Towards a Person-Centric Interface for Information Re-finding and Sharing Tasks

Paul Holmes

School of Engineering

Thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

May 2018

Acknowledgements

My thanks to my supervisors James Goulding and Sarah Sharples for their time, patience, and feedback. In particular James for encouraging my sentences to become shorter, his optimism about my thesis, and support in developing the prototype used in this research. Sarah for her contributions to the methodology, and knowing the right moment to tell me to stop equivocating about my next study and just go and do it.

I am grateful for financial support from the Horizon Doctoral Training Centre, which allowed me to complete this work. I would also like to thank everybody who participated in my studies.

Edward Anstead who helped me in classifying data and by talking to me about my thesis when I wasn't sure what I was doing. Emily Webber and Laura Lewis for their invaluable time during our tea breaks together and ensuring there was always cheese for us to eat on my birthday every year. My partner Martien for her extensive proofing and provision of comments on this thesis – including her unsuccessful attempt to remove all of the split infinitives – and her considerable emotional and practical support, for which I'm incredibly grateful. And lastly, my cat, Clement Cattlee, who sat in front of my monitor, and my fish Herbert Hoover, Calvin Coolidge, and Henry Kissinger whom I watched instead of writing my thesis.

Abstract

As the amount of information people generate and encounter continues to increase over a diverse array of devices, systems, applications, and websites, their ability to locate and manage information effectively to perform personal information tasks may become increasingly difficult. To understand more about how people perform personal information tasks this thesis explores a specific personal information task, that of information re-finding, and examines in detail what typical everyday information re-finding behaviours consist of and where greater support may be provided to help users perform them. It is posited that one means of providing support for personal information tasks, such as re-finding and sharing, is through a better understanding of how to use and design interfaces focused on connections between people and information.

To understand more about how daily information re-finding behaviours are performed, re-finding tasks were examined through a diary study in which instances where participants searched for information they had previously encountered were recorded over a two-week period. This study identified information about the difficulty users had in performing the information search tasks, the steps taken to re-find the information, the reasons for searching for it, and information remembered about the information sought. Analysis of the diary entries enabled the creation of analytical frameworks describing the goals of users performing re-finding tasks, what information needs they have in performing them, and what people remember about the information they are re-finding. The frameworks were then used to examine the process people undertake when re-finding information. This revealed

the importance of using memories of the information sought in the process of re-finding it. It was noted that participants would recall what the source of the information being sought was and navigate to it using aspects they recalled about it. Most commonly they would recall keywords describing the subject, the time it was encountered, and the people connected to it. Being able to recall these elements was often a critical part of the re-finding process.

When performing information tasks, prominent links were often created in the memory of participants between the information they interacted with and the people connected to it. These connections might be as authors of the information, recipients, sharers, subjects, sources or in relation to our future intentions for information. These connections, however, are often not exploited in the presentation or organisation of information in the design of systems and interfaces, to aid tasks such as re-finding, to the same extent as aspects such as time, or keywords. Providing equivalent support and understanding of how connections between people and information can be utilised was therefore identified as a means of improving support for personal information tasks.

After the identification of the role that the connections between people and information can play in supporting personal information tasks, some means of exploiting these connections to support information re-finding and sharing were considered. Some past research has examined the use of people in relation to information to perform information tasks, primarily applied to information re-finding and sharing. This small body of work, however, has not explored in depth a basis for how best to design interfaces focused on people, to support users in performing personal information tasks in this manner.

Two further studies were therefore conducted to explore how to design interfaces that support a greater focus on people – interfaces that are 'person-centric' in nature. The first of these studies provided a basis for how to design interfaces focused on the use of people for personal

information tasks, and the second evaluated a 'person-centric' design based on the findings of the prior study.

Designing interfaces that provide a means of accessing or sharing information through interaction with personal contacts requires a means of organising and representing those contacts. Just as the diary study revealed some of the prominent dimensions that information is recalled and ordered by, the design of 'person-centric' interfaces requires a similar understanding of what dimensions influence how to structure and order contacts. In order to design an interface based on supporting information tasks using people as a central component of the interface, how people think about relationships between their contacts and how people would choose to represent their contacts was examined.

This was approached in two ways. First, through use of a modified repertory grid method, whereby participants made comparisons between many combinations of their own personal contacts to elicit constructs which described the different forms of relationships they perceived between their contacts. Second, through participants making free-hand drawings that illustrated how they would choose to represent all of their contacts. From analysing this data a categorisation of the different forms of contact groups people perceived was generated. Applying this to the data revealed which forms of expressing relationships were most prevalent, which at the highest level were those related to organisations, locations, relationship types and events. Different forms of representations of contacts, which each structured the contacts in different ways, were also identified from the data. These representations each emphasised different aspects such as overlaps between groups of contacts, the relative importance of contacts, the location of contacts, and the links between contacts.

A prototype 'person-centric' system was then developed to evaluate a range of 'person-centric' design concepts for supporting people in performing information sharing and re-finding tasks. This incorporated design ideas based on examination of the re-finding behaviours

from the diary study and the examination of contact relationships and representations. The prototype system extracted contacts with whom the user had reciprocated contact from their email and Facebook accounts, as well as related messages, files and links. The user was then able to use the system to create groups of contacts, which could then later be used to aid performing information sharing and re-finding tasks through a series of different 'person-centric' interface presentations. The structure of these presentations could be changed, re-ordered, and filtered based on the results of the repertory grid and free-hand diagramming, which provided a basis for understanding different ways people may wish to order or filter contacts, and the different ways to structure contacts in an interface.

The prototype was evaluated in two parts. In the first part the process of contact grouping was studied, and in the second part the design ideas and the use of people as a primary unit of interaction was evaluated. In both parts of the evaluation participants performed think-aloud protocol while interacting with the interfaces to perform the tasks, which was recorded using audio and video capture.

In the evaluation of the contact grouping two grouping interface designs were used. The participant's task with the interfaces was to group all of their contacts in each interface condition. The aim of this was to explore what constituted a meaningful contact group and to understand what role interface design played in this. In the re-finding and information sharing interface evaluation participants were questioned about their personal data in order to generate tasks to re-find information, or items to share from their personal data using the prototype system. Using the identified information items, tasks were set using four different representations that supported information sharing and re-finding. These tasks aimed to evaluate the utility of the different 'person-centric' design concepts that had been identified. Following both parts of the evaluation interviews were conducted to understand the experiences and preferences of the participants in us-

ing the different interfaces.

From the two part evaluation, the decision-making processes of constructing groups and the factors that influence that process were outlined, as well as preferences and behaviour regarding different features of the interfaces that supported performing the information tasks. Evaluation of design ideas from the previous two studies through the prototype confirmed the value of utilising connections made between people and information when re-finding information, and allowed factors that have implications for 'person-centric' interface design to be identified in relation to how contacts and groups should be represented.

Through its examination of the use of connections that are made between people and information when performing personal information tasks and the representation of people in interface design, this thesis provides a greater degree of understanding of typical information re-finding behaviours, the representation and relationships of people in interfaces for performing information tasks, the process of constructing contact groups and what constitutes a meaningful contact grouping. It contributes guidelines that inform how to design 'person-centric' interfaces in relation to the structure of representations, ordering and filtering of contacts, and the linking of information to contacts. An improved understanding of the processes involved in contact group creation and factors that influence it with implications for supporting for manual, semi-automated, and automated group creation and identification. As well as a more complete picture of what information re-finding entails through analytical frameworks that describe the how, what, and why of re-finding tasks.

Contents

Contents **vii**

List of Figures **xiv**

List of Tables **xvii**

Nomenclature **xvii**

1 Introduction **1**

1.1 Aims and Objectives 5

1.2 Objectives 6

1.3 Thesis Overview & Structure 7

1.3.1 Chapter 2 – Literature Review 8

1.3.2 Chapter 3 – Methodology 11

1.3.3 Chapter 4 – Diary Study of Information Re-finding Behaviours (Study 1) 11

1.3.4 Chapter 5 – Contact Grouping: Categorisation and Representation (Study 2) 12

1.3.5 Chapter 6 – Person-Centric Prototype Design 12

1.3.6 Chapter 7 – Person-Centric Prototype Evaluation (Study 3) 13

1.3.7 Chapter 8 – Discussion 14

1.3.8 Chapter 9 – Conclusion 14

1.4 Research Contribution 14

2 Literature Review **17**

2.1 Introduction . . . 17
2.2 Overview of the Use of Time and Keywords in Performing Information Tasks . . . 18
2.3 Supporting Information Tasks . . . 22
2.3.1 Information Re-Finding . . . 22
2.3.1.1 Introduction . . . 22
2.3.1.2 Information Re-finding Behaviours . . . 23
2.3.1.3 Information Re-finding Interfaces . . . 26
2.3.2 Information Sharing and Contact Grouping . . . 32
2.4 Representation of People & Information Behaviour . . . 38
2.4.1 Structure . . . 38
2.4.1.1 Network Focus . . . 38
2.4.2 Behaviour . . . 44
2.4.2.1 Activity Focus . . . 44
2.4.2.2 Content Focus . . . 48
2.5 Conclusion . . . 51

3 Methodology **54**
3.1 Methods . . . 56
3.1.1 Study 1 – Information Re-finding Diary Study . . . 56
3.1.1.1 Alternative Methods . . . 57
3.1.2 Study 2 – Contact Grouping: Categorisation and Representation . . . 58
3.1.2.1 Alternative Methods . . . 60
3.1.3 Study 3 – Person-Centric Prototype Evaluation . . . 62
3.1.3.1 Alternative Methods . . . 63
3.2 Summary of the Approach to the Research Questions . . . 65

4 Re-finding and Information Tasks **70**
4.1 Introduction . . . 70
4.2 Methodology . . . 71
4.2.1 Study Design . . . 71
4.2.2 Participants . . . 74

4.2.3 Procedure . . . 75
4.3 Analytical Frameworks of Information Re-finding . . . 76
4.3.1 Information Needs & Tasks . . . 77
4.3.2 Known Attributes and Associations . . . 82
4.4 Results . . . 85
4.4.1 Information Re-finding Overview . . . 85
4.4.2 Information Sources . . . 85
4.4.3 Success and Failure of Retrieval . . . 88
4.4.4 Difficulty . . . 88
4.4.5 Information Needs & Tasks . . . 90
4.4.6 Known Associations & Attributes . . . 91
4.5 Discussion . . . 97
4.5.1 Limitations of the Research Method . . . 97
4.5.2 Information Needs and Tasks . . . 98
4.5.3 Known Attributes and Information Re-Finding . . . 99
4.5.4 Re-Finding and Person . . . 101
4.5.5 Difficulty and Recency . . . 103
4.6 Conclusion . . . 106

5 Contact Grouping: Categorisation and Representation 110
5.1 Introduction . . . 110
5.2 Methodology . . . 113
5.2.1 Study Design . . . 113
5.2.2 Participants . . . 114
5.2.3 Procedure . . . 115
5.3 Results . . . 116
5.3.1 Repertory Grid Results . . . 116
5.3.1.1 Affinity Analysis . . . 116
5.3.1.2 Construct Categorisation Results . . . 121
5.3.1.3 Multiplex Constructs . . . 127
5.3.1.4 General Themes . . . 128
5.3.2 Contact Diagrams . . . 130
5.3.2.1 Diagram Types . . . 131

5.3.2.2 Groupings 144
5.3.2.3 Annotations 146
5.4 Discussion 148
5.4.1 Constructs and Diagrams 148
5.4.2 Methodology Drawbacks 152
5.5 Conclusion 153

6 Person-Centric Prototype Design 155
6.1 Introduction 155
6.2 Contact Extraction 157
6.3 Contact Grouping 159
6.3.1 Grouping Interface A: List Grouping 160
6.3.2 Grouping Interface B: Egocentric Grouping 162
6.4 Person-centric Interface for Information Tasks 164
6.4.1 Representation of Groups 164
6.4.1.1 List Presentation 165
6.4.1.2 Egocentric Presentation 167
6.4.1.3 Gravity and Group Link Presentation 169
6.4.2 Association of Information with a Person 173
6.4.3 Grouping Criteria & Multiplex Groupings 174
6.4.4 Relative Importance of Contacts 175
6.5 Prototype Development 176
6.6 Conclusion 178

7 Person-Centric Prototype Evaluation 181
7.1 Introduction 181
7.2 Group Construction 183
7.2.1 Group Construction Methodology 183
7.2.1.1 Study Design 183
7.2.1.2 Participants 185
7.2.1.3 Procedure 186
7.2.2 Group Construction Results 190
7.2.2.1 Adding Contacts 190

