routine medical data being outside of their control because “if somebody got hold of that information, what could they do?” (F4) made the same observation about shopping data and a similar observation about exam results. (F2) said they did not care about library data because they had nothing to hide.

When data is gathered by default, or actively sought
One area where families would prefer state authorities not to hold the data they do, is when excessive data is collected as part of enrolment processes, and kept when not really needed. There was a general sense that too much data is being collected - “It’s like they need to know every single thing going on.” (F4). (F2) observed that “data’s going to be saved somewhere no matter what you say.” Families felt they have no choice about data held by others about them: “They’re all things in databases somewhere that we have no control over” (F2). “Sometimes there’s no way of avoiding it” (F1). Each family identified types of data that should not be stored – commercial (F1, F3), leisure (F4), and relationships (F2). (F1) felt that when asking about illnesses, employers were asking for data they had no right to ask for, and observed that private health insurers would love to obtain data about people’s smoking. (F1) said that health data should never be sold, citing a recent incident where Google obtained NHS data. However, (F2) acknowledged that making data available to companies can be a trade-off: “I wouldn’t get my money off if I didn’t have [a loyalty card that records my purchases].” In this sense, personal data becomes a commodity [33]. Families do not have the right to total control of their data, though: (F2) said while trying to think of a good example for the Family Facts game, that “it’s hard to think of something that is only our data, that they [the authorities] have no claim to.”

When data sticks around
Most families expressed a view that authorities are routinely keeping data around for a long time, longer than is needed. Speaking of a breakdown in her past, (F2) said “I would rather people not know […] Obviously, it’s somewhere on the doctor’s computer, but it’s more your data than his, isn’t it?” (F3) talked about how the “drug addiction” label had stuck to her family so that they are permanently in that category and treated differently by social and medical care professionals. (F1) suggested that the problem was the lack of any legal obligation for authorities to delete families’ information.

How should Family Civic Data be handled in the future?
A set of basic rights - to be informed, involved and accurately represented
As families considered who should control different types of Family Civic Data, and what a future computer system should do, a phrase that came up often was ‘the right to…’. The families would like the right to be informed about and involved with the handling of their data. (F1) said while looking at the TFP interface mock-up, “Families need the right, without a long and costly procedure, without a lot of bureaucracy, to see the information social workers have about them.” (F2) and (F3) also envisaged a “right to know” and a default “right to see.” The mock TFP interfaces were thought to be practitioner-centric and clearly not supportive of these rights. It emerged that access is not enough; families must be actively informed about what data exists and how it is being used. This is in line with Human-Data Interaction (HDI)‘s concept of legibility [25]. (F2) described the proposed social worker interface as “a load of crap”, and (F1) observed that it “treats families like subjects, not as individuals with rights like privacy,” a view that provides evidence for Cornford’s premise that we need to ‘move beyond using information systems that treat [families] as objects to be administered.’ [6] A commonly occurring theme was the idea that information should be fair and accurate; (F1) added a sticker for this, and (F3) talked about the damage caused by labelling or categorizing a family. (F1) and (F4) want (similar to HDI’s concept of agency [25]) the right to choose how data is used, and when to share or disclose it, especially when telling children about adoption. Families also are concerned about incorrect or incomplete data, and need associated rights to deal with that. (F1) described this as the “right of veto” and also wished for a “right of deletion”, (F2) talked about needing the right to explain or add comments when the data does not tell the full picture. The manner of engagement was also discussed – (F4) talked about wanting to go and see someone in person to talk about her data, reminding us that families’ interaction with data should be thought of not just as a technological interface design but as a human interaction, a relationship with representatives of the state. Families want their representation in data to be complete and true. A row in a spreadsheet, or a set of numerical values, is not enough to adequately represent the complexities of families’ situations: “You need to be able to write and explain, because every family is different.” (F2).