7.2.2.2 Group Creation 192
7.2.2.3 Group Content 198
7.2.2.4 Group Structure 208
7.2.2.5 Importance Rating 215
7.2.2.6 Group Satisfaction 216
7.2.2.7 Grouping Interface Design 218
7.3 Person-centric Interface Evaluation 219
7.3.1 Interface Evaluation Methodology 219
7.3.1.1 Study Design 219
7.3.1.2 Procedure 221
7.3.2 Interface Evaluation Results 224
7.3.2.1 Person-centric Interface Preferences 224
7.3.2.2 Gravity Interface 225
7.3.2.3 List 225
7.3.2.4 Group Link & Egocentric 226
7.3.2.5 Search Strategies 229
7.3.2.6 Restructuring Buttons 230
7.3.2.7 Filtering Buttons 232
7.3.2.8 Importance Filter 233
7.3.2.9 Sharing Tasks 234
7.3.2.10 Re-finding Tasks 236
7.4 Discussion 238
7.4.1 Group Structure 238
7.4.2 Contact Awareness 243
7.4.3 Group Confidence 245
7.4.4 Group Importance 246
7.4.5 Information Tasks 247
7.4.6 Person-Centric Interface Design 247
7.4.7 Limitations & Methodological Issues 249
7.5 Conclusion 251
7.5.1 Group Construction 251
7.5.2 Group Representation and Supporting Information Tasks . 253

8 Discussion **256**
8.1 Group construction . . . 258
8.1.1 Implications for Automatic Grouping . . . 258
8.1.2 Factors Influencing Group Construction . . . 260
8.2 Implications for Information Tasks . . . 267
8.2.1 Information Sharing Tasks . . . 267
8.2.2 Information Re-finding . . . 269
8.3 Person-Centric Design Guidelines . . . 270
8.4 Using a focus on 'person' to support information re-finding and sharing tasks . . . 273
8.4.1 Objective 1: To understand the process of information re-finding and identify the role of factors that enable or affect its completion. . . . 274
8.4.2 Objective 2: Having identified recall of links between people and information as an important, but underutilised part of re-finding search strategies in Objective 1, identify how to represent people in interfaces to support information sharing and re-finding tasks. . . . 277
8.4.3 Objective 3: Demonstrate and evaluate 'person-centric' interface approaches in providing novel support for information tasks, and identify design implications for their use based on the findings of Objective 2. . . . 281

9 Conclusion **285**
9.1 Introduction . . . 285
9.2 Research Contributions . . . 287
9.2.1 Study 1 – Re-finding Information . . . 287
9.2.2 Study 2 – Contact Grouping: Categorisation and Representation . . . 289
9.2.3 Study 3 – Person-Centric Prototype Evaluation . . . 292
9.3 Implications of this Research . . . 297
9.4 Limitations . . . 300
9.5 Future Work . . . 302

References **304**

Appendix A – Information Re-finding Diary Data Examples **317**

Appendix B – Extracts of Code from Prototype Development **323**

Appendix C – Ethics Application Example **334**

Appendix D – Coding of Interface Observations **351**

Appendix E – Study 3 Interface Evaluation Interview Questions **361**

List of Figures

1.1 Thesis Structure 9

2.1 ContactMap (Whittaker et al 2004) 29
2.2 ReGroup (Amershi et al, 2012) 37
2.3 CommentFlow - (Offenhuber & Donath, 2007) 42
2.4 Vizster – (Heer & Boyd, 2005) 43
2.5 PostHistory (Viégas et al, 2004) 46
2.6 Ripples – (Viégas et al, 2013) 47
2.7 Lexigraphs – (Donath et al, 2010) 49
2.8 Themail - Viégas et al (2006) 50

3.1 Example Drawing and Network Visualization of a Croatian Migrant – Lubbers et al. (2007) 61
3.2 Thesis Structure & Methods 68

4.1 Diary Application 74
4.2 Time since last encountering information. 86
4.3 Information sources used. 87
4.4 Perception of Task Difficulty 89
4.5 Information needs categorisation compared to Dearman et al (2008) 92
4.6 Information Tasks Categorisation 92
4.7 Prior knowledge of attributes of the information sought 93

5.1 Axis Diagram 133
5.2 Venn Diagram 135
5.3 Location Diagram 137

5.4 Spider Diagrams . . . 139
5.5 Grid Diagram . . . 141
5.6 Mixed Diagram . . . 143

6.1 Contact Clean-up and Importance Rating . . . 159
6.2 List Grouping Interface . . . 161
6.3 Egocentric Grouping Interface . . . 163
6.4 List Interface . . . 166
6.5 Egocentric Interface . . . 168
6.6 Group Link Interface . . . 170
6.7 Gravity Interface . . . 172
6.8 Association of Information with a Person . . . 174
6.9 Interface Selection Buttons, Restructuring Buttons, and Filter Buttons respectively . . . 175
6.10 Importance Filter . . . 175
6.11 Data Collection Process . . . 177

7.1 Contact Grouping Diagram . . . 191
7.2 Contact Adding Strategies – Sequential Strategy . . . 193
7.3 Contact Adding Strategies – Search Strategy . . . 194
7.4 Prior Group Creation . . . 199
7.5 Group Creation Strategies . . . 200

8.1 Contact Grouping Diagram . . . 262

9.1 (a) Vizster (Heer and Boyd, 2005) (b) Spider Diagram . . . 290
9.2 (a) Grid (b) Axis Diagrams . . . 291

3 Ethics Application Form 1 of 3 . . . 335
4 Ethics Application Form 2 of 3 . . . 336
5 Ethics Application Form 3 of 3 . . . 337
6 Study 1 – Ethics Application Amendment Form 1 of 2 . . . 339
7 Study 1 – Ethics Application Amendment Form 2 of 2 . . . 340
8 Study 1 – Participant Information Sheet . . . 342
9 Study 1 – Participant Consent Form . . . 343

10 Study 2 – Ethics Application Form 1 of 2 345
11 Study 2 – Ethics Application Form 2 of 2 346
12 Study 3 – Ethics Application Form 1 of 2 348
13 Study 3 – Ethics Application Form 2 of 3 349
14 Study 3 – Ethics Application Form 3 of 3 350

15 Participant L – Sequential Group Creation Coding Example . . . 352
16 Participant P – Non-sequential Group Creation Coding Example . 358

List of Tables

2.1 Overview of Attributes Used and Remembered for Information Re-finding 27

5.1 Repertory Grid Construct Results 123
5.2 Repertory Grid Most Common Themes 124
5.3 Table of Multiplex Constructs 128

7.1 Participant L Ego-centric condition coding output extract 197
7.2 Participant P List condition coding output extract 198
7.3 Prototype Grouping Categorisation 204
7.4 Ego and List Condition Group Quantities 206
7.5 Free Grouping Condition Standard Deviation of Group Contact Importance Ratings 215
7.6 Participant Ratings of Interface Preference 225

1 Participant L Ego-centric condition coding output 356
2 Participant P List condition coding output 360

Chapter 1

Introduction

Over the last decade or more both the amount of digital information generated as well as access to information has vastly increased (Carpenter, 2009; Lenhart et al., 2010). At the same there has been an expansion of the types and number of devices that people use to perform information tasks (Newman, 2008). This has contributed to making our personal information become ever more fragmented and distributed across devices (both within and between devices), applications, file types, and an array of Internet services (Jones and Anderson, 2011). Despite this, tools that support the way information is viewed and manipulated – such as email clients, file explorers, and instant messaging – in many cases have remained largely the same in their operation over the last decade or more, changing only slowly with much of the research on personal information management often failing to transfer to industry software (Kljun et al., 2013). As a result of the increasing diversity of information with which tasks are performed across tools (Katifori et al., 2008) it can be argued there is an increasing challenge in re-finding and managing the wide variety of information that users may be interested in. The increasing amount and diversity of information that people encounter formed the starting point for this thesis, which begins by asking what are the daily personal information tasks that people perform? What sorts of information do people need to perform these tasks? And how do people perform those tasks?

The term 'personal information' is used in this thesis to encompass all forms of information that might be considered personal to a person, as outlined by Jones (2008). These include all of the information owned by a person (e.g. documents created), about that person (e.g. address), directed to that person (e.g. emails received), sent by that person (e.g. instant messages sent), and previously experienced by that person (e.g. websites browsed). This description applies to the information used in performing information tasks discussed throughout this thesis. This broad definition is used in order to encompass both information that is generated and directly managed by a person, as well as information out of their direct influence, such as information on websites, which people may also require in performing personal information tasks.

By examining our daily personal information behaviours the intention of this research was to identify areas where greater support might be provided to people in performing every day information tasks with their personal information, and to identify potential solutions. Examination of previous literature identified information re-finding as a common personal information task, which could benefit from greater support (Jones and Anderson, 2011). Information re-finding tasks being those tasks where a piece of previously encountered information is searched for again. Commonly this might be re-finding particular web pages, emails, documents, and events (see Chapter 4). It is shown to be a distinctly different process from those undertaken in classic information search models (Capra et al., 2005) – due to its focus on memory rather than an exploration of a search space, as happens in information search tasks for new information.

Having identified information re-finding as a candidate task, detailed exploration of the process of information re-finding, for different task activities was then conducted in order to propose support options. Examination of information re-finding behaviours highlighted aspects such as the information source, time, keywords, and people as being prominent memories and critical cues for navigation that are recalled and used as part of the process of information re-finding. Of these, the recall of people in relation to personal information was identified as having the greatest potential to be used to improve support for personal information tasks. This element was found to be equally as prominently recalled as those

of time and keyword or subject related memories in the re-finding process, each of which have been demonstrated as being valuable means of interacting with information (e.g. Kules et al., 2010; Gemmell et al., 2002; Rekimoto, 1999), but also less readily usable to locate information. Exploiting an understanding of how people are recalled in relation to information for performing personal information tasks should have equal scope for providing support to personal information tasks, but compared to those other elements used in the re-finding process, using the connections between people and information is an area which is currently underexploited in interface design. The use of time and keywords as a means of ordering, or searching for information by contrast have been more widely researched and implemented into interface design (Fisher, 2004; Kules et al., 2010), for example in email clients, operating systems, and search engines.

Enabling the connections that are made between information and people to be used to a greater extent could therefore have the potential to improve the experience of performing personal information tasks, by providing another alternative means of accessing information. The connections people perceive might take multiple forms, including aspects such as memory of who shared or created a piece of information, people present when it was created or involved in its use, people the information is about, or future intentions for the information in regards to other people. This thesis argues that the use of connections made between people and information – such as the use of a person as an attribute linked to information or as a means of facilitating interaction – has value in supporting information re-finding and sharing tasks, and that a greater understanding of how to design interfaces focused on people can facilitate this.

Information is inherently social in nature, in that it is created for, by, and about people, and as this research demonstrates, it is remembered in terms of people. Yet interfaces rarely exploit this to any great extent. The lower degree of the use of links between people and information is likely in part due to the difficulty of capturing social information and relating it to files and pieces of information, as well as the difficulty for computers in interpreting social information. There has therefore been less research into the design of interfaces that are 'person-centric' – focusing on information related to people to support information tasks. How-

ever, with the continuing rise of ubiquitous computing technologies in everyday life (Marinagi et al., 2013; Khan et al., 2012), and the continuing integration of social networking websites into our communication and information gathering behaviours (Bakshy et al., 2012), the opportunities to record and use these connections between people and information more broadly, as well as the need to manage privacy in relation to people and entities, will increase. Understanding how to design an interface to make the most of the social nature of information in performing information tasks will allow this to be more fully exploited.

To explore the value of making greater use of connections between people and information for personal information tasks, and how to design for it, this research identified different forms of representation and ordering of people within interfaces for performing personal information tasks. In past research the most common means of facilitating the representation or organisation of contacts within interfaces to perform information tasks has been to group together contacts related to each other by a specific uniting criteria (Sawyer et al., 2012; Kelley et al., 2011; Mazzia et al., 2012), such as the number of ties between a group of contacts. This approach has been used to support information re-finding tasks, for example, relocating documents based on sets of people related to tasks being performed with those documents (Sawyer et al., 2012), however, more commonly it has been used to help support information sharing tasks. These are tasks where a piece of information needs to be shared with one or more other people, such as in an email, via a social networking post, or through face to face communication. Due to the prominence of 'person-centric' interfaces in the literature being used for information sharing tasks, this research also considers how information sharing tasks are supported by a 'person-centric' approach in addition to those of re-finding. As a task that has been more widely researched with respect to 'person-centric' interfaces, considering information sharing provides a greater degree of literature to inform the design and discussion of the research, and complements the use of information re-finding, being a task for which re-finding is performed for.

In order to understand more about how to represent contacts and contact groups to support personal information tasks this research considers how people perceive relationships and structures among their personal contacts to use as a basis

for guiding the design of 'person-centric' interfaces. A range of observation and knowledge elicitation methods are used to examine questions such as: how should contacts be ordered and grouped in interfaces to support information tasks, how should representations of contacts be structured, what constitutes a useful contact grouping, and what factors influence the decision making process in creating contact groups. Research into these questions provides a range of design concepts for 'person-centric' interfaces that are tested in this work to identify factors that should influence the design of interfaces for supporting personal information tasks through links between people and information, facilitated by contact grouping.

1.1 Aims and Objectives

The aim of this research is to explore how making explicit the links that exist between people and information can be used to organise and interact with information, in order to support users in performing personal information tasks. To this end it seeks to address how the use and design of 'person-centric' interfaces might support information re-finding and information sharing tasks.

Information re-finding tasks are defined as tasks where information has been previously encountered and needs to be located again as part of a wider task. This might be locating a document that is being worked on, locating an email sent yesterday or even several years ago, or finding a news article on a website. Information sharing tasks are tasks where a person has created, knows, or locates a piece of information and then communicates it to one or more other people. For example, this could be emailing a colleague a file, or sharing personal news with hundreds of people on a social media platform.

In both information re-finding and sharing tasks people are a prominent component of how both information re-finding and information sharing tasks are performed. In information re-finding tasks, a common link made between people and information is for the user to recall a specific person who sent or created piece of information. In information sharing tasks, information is considered in relation to particular people with whom the information is or may be shared.

This research considers how links such as these can be utilised to improve support for personal information tasks through interfaces designed with a focus on people.

In the case of information re-finding, being able to re-find information in terms of links to people can be beneficial as it is common for information to be recalled in terms of people (see Chapter 4). For example, a file might be required to be found and the user may remember it was a file sent to them by or authored by a particular colleague. In cases where this is the most effective means of locating a piece of information, for example, where it is the most strongly recalled attribute or where other useful attributes, such as keywords that might locate the file through a keyword search can not be recalled, having the ability to locate the information in terms of people would be beneficial. Creating more opportunities for those connections to be utilised and understanding how to design interfaces that support them has the potential to increase the number of successful re-finding tasks being performed, or to reduce the time taken to perform them, or the frustration of performing a proportion of information re-finding tasks. Similarly, in information sharing how that task is enabled by designing around people is also important for its completion. In information sharing tasks, how contacts are organised and structured to enable both locating, selecting, and recalling people to share information with is critical to facilitating the task. This provides a complementary task to consider in identifying what are valuable approaches for designing interfaces focused on people. This research therefore seeks to contribute a greater understanding of how to design to support people performing these common tasks through 'person-centric' interaction, organisation, and representations.

1.2 Objectives

To achieve this aim, a series of objectives emerged which sought to understand the processes that are undertaken to re-find information, and how groups of people are viewed as a means of performing re-finding and sharing tasks. That understanding of re-finding processes and representations of people was then used to guide the design of, and evaluate 'person-centred' approaches to supporting

information re-finding and sharing tasks. The objectives were as follows:

1. To understand the process of information re-finding and identify the role of factors that enable or affect its completion.
2. Having identified recall of links between people and information as an important, but underutilised part of re-finding search strategies in Objective 1, identify how to represent people in interfaces to support information sharing and re-finding tasks.
3. Demonstrate and evaluate 'person-centric' interface approaches in providing novel support for information tasks, and identify design implications for their use based on the findings of Objective 2.

This research at various points focuses on the context of email and social networks as an application area, such as in its implementation of design ideas, and elicitation of data about participants' contacts. These contexts were focused on as they are commonly accessed sources of information for the tasks being studied, that provided readily available data containing explicit links between information and people. As this research aimed to understand more about the design of 'person-centric' interfaces, this was preferred over implementing a novel form of social data capture to provide those links between people and information. This approach also provided greater scope for comparisons with the literature, much of which has also focused on these areas for similar reasons.

1.3 Thesis Overview & Structure

To address these objectives a literature review and three studies were conducted. Each study in turn was primarily focused on addressing one of these objectives. Study 1 focusses on understanding the process of re-finding information and demonstrated that recalling people is an important component of what is recalled about information when performing tasks with it. Study 2 looks at the relationships people perceive between contacts and how they choose to represent their contacts. Study 3 takes findings from the first two studies to design and evaluate a prototype system that uses people as its central point of interaction.

From the results of these studies factors driving and supporting information re-finding are identified, broad guidelines are generated that describe a range of factors to consider which have implications for the design of 'person-centric' interfaces, and structures and relationships for representing people in interfaces are identified.

Figure 1.1 (p. 9) outlines the structure of the thesis described in this overview. In the centre and left-hand side of the figure grey boxes represent individual chapters in the thesis. On the right hand side are the different methods applied as part of three different studies, each colour coded to indicate the study which generated the result – as delineated in the key in the bottom right. The links between the results sets and the chapters indicate that the results of applying a method contribute to the topic of that chapter either directly – in the chapter – or by informing the discussion of that topic in the final discussion in Chapter 8.

1.3.1 Chapter 2 – Literature Review

As an outcome of the research in Chapter 4, the use of attributes commonly recalled about or in relation to personal information, that can be used to organise and locate information are highlighted.

The research described in Chapter 4 highlighted how a number of different attributes are commonly recalled about or in relation to personal information, that are important for enabling the process information re-finding. These attributes include a subject or keyword, time, and persons. The literature review first considers time and keywords as means of organising and accessing personal information, and shows that they have been widely researched and understood as means of interacting with information. Common approaches to using these attributes as well as instances of specific novel uses of them are reviewed, demonstrating a substantial existing research body utilising them. This is in contrast with a lower degree of research into how links between people and information can be used to support personal information tasks, which is the focus of the rest of the literature review.

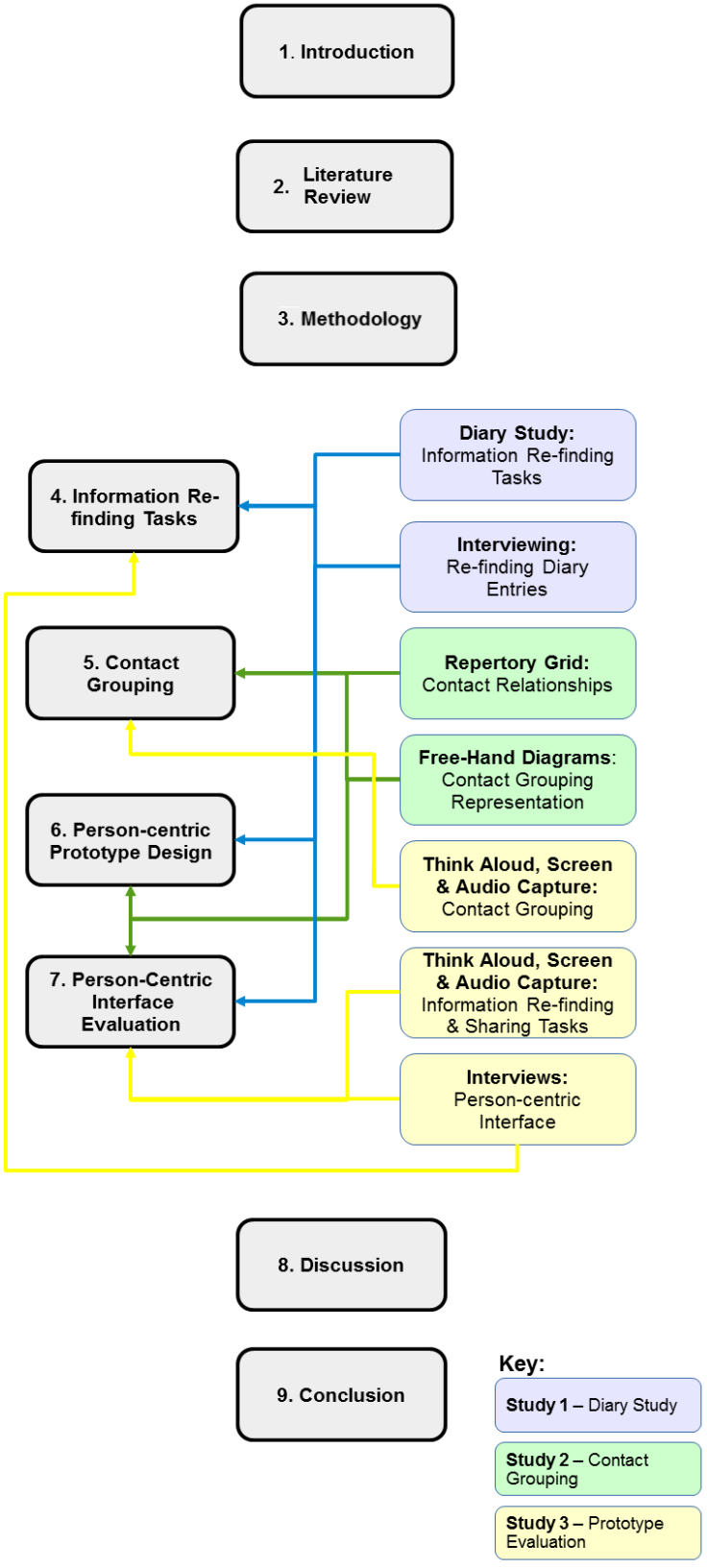


Figure 1.1: Thesis Structure

Specifically, information re-finding and information sharing are identified as two tasks where an improved understanding of how to design interfaces that utilise links between people and information could contribute to providing greater support to performing those tasks. These two tasks and approaches to performing them are reviewed, demonstrating how use of links between people and information in performing the tasks has merit as an approach to supporting those tasks. Previous work on information re-finding tasks is reviewed, with regards to information needs, and the re-finding process in terms of navigation and memory. Interaction with people in the form of personal contacts for information sharing, and the use and construction of contact groups to facilitate it, is also examined. Additionally, interfaces which facilitate these two tasks through the use of links between people and information are examined.

The review then goes on to consider research where interfaces use people as a primary, or prominent, unit of interaction. The different ways that interfaces that focus on the use of people for representation, visualisation, and interaction in the literature have been designed are reviewed. This identifies three different categories of representations in the literature. These are; interfaces and visualisations that seek to provide awareness of various communication behaviours in terms of the structure of communication and relations between contacts, the content of information exchanged with contacts, and those revealing activity with contacts.

The review concludes by highlighting that there is not an extensive body of research into the utilising links between information and people to support personal information tasks, while there is a wider body of research concerning the use of other attributes connected to supporting personal information tasks. In particular, it concludes there is a research gap in understanding how representations of people in interfaces can be designed to support how information re-finding and sharing tasks are performed. It also shows that there is a need to establish what factors are important for representing people in interfaces to perform information tasks effectively. Having a greater understanding of the representation of contacts in interface design will enable more effective interface designs to be created.

1.3.2 Chapter 3 – Methodology

The methodology chapter sets out the approach of the research carried out in the thesis and describes each of the methods that were applied to address the research aims outlined in this chapter. The background for each method is described and the reason for its selection are made clear. Each of the chosen methods are compared with methods that might also have plausibly been adopted as part of the justification for its selection.

1.3.3 Chapter 4 – Diary Study of Information Re-finding Behaviours (Study 1)

This first study examined how people behave with their personal information when performing information re-finding tasks. The aim of this study was to identify where greater support might be provided in performing information re-finding tasks. Participants were asked in the study to keep a diary of the occasions when they sought to re-find information over a period of two weeks (Palen and Salzman, 2002). This included information re-finding across all of their devices, and both work and personal contexts, in order to capture the full extent and range of re-finding behaviours that occurred. From the collected diary entries, the information needs, information tasks, task difficulty and the participant's memory of the information they sought were examined. The diary entries showed that when participants performed information tasks, such as locating a website, file, or email, they often recalled the source the information was contained within, a relative time it was last seen, its subject or a keyword, and other people related to the information that they were trying to find. These memories were used by participants as part of search strategies to navigate to the information they sought. However, the recall of people in relation to the information was less often able to be utilised as a strategy to re-find information compared to the other attributes. As recalling people was as common as the other attributes used to navigate to information, improving the opportunities for utilising connections between people and information in re-finding tasks is highlighted as a means to improve support for those tasks.