One safe place for data, with sensible access controls
When discussing current practices, families felt that emailing spreadsheets, and even social workers retrieving search results of family data as needed, represent a data handling approach which is too scattered, too uncontrolled, and too risky. It was proposed that Family Civic Data should only exist in one place. (F2) said that “you should go to the data rather than have it sent to you” and suggested accessing your data like requesting a record in the library. (F4), viewing herself as not computer-literate,
argued for a place you could physically visit to see, and talk to someone about, your data. (F1) said “data should never leave its source” and that families and practitioners should “go to a place to get the data” in order to keep the data safe. (F3) said “All data should be protected [...] from whoever [we] choose.” A young child in (F4) described how important it was to ‘keep the lock on it’ and have different accounts for each person allowed to access the data, highlighting the importance of being able to tell who has accessed information after the fact. (F1) imagined a sole organization for citizen data management to make families’ relationship with their civic data really work. Interestingly, contrary to our expectations, all the families think systems should not allow them to directly edit their data. (F2) said she lies about her smoking habits whenever the doctor asks her, and observed that families lie to get themselves out of trouble or to deflect official attention. (F3) also cited smoking as something they would lie about, and talked about how people under suspicion can “blag [the authorities] to death and they go away thinking, ‘They’re doing amazing.’” In general, families’ requirements originate from a desire to ensure their Family Civic Data is always fair, accurate and representative – consistent with Cornford’s premise that families need to be observable through data [6]. The requirement to keep data safe and secure will soon become a legal requirement, through the EU’s General Data Protection Regulations [16].

To work together with the state in a positive relationship

One of the most interesting themes that emerged was that the families are willing, and indeed actively want to, work together with authorities to get the best help and support they can. This began with suggestions of discussing, querying and explaining their data, but went further into families wanting to “have a conversation” with the state authorities (F2), and wanting authorities to get to know them as a family with its own unique story – which they are only too happy to tell – not just as a database record. (F3) recognized the importance of having a good relationship with data-holders, and offered a positive vision for the future, suggesting that with better data sharing, “There would be no loopholes, no uncovered stones. People would be working together. To protect something, to move something on, you can’t do it all on your own, you have to work as a collective.” The idea of families being involved in the state’s recording of their data is a radical and beneficial departure from current approaches. (F1) observed that if you invite families to view and comment on their data, you “get them involved, and they think, ‘Oh we are not [just] subjects anymore.’” He further suggested that engaging families like this on a human level could inspire passive families that have “given up” to seize more control of their lives; “Maybe that changes the whole family.”

Quantitative Analysis: Who should own families’ data and what are the risks?

While the sample size was not large enough to be statistically representative, it was possible to analyze the data from the card sorting exercise in a quantitative manner, in order to identify trends in opinion for specific super categories of data. This was done by assigning numerical scores for control and risk as x and y values according to the placement of the Data Cards on the corkboard.

On average (Figure 6), we see that families thought that those types of data that concern society as a whole (Civil, Money/Work, Crime and Welfare) should be more in the control of the state, whereas those data types closer to the individual (Family Information, Education, and Leisure) should be more in family control. One exception was Health data, thought best left to the professionals. For risk, we see a different pattern, with Money/Work data being thought risky, and Crime data especially so. The more “individual” types of data such as Health, Family Information and Education were thought moderately risky. Civil and Leisure data were considered the lowest risk. Looking at variance and standard deviations, the areas where families expressed most varied opinions were around Health, Family Information and Leisure. Crime, Civil, Money/Work and Welfare data were most agreed upon.

![Figure 6. Mean average control and risk per data type.](image)

Size of circles denote relative variance (not to same scale as axes).

In general, families thought Family Civic Data should be slightly more in the data-holder’s control (-1.1, \(\sigma=3.2\)), and that it carries an above average risk (0.5R, \(\sigma=3.3\)).

Ethnographic and Methodological Observations

In all four sessions, adults and, contrary to expectations, the children present (especially those aged between 5 and 12) engaged enthusiastically with the activities, rather than being passive observers to their parents’ participation. All participants seemed to have fun and be stimulated and interested in the activities, which managed to keep the atmosphere light and playful – as seen when (F1) drew her third purple (Crime) card in a row and exclaimed, “Oh no, Alex, not another purple one!” The designs of our cards, stickers and activities had been designed to be child-friendly (in the hope that this would encourage both adult and child participation), and indeed the children were most interested in the game-like aspects – turn-taking, card-picking, pinning cards and placing stickers; even the researcher’s camera was of interest. The drawing and writing elements of the Family Facts task were embraced
by younger participants – e.g. a five-year old boy created a Family Fact by drawing a picture of his late pet tortoise, and labelled it with help from his mother, despite his undeveloped writing skills (F2). In this family setting, literacy was not essential for participation: both parents and researchers took the time to explain things or assist the children, which gave the sessions an unplanned educational aspect. The child-friendly designs did not “dumb down” the topic, and also helped adults to engage easily with the topic.