1.3.4 Chapter 5 – Contact Grouping: Categorisation and Representation (Study 2)

The second study explored what criteria and structures are important for the representation of people in interface design for performing personal information tasks. The number of contacts people tend to have – an average of 150 actively maintained relationships and an average total 214 contacts on social media (Ugander et al., 2011; Dunbar, 2016) – is large enough that displaying all of a person's contacts in an interface poses design challenges including whether all contacts are displayed, how to structure them, and how to order them. One of the most common approaches to representing people in the literature for performing information sharing and re-finding tasks has been by means of organising contacts into groups. This allows contacts to be managed and accessed for information tasks by chunking together contacts that are similar in one or more respects. The study utilised a form of repertory grid technique (Fransella and Bannister, 2004), using participants' contacts to elicit constructs that describe the different ways they perceive the relationships between people for contact grouping. From the repertory grid technique, an analytical framework describing the types of relationship/groupings encountered is described. In a second part of the study, participants created free-hand diagrams in order to examine how people choose to represent groups of contacts. From this five types of representation are identified, each emphasising different attributes and structures of contacts. The results of these activities were used to inform how interfaces could reflect the structures participants valued in order to describe and represent the relationships between contacts and groups of contacts.

1.3.5 Chapter 6 – Person-Centric Prototype Design

Following on from the results and analysis from the second study, a prototype person-centric system was developed, which incorporated different means to view contacts into a series of interfaces that were designed in terms of the representations and relationships from the contact grouping study in Chapter 5. The design of the prototype is described in three parts, which reflect three different stages of use. The first part describes how the prototype extracts contacts and per-

sonal information related to those contacts from the user's Facebook and email accounts, and then cleans up this information for use in the system's interface. The second part describes a contact grouping system used to generate contact groupings from the identified contacts. The third describes the design of 'person-centric' interfaces for performing information re-finding and sharing tasks. The features of the interfaces are described and the design decisions based on the second study are explained. The prototype's interface presents the contact groupings the participant previously made and provides a mechanism for accessing related information via the contacts, as well as sharing information with them. Four different interface designs based on the second study's free-hand diagram results are available to present the user's contacts, and interface controls which allow the groupings to be re-ordered according to criteria identified in the second study were also included in the design.

1.3.6 Chapter 7 – Person-Centric Prototype Evaluation (Study 3)

Chapter 7 describes the final study, which evaluated the 'person-centric' prototype described in the previous chapter. The study is broken up into two sections, which each describe one of two distinct tasks that were set as part of the study. In the first section, participants used the prototype to group their contacts using two different interfaces, which are compared. The grouping process is analysed via coding screen capture, think-aloud protocol, and interviews. Additionally, the grouping framework from Chapter 4 is used to describe the groupings created. In the second part, participants use their contact groupings to perform two types of tasks with each of four alternative person-centric interfaces. One half of the tasks involve information re-finding, and the other half involve information sharing. To draw out design implications for person-centric interfaces screen capture, think-aloud protocol, and interviews were again used.

1.3.7 Chapter 8 – Discussion

From the findings of the research reported in this thesis, factors that influence how contact groupings are generated were identified and broad guidelines for the design of 'person-centric' interface are proposed. The implications of the research are then discussed and put into context in relation to the literature described in Chapter 2. Finally, how the objectives outlined in this introduction were met is described.

1.3.8 Chapter 9 – Conclusion

In the conclusion the findings of this research are summarised and the contributions made to the literature are detailed. This is also summarised in Section 1.4 of this chapter. The implications of the research are highlighted with regards to application areas. These application areas included supporting personal information management within specific contexts of use such as operating systems, email, and social network websites and other information tasks such as managing privacy, projects management, and filtering content. The limitations of the research conducted are highlighted, and future research opportunities to build upon this work are identified.

1.4 Research Contribution

The research activities outlined in this thesis highlighted the role of recalling people in the information re-finding process and produced a range of outputs that contribute to understanding how 'person-centric' interfaces, which leverage the connections made between people and information, can be designed to support information re-finding and sharing tasks. These are summarised here and described in greater detail in Chapter 8.

Recording everyday re-finding behaviours (in Chapter 4) reinforced the results of past studies with regards to information re-finding events and the process of re-finding, and added to that work by examining re-finding across both work and home domains, as well as across digital devices. This showed that while

information re-finding tasks are not perceived as being as difficult to complete as some of the literature has implied, a small but still considerable number of tasks (10%) are considered explicitly difficult and could benefit from greater support. In order to describe the results from this study two new and one modified analytical frameworks for the analysis of re-finding events in context were developed. These described the what, why, how of information re-finding events. Applying this to the data highlighted that recalling other people related to the information sought was a prominent feature of the re-finding process, but one which has not been commonly designed for and has had only a limited amount of research exploring its use (Chapter 2).

In order to determine how people should be represented in interface design to support information sharing and re-finding tasks this research adapted repertory grid technique to provide a new means of eliciting relationships between people, in this case in the context of personal contacts. Analysis of the data generated from this resulted in the description of a categorisation which identifies the breadth of forms and ways people think about relationships between their contacts. This was done at both a high level – identifying the concepts the groups were describing – and at a lower level – highlighting the different forms of expression of a particular concept. In addition, distinct types of representations of contact groupings that people perceive in order to structure their contacts and make visible relationships between them were identified. These two findings contribute towards understanding how relationships are perceived and visualised, and provided a basis for identifying the design ideas that were developed into a prototype 'person-centric' interface.

Evaluating the prototype 'person-centric' interface enabled the decision making processes involved in group construction to be described. Understanding these processes was important for identifying factors that influence group construction (this is summarised in Figure 8.1). Understanding the group construction processes contributed to literature regarding how group creation should be supported for manual, semi-automated, and automated group creation and identification processes. In particular, understanding the requirements for what makes a group meaningful and useful for a user may contribute to improving accuracy of au-

tomatic detection of groups. Additionally examination of grouping behaviours also evidenced how interface design can impact on how users choose to construct user generated groups, and validated the categorisation of contact relationships described in Chapter 5 as being suitable for describing the make-up of contact groupings.

Evaluation of design ideas from Chapters 4 and 5 through the prototype confirmed the value of utilising connections made between people and information when re-finding information, as identified in Chapter 4. It also allowed factors that have implications for 'person-centric' interface design to be identified in relation to how contacts and groups should be represented. From the identified factors, high level guidelines for person-centric interface design were generated.

Chapter 2

Literature Review

2.1 Introduction

The research in this thesis primarily considers the design of interfaces that are broadly termed 'person-centric'. These are interfaces that have an approach to personal information management that focus on performing information tasks through views and organisations of people in relation to personal information. This might be information we have exchanged, received, viewed, created, as well as other facets such as our information behaviours with others. These interfaces have the potential to be valuable as much of the information we generate is created, for and in relation to, other people (Fisher and Dourish, 2004), and as such utilises the context in which information and relations between people are situated could help provide support in performing information tasks. However, the value of such an approach has not been clearly established and further work needs to take place to establish the extent of the usefulness and appropriate application of this approach. For example, to understand in what contexts the links between people and information might provide the most valuable support in performing information tasks, or how best to design for interfaces of this kind.

In the study described in Chapter 4 on information re-finding, three attributes are identified as being particularly important in the process of navigating to information. These were subject/keyword, relative time, and people linked to

the information. This literature review provides an overview of research using both time and keywords as attributes to support information tasks, and argues these have been researched and implemented to a greater degree than interfaces which make use of links between people and information. For the remainder of the review research into interfaces described as 'person-centric' are focussed on. This research has primarily centred on two tasks, those of information sharing and information re-finding, which are each considered in the review, as well as the different forms of design that 'person-centric' interfaces have taken.

2.2 Overview of the Use of Time and Keywords in Performing Information Tasks

Two common means of supporting the processes of organising personal information and re-finding information have been the use of time, to provide context to information in order to re-find it, and the identification of keywords in a subject, title, text, or other meta-data associated with a piece of information (see 2.1). These are useful ways of organising information, in part, because we typically remember to some extent at least one of these attributes of the information we are seeking. However, as is illustrated in Chapter 4, links between people and information are also typically remembered alongside these two other attributes.

Time and keywords are widely used a means of locating and ordering information in interfaces, and there has been a variety of research into interfaces designed to make use of these two means of supporting information re-finding. Systems that support information retrieval via the use of keywords, in particular, have been extensively researched within literature related to information search, leading to a variety of approaches to interface design and interaction. Kules et al. (2010) review a range of systems supporting information search tasks and describe a taxonomy of the approaches taken, which classifies the interfaces and describes the level at which the interface is providing support to search tasks, the depth of evaluation of the interface, and its prevalence online. Their classification consisted of the following types of interface:

- Hierarchical Classification – One of the most well known examples of this approach was Yahoo! Directory (Liu et al., 2005), which allowed users to navigate to relevant information through a hierarchy of increasingly specific subject areas, within which links to associated websites were listed. The high associated cost of manually updating and maintaining the classification is a drawback to employing this approach.
- Automatic Clustering – This approach attempts to automatically identify the pertinent themes (clusters) within matching datasets to search queries. The prominent themes then provide a means of navigation through the results of a search domain to pertinent information. This then allows navigation in a similar manner to Hierarchical classification, as in Rivadeneira and Bederson (2003)'s study, but allows much greater scalability due to the automation.
- Faceted Classification – This has been a popular approach for commercial websites (Dawson et al., 2006), for example, the online auction website eBay is a well known example of a faceted search approach. In faceted classification the results of the search can be refined by applying constraints on the search along particular dimensions or facets, which in the case of eBay could be new or used items, location, price etc..
- Social Classification – This makes use of often large numbers of people contributing their own personal classifications to resources in a dataset, which are then available for all other users to make use of as a means of information retrieval (Zubiaga et al., 2013). For example the Goodreads website encourages users to classify books they read into genres or themes, all users of the website then benefit from the collective classification, using the user generated classifications as a means to search for books of interest.
- Result List – This is the approach taken by many popular web search engines such as Google, whereby a list of matching results are returned in response to a keyword, ordered against a criteria to determine their relevancy. One area research within this, for example, is how interpreting user interaction with the results list can be used to increase relevancy of the

returned results by re-ranking them (Kim et al., 2013).

- 2D & 3D Result Representations – There has been a wide variety of research examining how visualisation research can be applied to the results of keyword searches. This approach typically involves graphical representations of the structure of a search space, e.g. linked or related content connected or clustered together, examination of which can lead to identifying relevant information. Risden et al. (2000), for example, describe a comparison of both 2D and 3D visualisations for accessing web content.

Time has also been well understood as a useful means to order and retrieve information, most obviously as a way to order information from most recent to least recent.Fisher (2004) highlights a range of research which make use of temporal structures to aid everyday collaboration, and presents the design of a system that visualises patterns of interaction with contacts. Other work such as André et al. (2007)'s Continuum has focused on the design of visualisations using time as a central component. Continuum, for example, presents a means of temporally exploring large hierarchically nested datasets.

One of the largest bodies of work focusing on the use of time has made use of its relationship to episodic memory. Episodic memory (Tulving, 2002), in which time provides order to the memory of autobiographical events, which includes much of the surrounding context to an event, has been a particularly useful concept for the retrieval of information. Through this mechanism various research projects have used time as an important cue for retrieving information using the associations between information and the surrounding context of events. Gemmell et al. (2002)'s MyLifeBits, and other similar systems, in particular have taken this concept to an extreme through research on life-logging. Life-logging attempts to record all daily interactions, often by means of logging all digital interactions, as well as augmenting this information via wearable devices that record data such as location and images of the users travels and interactions. From this it records more of the contextual memories that might be recalled and linked with information, which can then be used as part of the process of re-finding information. Ringel et al. (2003)'s Milestones interface took a similar approach to this, providing a unified access point to all the documents, emails, and web pages the

user encountered, and then presenting these along a timeline. This timeline was then enriched with contextualising events from both the user's personal calendar and photos, and public events and news. The aim was to aid users in locating documents using memories contemporary to the information sought. Rekimoto (1999) looked to support a 'Time-centric' information environment, in which its users could return to the information state of the computer at any given time in the past. Krishnan and Jones (2005) developed TimeSpace which took a more nuanced task-based approach to relating files to time by associating files with specific tasks or projects within timelines to provide a greater level of context. Taking this approach allowed greater integration of time as a point of organisation with the tasks people were performing, this allowed recollections of both currently active and past sessions of performing tasks to be used as means of recalling information and making the documents related to them retrievable.

While there are some common examples of interfaces that concentrate on people in their interaction, these are for the most part related to messaging and video conferencing. The use of people in these interfaces is typically as a functional necessity, only serving the purpose of allowing direct contact with a person, rather than using that information to support a task, such as re-finding information. Communications systems, such as email and social networking websites, help to facilitate a variety of information tasks in relation to both individuals and groups of people. These systems are typically message-centric in nature, focused on the process of message exchange, helping us to process the most recent incoming information. This is a useful design for many of the tasks we perform where we are responding to recent information, but often ignores – or does not fully leverage – the importance of the contacts involved in the exchange for performing other forms of information task in relation to the information, and the complex social relations within which the tasks we are performing are situated.

The following part of this chapter identifies information re-finding and sharing tasks as areas where a greater use of *person* as part of the interface may be beneficial, and reviews research describing approaches to supporting those tasks through the use of people as a major component of the interface. And in the final part of this chapter research on approaches to the representation of people

in relation to data is reviewed.

2.3 Supporting Information Tasks

Where research has focused on using people as a central component for supporting specific information tasks, it has largely involved two distinct areas of research. Firstly in re-finding information, which has focused particularly on email, due to the availability of data, due to people being recorded alongside the information as meta-data. Or secondly sharing information, or in order to filter information or control access to information. This has primarily been driven by research into social networking websites, such as Facebook, as means to address privacy issues (Mazzia et al., 2012) or share / filter content with / from groups of people (Amershi et al., 2012). These two information tasks and the approaches to supporting them through use of people as a central component of the interface are described in the following sub-sections.

2.3.1 Information Re-Finding

2.3.1.1 Introduction

It has been suggested that there is a case for greater integration of information (Jones and Anderson, 2011) in response to the challenges of fragmentation and volume of information, and that through integration easier access and efficient retrieval of information might be achieved. Over the years a variety of desktop tools have been developed that offer a more unified point of access for the management of personal information. These have included everything from search tools, such as Stuff I've Seen (Dumais et al., 2003) and Google Desktop, to semantic desktop tools mapping the user's mental models of personal information, such as Haystack (Quan et al., 2003) and Gnowsis (Sauermann et al., 2005). Another strand of work has harnessed the concept of life-logging where all daily interactions are recorded in order to retrieve useful information through episodic memory, this has produced tools such as MyLifeBits (Gemmell et al., 2002) and YouPivot (Hailpern et al., 2011). For the most part, such systems have received only piecemeal implementations, and have seen only limited uptake. Lack of

adoption may be due to users simply preferring their existing methods of accessing information, even if there are quicker alternatives (Bergman et al., 2008; Jones et al., 2004), or a desire to view information within the context it was encountered in. Another possible explanation, suggested in this chapter, is that it may also be due to a lack of understanding of the considerable differences that exist between the activities of information finding and information re-finding, differences that have to some extent been overlooked.

While the process of information seeking has been extensively researched and modelled by researchers such as Wilson (1997), particularly with regards to information retrieval behaviours, re-finding as a distinct process in its own right has received much less attention. Information retrieval as part of information seeking is often represented as an iterative model of search and refinement until the searcher believes their information need to have been satisfied (Belkin et al., 1993). This process involving elements such as iteration, judgements about retrieved information, adaptation of information needs and understanding of the search area is less relevant in the case of re-finding, where the information sought is already known to exist and iteration need only occur as process of navigation rather than exploration. Knowing that the data being sought is certain to exist causes people to undertake markedly different processes to locate that information. The processes involved in re-finding are more directed. They depend on memory and context to locate information to a greater degree than the more open-ended search processes involved in locating information for the first time (Capra et al., 2005).

2.3.1.2 Information Re-finding Behaviours

Despite being an extremely common activity, re-finding has only received limited attention - research by Tyler and Teevan (2010) and Teevan et al. (2007) has suggested figures for repeated queries in web search occur somewhere between 22% and 39% respectively. Similarly Broder (2002) in his taxonomy of web searches found 24% of search queries were to websites participants already had in mind, which he classified as *Navigational* queries. Outside of web search, re-finding is also undoubtedly a significant behaviour across many other mediums

of information, particularly email, and our various folder systems, which many people will often organise actively for the purpose of re-finding (e.g. Whittaker and Sidner, 1996).

In order to examine the nature of information re-finding it is necessary to understand typical information re-finding tasks and behaviours. Previous research examining natural information behaviours has focused on the information needs that cause people to search for information (Sohn et al., 2008; Dearman et al., 2008), and the processes people go through to locate the information they desire (Capra and Pérez-Quiñones, 2003; Teevan et al., 2004). Much of this research has made use of diary studies to collect representative instances of when and what information is sought, and used this to define categorisations of types of information needs and to then illustrate their frequency. For example, Sohn et al. (2008) focus on capturing information needs that arise while mobile, or Dearman et al. (2008) who take a broader approach, revealing a cross-section of information behaviour across all daily activities and locations. In both cases, structured diary entries recorded occasions when participants needed to locate information, which were then categorised in the analysis by grouping entries together into distinct mutually exclusive groupings.

In a similar vein to the diary study methodology, Teevan et al. (2004) focused on the varying behaviours used in the process of finding information. Rather than rely on participants to record diary entries, they initiated the recording themselves by interrupting participants in their offices, unexpectedly, on a daily basis. They conducted 5-minute interviews in which they asked them to describe what it was that they had most recently looked for in their email, files, and on the Internet. Through this method they identified and discussed two processes through which people navigate to their personal information, *Orienteering* and *Teleporting*. *Orienteering* described the behaviour of navigating in a series of small steps to incrementally close in on the information being targeted, usually relying on prior knowledge and contextual information to make each step closer. In contrast the behaviour of *Teleporting* described participants attempting to navigate directly to information, for example, by an explicit search engine query hoping to retrieve the exact page desired.

Capra and Pérez-Quiñones (2003) also describe a process of people re-finding information that shares similarities to *Orienteering* in their study of how people re-find found things. They explored re-finding behaviours by setting information search tasks on the web that people would perform, both of a specific nature, to find a particular website, and of an exploratory nature, such as looking for sweatshirts you might want to buy. They then at a later date they asked participants to describe over the phone to another participant how to locate the page from memory. From this they noted that a two-stage approach to information re-finding was used in most cases, whereby people first concentrate on locating the source of information, by searching, and then proceed to navigate by browsing until the specific information was located.

The desire to browse to information was also noted by Boardman and Sasse (2004) who observed a strong preference in their study of cross-tool filing and retrieval behaviours for people to browse to information rather than to search for it, stating that search was often viewed as a '*last resort*' when re-finding information. An observation which was also reported by Bergman et al. (2008) studying the effects of improved search on re-finding. They describe browsing behaviours as consisting of location-based browsing, such as navigation through folders or locating an icon on the desktop, and sorting /scanning items ordered by meta-data e.g. date, with people using a combination of both to satisfy their needs. They note that differences occur across different tools, for example, email was often more focused on sorting /scanning, and that search was used to a greater extent within that context. While it appears that there are different broad tendencies that exist within specific contexts, a general preference towards navigation over search seems to exist within all re-finding behaviour (Capra and Pérez-Quiñones, 2003; Boardman and Sasse, 2004; Teevan et al., 2004; Bergman et al., 2008).

Underpinning these navigational processes, undertaken to satisfy information needs when re-finding information, is the question of what people remember about their information. Exactly what people remember when searching for information has only been explored to a limited extent, with Capra et al. (2005), Elsweiler and Ruthven (2007), and Blanc-Brude and Scapin (2007) covering three of the main

areas of information re-finding: web, email, and local system files respectively. This is arguably an important aspect for understanding re-finding behaviour as it will influence how people undertake the process of re-finding. From Capra and Pérez-Quiñones (2003)'s telephone task they observed that people would remember and use three central 'waypoints' to navigate to the information they were re-finding, these were *Page/Site Titles*, *URLs*, and *Page Descriptions*. In contrast Elsweiler and Ruthven (2007) examined a different store of much of our personal information by focusing on email re-finding. After each email search task the participants performed they were asked whether they could remember the following attributes about the email: when it was sent, the sender, the topic, the reason for it being sent, other recipients of the email, any attachments to it, images contained within it, and any links within it. Blanc-Brude and Scapin (2007) approached this by using a combination of these approaches, using both an open '*free-recall*' of attributes by participants, and a list of eleven attributes derived from the literature, which participants might remember. These were: location, type or format, file name, title, size in terms of pages, time of last accessing the document, keywords, visual elements, associated events, documents linked to the target document, and actions performed on the document such as printing. Table 2.1 gives an overview of the results of these studies, illustrating the extent to which these memory cues were remembered or used for re-finding.

2.3.1.3 Information Re-finding Interfaces

The most notable research attempt to make people a primary unit of interaction was Whittaker et al.'s (2004) ContactMap system (see Figure 2.1). ContactMap is 'a personal social desktop', which provides a 'person-centric' representation of the people who are important in a person's work and social life. Where accessing and sending emails is mediated through a visual representation of people. Contacts are indicated by large images, their layout determined by the user spatially placing the contacts where they wished, and sorting the contacts into various groups that were meaningful to them. These were typically reflective of their personal relationships and were largely based on affiliations, work projects, and social categories, rather than ties of varying strength. The contact icons provided

Table 2.1: Overview of Attributes Used and Remembered for Information Re-finding

Files: Blanc-Brude & Scapin (2007)	Recall % Correct / Partial / False	**Webpages:** Capra & Perez-Quinones (2003)	% Use for Re-finding	**Email:** Elsweiler (2007)	Recall %
Location	36% / 60% / 4%	Page / Site Titles	17.8%	Date Sent	57.5%
Type or format	93.3% / 3.3% / 3.3%	URLs	37.8%	Sender	77.3%
File Name	25% / 66.7% / 8.3%	Page Descriptions	44.4%	Topic	85.1%
Title	33.3% / 46.7% / 20%			Reason for Email	80.5%
No. Pages (Size)	15.4% / 30.8% / 53.8%			Attachments	12.77%
Time Accessed	4.8% / 47.6% / 47.6%			Images	2.13%
Keywords	32% / 68% / 0%			Links	21.81%
Visual Elements	76.2% / 23.8% / 0%			Recipients	46.8%
Events	100% (Any Recall)				
Related Docs	90% (Any Recall)				
Actions	80% (Any Recall)				

visual cues indicating aspects such as new messages and pending commitments to people, while interacting with the contact icons allows the user to access and send associated communications with them as well as view reminder notes. Figure 2.1 shows the main screen of the ContactMap interface with contacts sorted into groups of related contacts, as the user saw fit. Information related to those contacts could be accessed by clicking them. As in Vizster (Heer and Boyd, 2005) colour has been used to denote related contacts within a space.



Figure 2.1: ContactMap (Whittaker et al 2004)

Whittaker et al. (2004, p.447) state that message-centric interfaces '*...do not represent the relationships between co-workers that are useful in the associative retrieval processes characteristic of social data mining.*' They propose ContactMap (Figure 2.1) as a system that aims to support the processes of social reminding and social data mining that are normally a by-product of physical proximity to people, that are now missing in our increasingly distributed online lives, and inadequately served by message-centric interfaces. And, indeed, the results of their studies did indicate that people preferred ContactMap as a system over traditional email interfaces when accessing information related to those contacts and performing social tasks. These were social remembering tasks, such as remembering to honour communication commitments and remembering to keep in touch, or social mining tasks such as exploiting contacts and keeping track of projects.

As was expected, ContactMap was found to be much less useful for reading, replying to, and composing messages, activities that are the strength of message-centric interfaces. And as Elsweiler and Ruthven (2007) points out, this approach has some drawbacks in that not everyone works in only a person and project focused manner, and it also forces people to remember information in terms of who sent it, which while common, is not always the way in which information is remembered. Ultimately, Whittaker et al suggest interfaces such ContactMap should not seek to replace message-centric interfaces, but rather provide an alternative view of one's personal data that takes into consideration social relationships as one of a number of views that people might use, depending on the tasks they need to perform.