Although the sessions were structured as games, the activities played out mostly as discussions between family members about where to place Data Cards or Petals, or what should happen to certain data types. The researcher’s role was mainly that of facilitator – occasionally clarifying or sense-checking an idea, or offering an alternative point of view to stimulate the discussion, but largely the researcher was able to step back and observe the family talking with itself. It was these discussions that provided the majority of the findings detailed above. In our role as ethnographic observers, the researchers were able to infer a lot about the lifestyles, relationships and living patterns of the families, even the absent family members, e.g. from photographs or domestic interruptions, that did not occur in the removed setting of (F1)’s interview. Many interesting dynamics between family members were observed – children vying for control of the game, parents ‘talking over’ each other, parental realism contradicting childhood optimism, children revealing past parental conflicts, parents making data decisions for their children, jokes and banter between couples about keeping spending secret from each other or exasperated-but-playful expressions of despair at children’s interruptions. These observations are simply the substance of normal family life; clearly the familiar setting of the home put families at ease, while also revealing hidden aspects of family life and providing researchers with an unexpectedly rich window into families’ lived experiences.

**DISCUSSION**

Through our lexicon of Family Civic Data and the use of Family Design Games in family homes, we have successfully conducted meaningful conversations with families. Our analysis has revealed a new understanding of how families would like their data to be handled. While recognizing that Family Civic Data is private and risky, they are generally happy for their Family Civic Data to be in state control, but are concerned about losing their influence, especially where the stored data may be misleading. The families share a belief that that numerical metrics, problematized labels like ‘domestic violence’ or ‘troubled family’, and rows in spreadsheets, cannot adequately or fairly convey the reality of their lives. This seems to be a systemic problem, as beyond the aforementioned “point of severance” [19] that a simple opt-in/opt-out model of consent [41] creates, the data becomes embedded in the organizations that hold it – it becomes data-in-place [37]. Privacy researchers agree that meaningful consent must be more than a moment-in-time decision [31,39]. Families want to be involved in the stewardship of their data to ensure that those making decisions that affect their lives can always see a full, fair and accurate story of their lives – in this, we see strong evidence that families’ needs can only be met by the dynamic consent model [14,15,40], where management of Family Civic Data becomes an ongoing conversation between family and state.

Implementing a dynamic consent model will be a daunting task for any data-holding organization; it will require cultural change and significant redesign of both analogue processes and data management systems, as outlined by Kaye et al. [40]. It may be hard for a system designer to know where to start. Building upon the ideas and insights gained from our participants, we propose three precepts which can serve as guiding principles to inform the design of a new dynamic-consent-based data-handling system:

**Families must be given continuing rights, control and visibility over their Family Civic Data and its handling, to keep data accurate and meaningful.**

A core principle of dynamic consent is that the citizens about which the data is held must be able to see what data is held about them and how it is being used, at any time, and change their consent. This is not just a matter of displaying the data in an interpretable way, but giving negotiability [25] to families so that they can discuss, challenge and communicate over it and give or revoke consent. Patrick and Kenny’s aforementioned design principles could be helpful here [29]. It is also important to note that this requirement is not just about “doing the right thing for families” – without their ongoing involvement, the gap between the family’s reality and that suggested by their stored data will grow, increasing the risk of the family being mismanaged as a result. Having the family involved serves as a quality control check for the data records.

**Family Civic Data must never be allowed to serve as a proxy for the family’s involvement in decision-making, to avoid disempowerment.**

Our analysis of families’ views shows that in isolation, data about a family stored in an official database will never be enough to adequately capture the complexity of their lives – if you want to tell the whole story of a family, they have to be involved. Similarly, families’ accounts of their lives cannot be assumed as fact, since as our participants told us, families will lie or bend the truth in pursuit of their own interests. As such, to find the objective truth requires an ongoing interpretation of both the state’s data and the family’s statements. The family and the state must work in partnership, and where there is disagreement, there must be discussion and collaboration. The implications of this shift to user ‘control’ of organizations’ data is further explored by Whitley et al. [39].
Families must be able to have a relationship with their Family Civic Data, not just its holders. This requires a mindset shift for data-holding organizations, who must co-operate with each other in a family-centric way.