One approach to complement the message-centric approach of email clients is to integrate people into them more actively. In Gruen et al.'s (2004) reMail prototypes their 'correspondent map', highlighted in a side tab the people with whom the user corresponded most frequently over email. The messages of these frequent contacts were collected together and the contacts grouped by their affiliation to an organisation. Colours indicated the age of the most recent message from a contact and whether a message had been responded to. Gruen et al. (2004) found that their correspondent map was interesting to the participants, but used only

to a small extent in practice, in comparison to the other features they employed to aid with email tasks. They suggest this may have been in part due to unfamiliarity with the workings of the visualisation. They go on to say that simpler visualisations, which enrich the context of already existing items may be a better approach to provide useful contextual information. This is also the approach taken by Fisher et al. (2007), whose participants found person-based organisation of information, with their filtering strategy that was integrated with a more typical email client, to be a useful approach.

Neustaedter et al. (2005) found that the cue of seeing a person's name in email was central to judgements of importance. When examining new unread email people look first for the most important messages, where this perceived importance is linked to the person who sent the message. In Fisher et al.'s (2007) SNARF system the goal was to support the practice of email triage by providing support for a variety of strategies for increasing the relevance of email messages displayed, by filtering out irrelevant ones. This was achieved by filtering emails based on criteria such as people with whom they frequently interacted, whether the emails have been read, and how it is addressed e.g. just to the user or to others. As part of Themail (Viégas et al., 2006) contacts were displayed to the user based on their history of interaction. The most frequent contacts were displayed to the user and were selectable as a means of filtering messages.

'Person-centric' interfaces have predominantly been built on data from email and social network websites, due to their wide adoption and embedded use within organisations and personal life, and readily available meta-data about people in networks of relevant contacts to the tasks being performed with them. Sawyer et al. (2012), however, provide an example of how people, and groups of people, can be used beyond the confines of email. They do this by utilising relationships based upon physical proximity, not only to other people, but also information. This is achieved by recording the physical and social interactions between people and information. They have created a system which enables searching of email, system files, and web pages based upon the physical-social context. People were placed into social groupings, which may overlap, by participants. The system then enabled re-finding by tagging documents as being associated with particular

people and the social groups based on their physical presence when a document was created or edited. This could be done either manually, via a presentation of individuals and the created social groups, or automatically via sensors detecting presence of people. This enabled later retrieval using the memory of people or social groupings associated with documents.

2.3.2 Information Sharing and Contact Grouping

Interfaces that support information sharing, or facilitate access to information through associations between people and information, require a means of representing and organising those people in a form that is meaningful for the user. As it is commonplace for people to have a large number of contacts, not all of their contacts can be displayed simultaneously in any usable form. As a result, a manner of grouping together or filtering contacts is typically utilised; this enables the number of contacts shown to be reduced to a manageable level. For information sharing this is valuable in that it can provide the ability to perform actions across large numbers of people, locate specific people more effectively, and enables people to be recalled by means other than their name.

Grouping contacts together allows the representation of the everyday social structures people perceive. We often communicate to groups of people and need to contact members belonging to particular groups of people, and we also associate information with particular people and groups of people. As (Fisher and Dourish, 2004) point out, many tasks carried out by single user tools are collaborative in nature. Documents are created in relation to activities and information provided by others, designed with others in mind, or to coordinate others. This is what Fisher and Dourish describe as 'everyday collaboration'. As a result of this there is also often information that is associated with and holds shared relevance to all members of a group. Representing people in groups also provides another dimension through which to associate, remember and access people.

Within various social networks people are often encouraged to some extent to group their contacts. Social networks, such as Facebook and Google+, encourage the grouping of people as being central to the process of controlling privacy and

filtering content – particularly within Google+ and its concept of 'social circles'. Research exploring how people choose to group their social network contacts for, the purpose of sharing or filtering content, reveals the range of forms of distinct groups people perceive (e.g. Kelley et al., 2011), and might be aided in creating to manage their contacts. Organising the many varied types of social groupings a person categorises people into requires eliciting information about a person's complex social network and representing it in an interface. As these groupings vary greatly from person to person, this creates a challenge in that any technique used to elicit and display them would need to be representative enough of their mental model of their relationships to be recognisable and useful (Jones and O'Neill, 2010), whilst also not being overly labour intensive as to deter people from grouping the contacts (Whittaker et al., 2004). Whether for sharing or re-finding, the literature makes use of a mixture of four broad forms of group creation in order to address these challenges: groups created based on the behaviour of the user and/or their contacts, manually created by the user, automatically detected groups, and a mixture thereof. These variances in approach reflect a variety of trade-offs in the need to make groups quick to create, but also accurate.

Kelley et al. (2011) found that participants mostly grouped people based on when or where they met them. In a card sorting task, they found that 89% of participants[1] created a group of people from school or university, with 76% of participants specifically referring to university. Within this categorisation of university groups, they also identify university clubs and groups as being present for 33% of people in the study. Family related groups were also very common, being present in 63% of participant's groupings. Locations other than school or university were created by 30% of participants, and 30% of participants also had a group for miscellaneous people. They suggest that the place where a person is first met can be a deciding factor in determining how to resolve which group to place a person in, when an overlap occurred between two groups. Jones and O'Neill (2010) similarly performed a card sorting task using cards made

[1]Participants in the study were in part recruited through a university website, which possibly explains part of the emphasis on university. This did not represent the whole sample, however, with other participants being drawn from advertising via Craigslist and fliers around the local city.

up of their participant's Facebook contacts, who were then asked to sort their contacts by grouping them as they would want to for controlling their privacy. They identified the following categories of groupings present in the groupings of participants: Social Circles & Cliques (87%), Tie Strength (73%), Geographical Locations (67%), Organisational Boundaries (60%), Temporal Episodes (40%), and Functional Roles (40%).

Separate to manual grouping tasks are groups identified algorithmically, generated from social networks graphs similar to those described in the network focus section of the chapter (see 2.4.1.1). Research in this area has been put to a number of uses, but significantly it has been used to aid people to create groups for use in social networks. People can often be apathetic towards manually grouping people as it can take considerable time and effort (Whittaker et al., 2004). While the accuracy of automatic detection is often variable, MacLean et al. (2011) argue that the value of automated group detection lies within people's inability to update and recreate their social structures manually, due to the volume of people and complexity of the structures' nuances. There are a variety of broad approaches to the algorithms used in solving this problem; groups might be identified based on communication behaviours, mutual friends, communication and co-presence in tagged photos or events (e.g. MacLean et al., 2011), communication and profile information (e.g. Leskovec and Mcauley, 2012), and a combination of all of these elements (e.g. Amershi et al., 2012).

Jones and O'Neill (2010) went on to compare their participant's card sorting to automatically generated groupings based on network structure. This approach produced wildly varying correlations from between 18.1% to 79.5%. They note that this can be explained by how clearly defined a person's social network is. Those who have very clearly defined clusters of contacts would group people in a similar manner to the structure of their social network ties, while those who had less distinct clusters of contacts had to find different ways to differentiate sections of their network. In many cases this was done based on tie strength, splitting what might be one group into two sub-groups of closer contacts, and those less close contacts. When Jones and O'Neill (2010) combined these subdivided groups from the card sorting the level of correlation sometimes doubled. They suggest

that factoring aspects such as tie strength into algorithms will therefore lead to automatically generated groupings that reflect more accurately participants' own perceptions about their social networks

Between the extremes of manual and automatic group detection lie interfaces for group suggestion, systems such as Social Flows MacLean et al. (2011) and Re-Group (Amershi et al., 2012). These systems seek to act as intermediate points between manual and fully automatic creation of groups, aiming to generate groups automatically as suggestions for users, which act as a starting points, from which people modify and create final groups as necessary. In Social Flows, MacLean et al. (2011) identify common groupings of people in email based on communication patterns and Facebook photo tags of people. They display the groupings ordered dependant on the total number of emails to the individuals within the group and the similarity of membership between those groups, while common subsets of super-groups are displayed beneath super-groups in a hierarchy. To prevent large numbers of less relevant or less representative subsets - in a similar way to the previously discussed grouping problems of Jones and O'Neill (2010) - they merge highly overlapping groupings. However, they go to note the importance of membership of multiple groups. These commonly used groupings of people then act as starting point for the user to rename, merge, and remove groups, as well as add or remove people.

Regroup (Amershi et al., 2012) works in the reverse manner to Social Flows, waiting to see who users add to a new group and then suggesting automatically detected groups of contacts to add the new group based on commonalities with the current group members. In order to identify relevant types of groups to suggest to users they performed a survey garnering 69 responses and identifying 244 custom made groups, predominantly from Google+. From this they identified the following features which could be used for group suggestions from available social network data:

- Gender, Age Range
- Family Member
- Home City, State and Country

- Current City, State, Country
- High School, College, Graduate School Workplace
- Amount of Correspondence
- Recency of Correspondence
- Friendship Duration
- Number of Mutual Friends
- Amount Seen Together Description/Examples

Figure 2.2 shows the ReGroup interface (Amershi et al., 2012) as an example of a contact grouping interface. On the top of the interface are selected contacts, to the left possible useful criteria as groupings of people to filter down contacts by, and on the right the results of the contact filtering process as a grid of matching contacts.



Figure 2.2: ReGroup (Amershi et al, 2012)

2.4 Representation of People & Information Behaviour

Research that has been conducted to elicit information behaviours in relation to other people has broadly taken two forms. One is social network based, primarily focused on mapping our behaviours with others onto a network structure. The other is activity based, focused on presenting the levels of information and the nature of the content of the information we exchange with others. The two categories are not without overlap in their designs, but the separation of the two serves to highlight differences in the underlying premise. The first seeks to present the structure of our information behaviour while the latter to present the extent and nature of our activity. Much of the work conducted in this area, identified by this literature review, has been created by Judith Donath's MIT Sociable Media Group, which examined a variety of novel methods (Donath et al., 2010; Gruen et al., 2004; Offenhuber and Donath, 2007) for visualising social data – many of which focused on people in their interaction and presentation - intended to aid the exploration and understanding of social data.

2.4.1 Structure

2.4.1.1 Network Focus

Used since the 1930s (following the work of Jacob Moreno), social network analysis has provided a means to visually map social networks. Social network analysis reveals the structure of our social networks, analysing the patterns of interdependencies between actors in order to draw conclusions about consequent behaviours that emerge from those patterns (Papachristos, 2011). This approach has been used across a wide variety of disciplines to research the impact of our social networks. For example, Nicholas et al. (2007) have used it to understand how rates of obesity are influenced by social distance to other obese people within a social network, while Papachristos (2011) has examined the structure of criminal organisations and how relationships within them can be used within criminology to understand gang behaviour. Visual representations of personal networks have also been common within counselling psychology and social work (McCarty

et al., 2007) through approaches such as Hierarchical Mapping Technique and Genograms, where an individual creates a map of their own perception of their social network.

Forms of visually mapping the structure of our contacts from our personal data are one area where a variety of interfaces focused on people have been created. Revealing the structure and patterns of our social networks, Social Network Fragments (Viégas et al., 2004) was one of the first of a series of interfaces by the MIT Sociable Media group to present people with the patterns of their connections, latent within their personal data. It focused on the evolution of the connections between people and the user representing the emergence and dissolution of relationships between people's contacts over time. A graph was created using the names of email contacts as nodes, linked together with lines where communication had occurred between email contacts.

Similarly Fisher and Dourish's (2004) Soylent system attempted to not only examine the structure of people's contacts, but also how the structure changes over time. Much of the research in this area either notes as a problem or simply ignores how members of groups of people can be temporally sensitive, their membership changing over time or the relative importance of those members or groups shifting. They found that that there were certain structures and relationships that were recurrent within people's data. They were able to identify distinct topologies of interactions, for example, when examining the strength of connections. They could identify that certain people had important roles, and were involved in a range of tasks and contexts, while some individuals acted as bridges between the different groups of people. Looking at these over time they could see that groups had core members and peripheral members, that people would become more or less involved in groups, and the groups themselves might appear and fade away. They used top ten lists to reveal to users their changing communication patterns, displaying their most frequent correspondents for the previous month and week. Interviewing participants revealed that people could relate these to specific stories or events, such as the beginning and ends of collaboration, joining a new team, or the arrival or departure of transient contacts such as temporary workers, in contrast to those who remained consistently on the lists over a longer period of

time.

Offenhuber and Donath's (2007) Comment Flow (Figure 2.3) system approached the concept of viewing changes in our digital social network from a slightly different perspective. Rather than examining the changing face of contacts over time, they exposed the communications themselves over time. Offenhuber and Donath (2007) visualised data from the MySpace social network website, and again this took the form of a force directed graph, but here the people were represented by images as nodes instead of names and the links in the graph were used to illustrate social activity. They argued that in order to gain an understanding of the true structure of a person's digital social network, that links alone are insufficient to describe it. They suggest that seeing how information flows along a network, between people, provides a much more accurate picture of the social network than the linked structure alone can provide. In their visualisation the links forming the structure of the graph and representing the social relations were also used as a basis for visualising the dynamic flow of social activity in the form of comments and images exchanged between people. Messages and images are visualised moving along the links giving a sense of the amounts and areas of activity between people in their social network at a given time.

Heer and Boyd's (2005) Vizster (Figure 2.4) again utilised a social network graph as part of its interface, but focused on facets of social behaviour unrelated to the activity of the contacts, instead focusing on the attributes of the contacts themselves. Vizster created its social graph using the Friendster social networking website, displaying connections between people an individual knew, as well as people they did not know - but were connected to people whom they did know. Each of the people in the graph provided a means to access a variety of profile information for each person, such as their interests and biographical details. While connections could be highlighted and explored in a variety of ways, such as highlighting direct connections between an individual, or automatically detecting and revealing groups of people. Heer and Boyd (2005) state that the interface helps users to '*facilitate discovery and increased awareness of their online community.*' They go on to say '*we wanted to support the exploratory and playful aspects of Friendster while also giving users easier access to search and group*

patterns'.

Figures 2.3 and 2.4 show example images of the network focused interfaces of CommentFlow (Offenhuber and Donath, 2007) and Vizster (Heer and Boyd, 2005). Both images illustrate the network of contacts used in these interfaces and the additional meta information graphically communicated about the members of those networks. In the case of CommentFlow communications along the networks can be seen and in the case of Vizster overlays of different groupings within the network are visible.

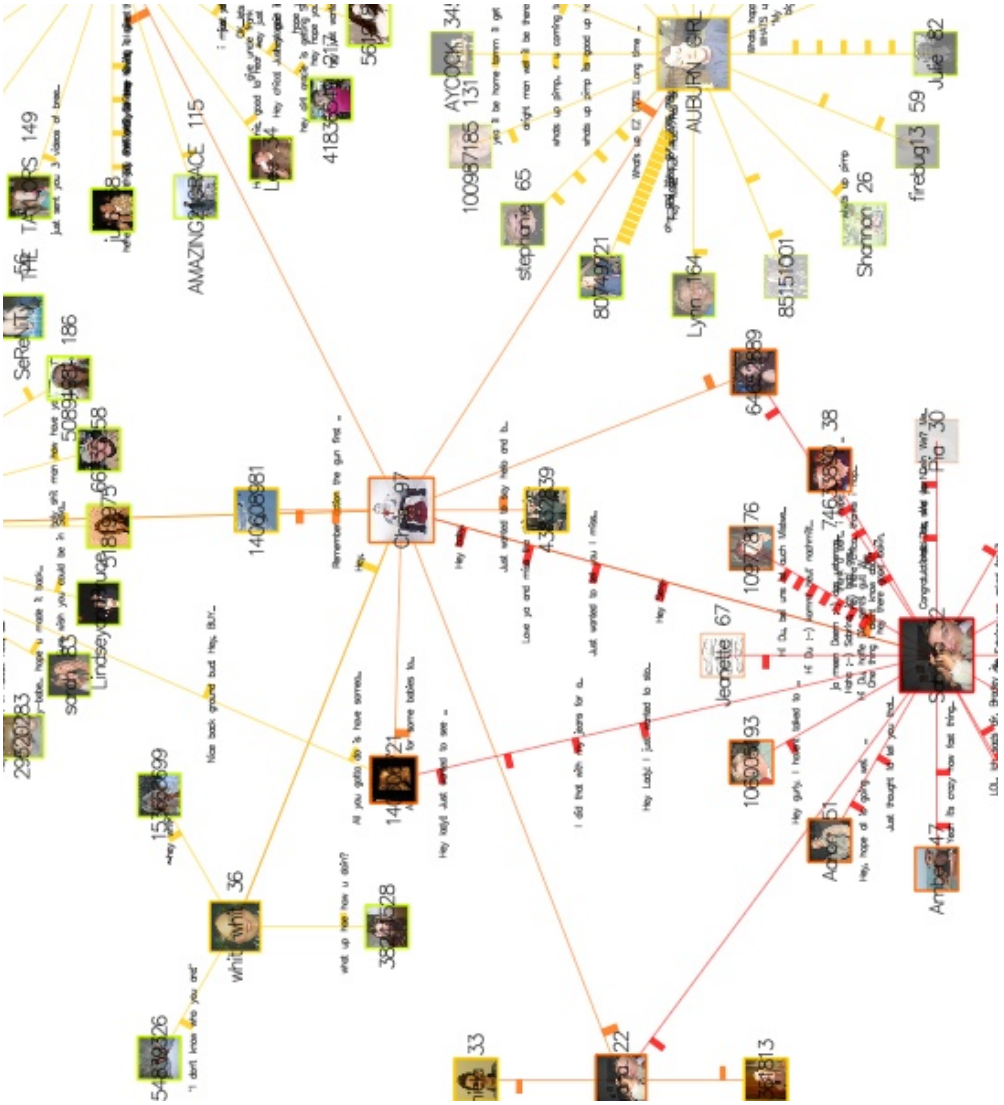


Figure 2.3: CommentFlow - (Offenhuber & Donath, 2007)

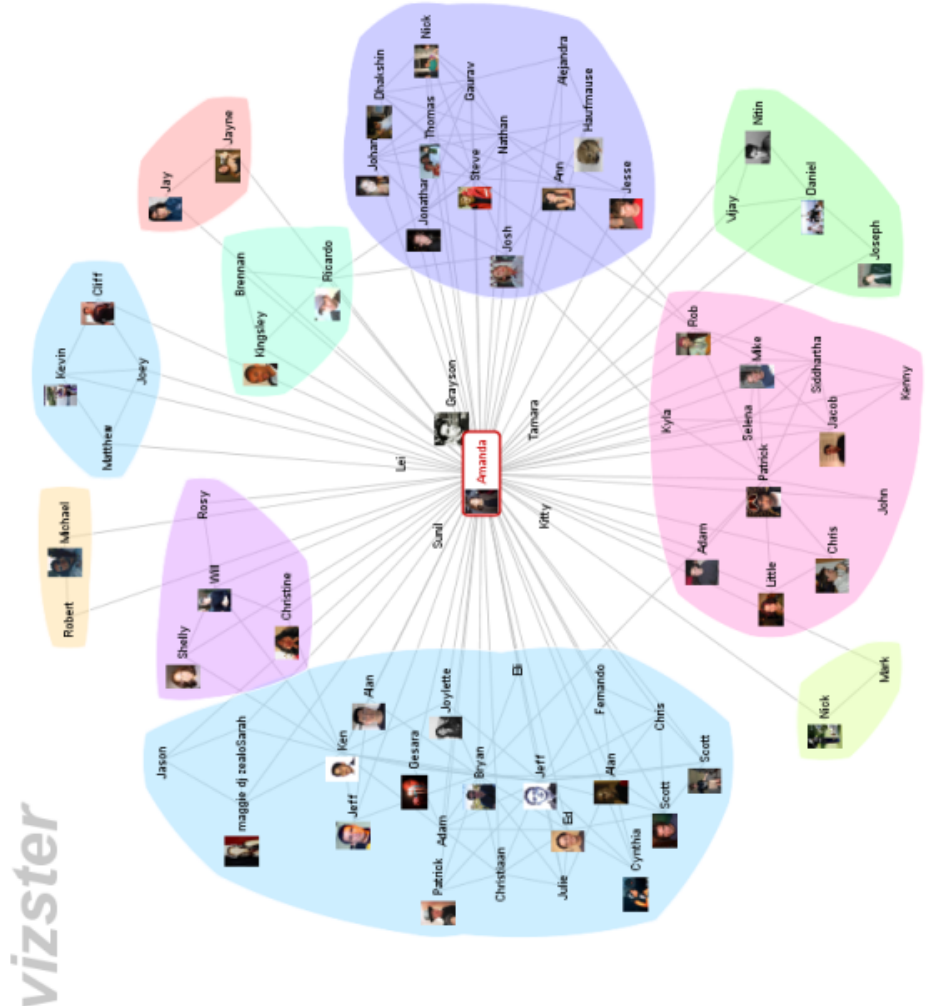


Figure 2.4: Vizster – (Heer & Boyd, 2005)

2.4.2 Behaviour

2.4.2.1 Activity Focus

In contrast to network based interfaces, a variety of research has also focused more closely on the nature of our information behaviours with and of individual people, eschewing the broader focus on the structure of our behaviour. Viégas et al. (2004) created Posthistory (Figure 2.5) to explore email exchange rhythms and patterns. These systems visualised people's email archives focusing on the elements of time and people. PostHistory visualised the amount of email exchange between the user and their contacts over time, on a daily level, and over longer periods of time, allowing users to see the changing rhythms of their communication. The names of contacts are used to visualise the amount of contact with the user in a particular time period. Names which appear closer to the user's name are those who at that point in time the user has been talking to most, while those further away the user talks to less. This exploration of communication rhythms was similar to that of Fisher and Dourish (2004), who found that these automatically derived rhythms and patterns were recognisable and identifiable to users as events in their everyday lives , and that the patterns could be used to augment collaboration tools such as email to provide greater awareness of how they were collaborating in what is traditionally an egocentric medium.

The rhythms of how we share information are something that Google has sought to expose through its Ripples visualisation (Viégas et al., 2013). The visualisation, illustrated in Figure 2.6 seeks to represent the propagation of posts shared by users in their Google+ social network. While arguably similar in concept to some of the network diagrams previously described, this approach focuses more on highlighting to whom information is flowing than being concerned with eliciting a complete structure of a network. People are represented as circles, their size indicating the number of people directly re-sharing the post in question from them, with circles connected together, both spatially and with directional arrows, to indicate the movement of the propagation. By making visible the paths information is taking through people, the individuals responsible for propagating a post to the most other people, and the unique unconnected sources of it become readily clear. This provides people with an awareness of how content they share

is spread by those around them. Arguably marketers have found the greatest utility in this, and other similar systems, using them to identify those actors who most readily spread messages they desire.

Figures 2.5 & 2.6 of the PostHistory (Viégas et al., 2004) and Ripples (Viégas et al., 2013) interfaces show how these activity focused interfaces illustrate the patterns of interaction within communications. The PostHistory image shows daily levels of interaction on the left and intensity of interactions with contacts on the right. The Ripples image shows the propagation of a piece of information between individuals and their impact on the total amount of times it was shared.

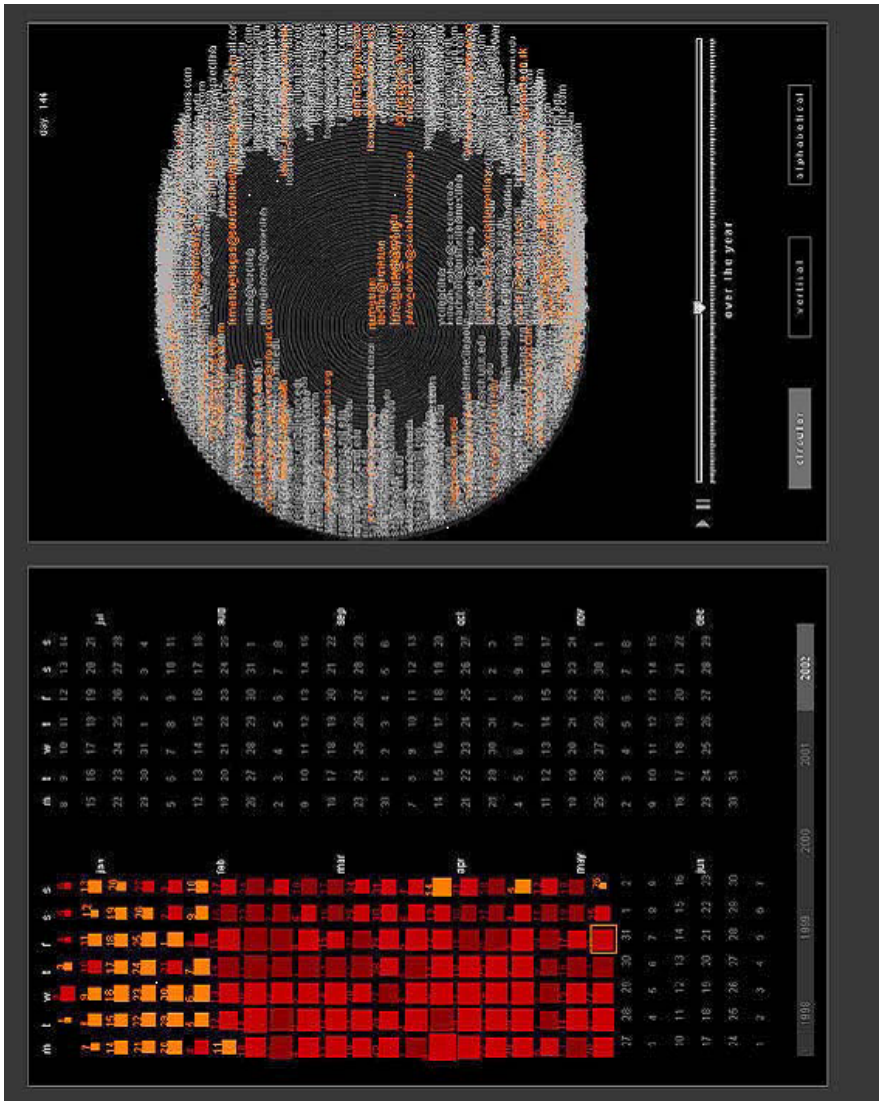


Figure 2.5: PostHistory (Viégas et al, 2004)