Through the lens of our research, we can see that Family Civic Data exists in the world, not as a simple description of a family or just as a side-effect of family-state relations, but as a distinct entity with an ongoing lifespan, which can encapsulate facts, opinions and possibly incomplete views of a family’s life. As such, designers should develop interfaces for family involvement that support Human-Data Interaction [25] – where the Family Civic Data is an artifact co-operatively managed by both family and state. The relationship between family and state will need to have this data at its core, rather than in the background, out of the family’s view or influence. The reality is that Family Civic Data is scattered throughout many, often disconnected, databases and systems, and so what is needed is not a unifying technological solution, but rather a paradigm shift in the way that civic organizations think about Family Civic Data. The family’s data should become a living, breathing thing that both the family and the organization(s) will nurture together. There must be no point of severance. This shift in focus will begin to address the families’ need for safe and controlled data handling, while the shift in focus towards co-operative family-state collaboration will improve data quality, by encouraging reconciliation of discrepancies and the establishment of shared and accepted truths. This will leave organizations better placed to meet the requirements of forthcoming GDPR legislation [49].

Reflection: Developing the 'Family Design Games' approach into a re-usable method for engaging families

As well as our insights into families, we believe our evolving approach of Family Design Games can be valuable to other researchers who wish to uncover family perspectives about other dry or complex topics. It can be adapted to other study topics by following these steps:

1. Gather Examples – Find examples of the type of data or thing you wish to explore: in our case, we found real-world examples of Family Civic Data.

2. Categorize – Organize your examples into groups, and name them according to what they represent; this process gave us our lexicon and taxonomy of Family Civic Data.

3. Create Physical Representations – Design a physical object to represent each category that can serve as a “thing-to-think-with” [3,28]. In our case, our Data Cards represented the different types of Family Civic Data.

4. Design Sensitization Activities – Design conversation starter activities that will allow the participants to think about their own life experience in terms of the questions you want to ask. Our Family Facts activity helped the family to think about their own life as data, and to be better equipped to consider questions of data control and risks.

5. Design Activities that ask Questions – Decide what questions you want to explore, and design activities that invite participants to do something with the physical objects that forces them to consider these questions. In our card sorting task, placing each card required consideration of the questions of data control and risk.

6. Conduct the Interviews – Run your design-game-based interviews with your participants, to give a rich transcript that can be analyzed to find their answers to your questions.

A key element of the success of our Family Design Games was our Data Cards, which made the boring, abstract topic of Family Civic Data tangible and relatable. In Tim Coughlan et. al.’s work studying technology use in the home [7], the research team used a dollhouse as a ‘stand-in’ to embody a representation of ‘the home’ in a way that is both relatable to the participants’ own home and life, yet also dissociated enough to allow participants to relax and not feel like they are being examined. Our Data Cards functioned just like Coughlan’s dollhouse, as a focal point that allowed families to talk freely about their own lives and views without feeling personally interrogated. By putting something before participants that is both tangible and separated from real life, you can facilitate a more revealing and personal conversation than a traditional interview.

CONCLUSION

Using our novel Family Design Games technique, we have shown that families – both adults and children – can be engaged in meaningful conversation about the complex and esoteric topic of Family Civic Data. Through these home-based activities we have gained a rich qualitative understanding of families’ requirements for the handling of their data, and derived guidelines that can inform the design of future data-handling systems and associated organizational processes: state organizations must support a dynamic consent model of data handling, and plan for a new paradigm of co-operative, data-based relationships with families, one where meaningful, representative data is nurtured for mutual benefit and families remain involved throughout.

ACKNOWLEDGEMENTS

Thanks to our colleagues at Newcastle University and in local communities across the North-East for their valuable support. This work was funded by the EPSRC CDT in Digital Civics (EP/L016176/1) and partially funded by the Department of Health and Social Care’s Connected Health Cities (North East & North Cumbria) initiative (SILVER sub-project [48]). Data supporting this publication is not openly available due to confidentiality considerations. Access may be possible under appropriate agreement. Additional metadata record at http://dx.doi.org/10.17634/154300-65. Please contact Newcastle Research Data Service at rdm@ncl.ac.uk for further information or access requests.
REFERENCES


50. Controversial database of medical records is scrapped over security concerns | The BMJ. Retrieved August 24, 2017 from http://www.bmj.com/content/354/bmj.i3804