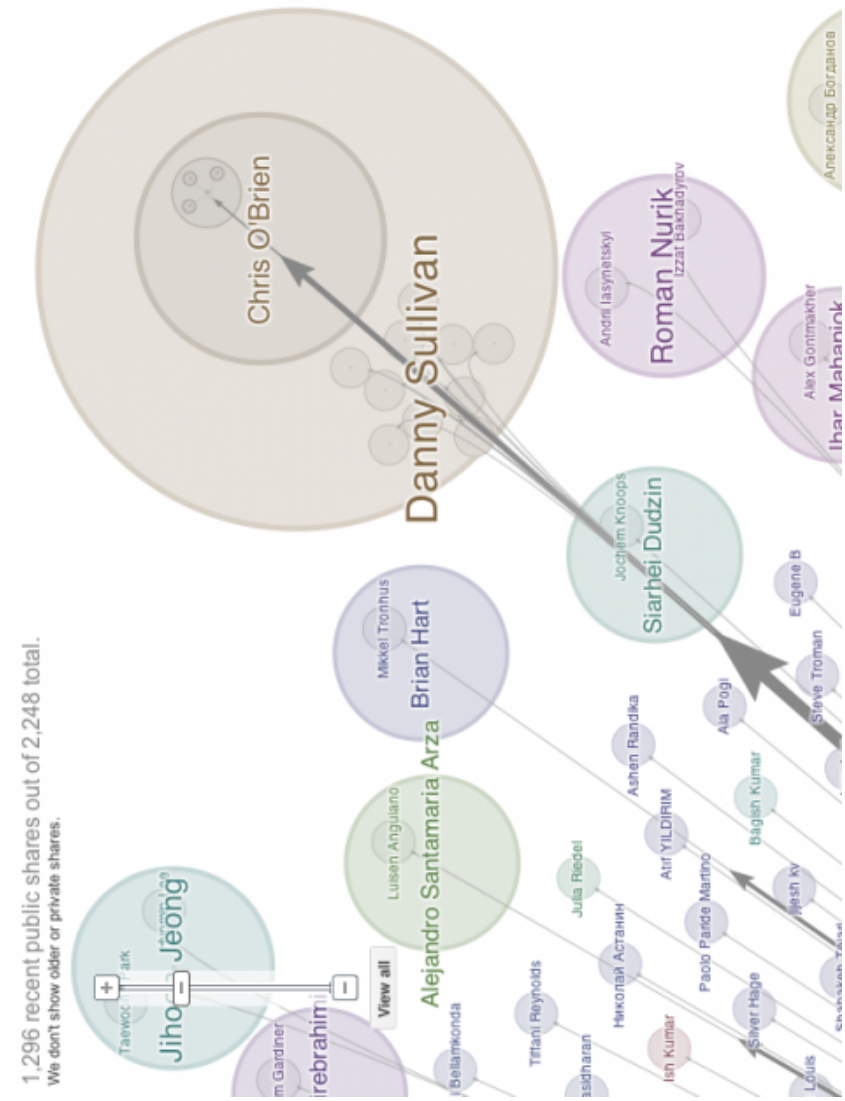


Figure 2.6: Ripples – (Viégas et al, 2013)

2.4.2.2 Content Focus

While these interfaces have looked to explore the varying levels of activity among those with whom we communicate, other work has sought to explore the content of the information we exchange. As an extension of their earlier work, such as PostHistory, Viégas et al. (2006) went on to create Themail (Figure 2.8) Themail illustrated to the user their relationships with their email contacts over time by drawing out the themes of the key topics discussed between the user and individual contacts. Themail displayed keywords used in emails with contacts on a timeline of monthly periods, varying the size of the keywords dependent on the uniqueness of the word to the contact in contrast to other contacts. This allows the user to gain insights into how the character of their conversations with people changed over time, but also what made them unique. In an approach similar to Themail Donath et al.'s (2010) various data portraits – but in particular Lexigraphs (Figure 2.7) – have also attempted to reveal the content of our relationships with people. Lexigraphs focused solely on what an individual had said on their Twitter account. It extracted the user's most often used keywords and used these words to construct the outline of a person's head and shoulders, creating a literal portrait of a user designed to represent the nature of the content they commented upon.

Figures 2.7 & 2.8 show the Lexigraphs (Donath et al., 2010) and Themail (Viégas et al., 2006) interfaces which are focused on the content of information linked with people. In both instances an approach has been taken of associating size with the frequency of a type of topic appearing in relation to individual people, revealing the content of interaction with those people.

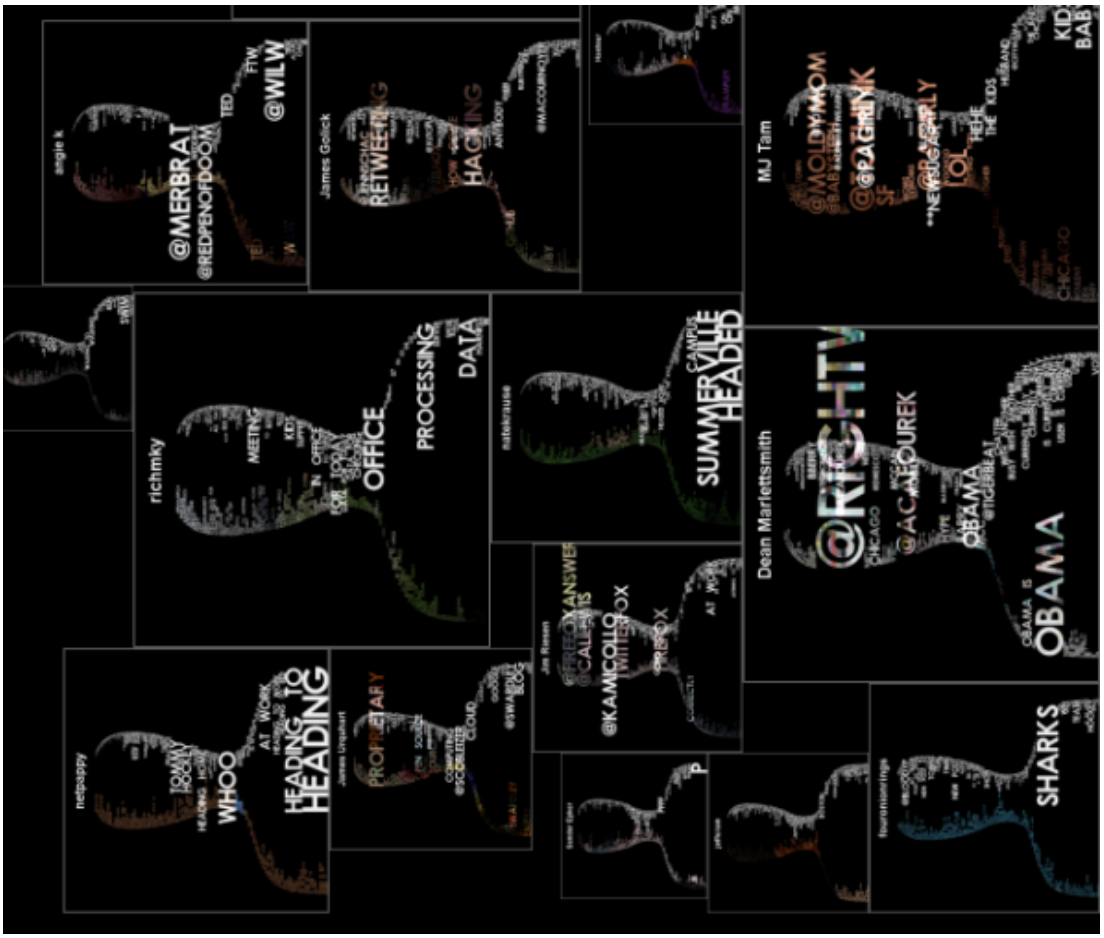


Figure 2.7: Lexigraphs – (Donath et al, 2010)

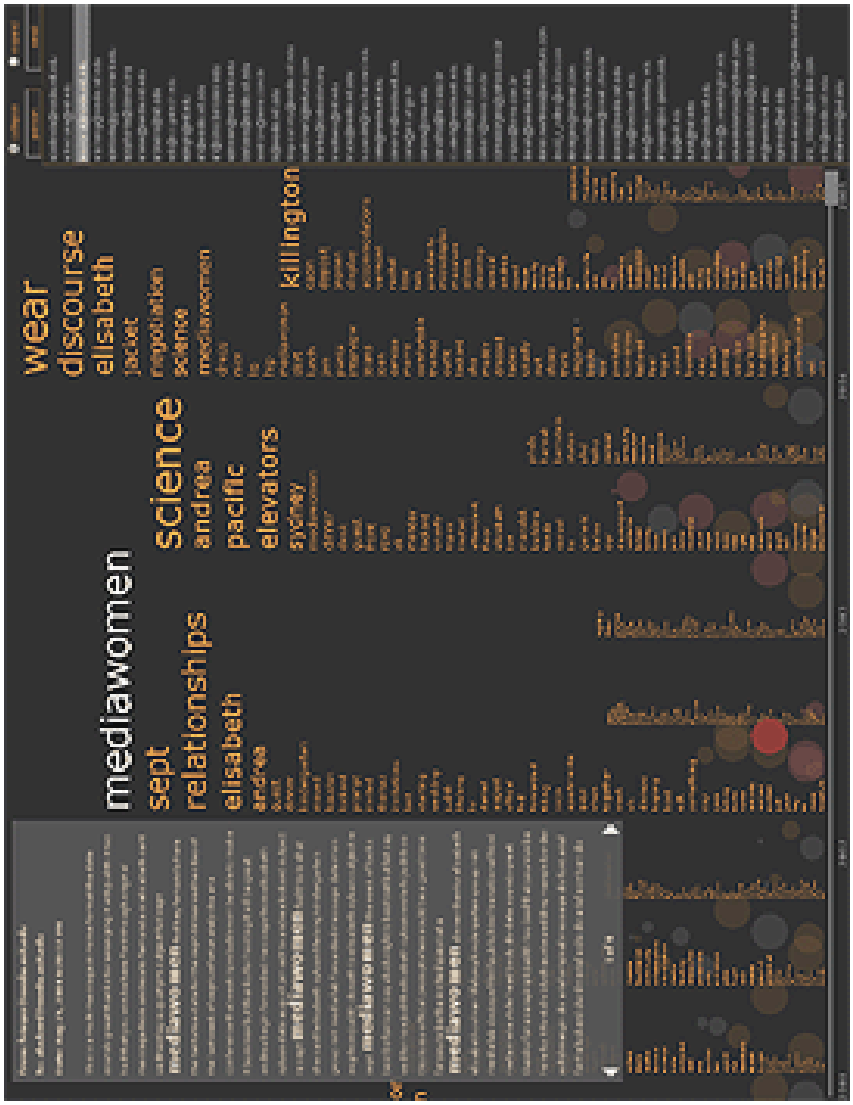


Figure 2.8: Themail - Viégas et al (2006)

These interfaces focused on information behaviour enable people to explore their behaviour by focusing on three separate, but interrelated, facets. The structure of the network, the activity within the network, and the content of the activity within the network. Donath (2002) suggests that visualisations and interfaces such as those developed at the Sociable Media Group are useful because they allow people to reflect upon and so better understand their online social space. While Heer and Boyd (2005, p.32) say that Vizster was useful in '*the discovery of people, connections, and communities to promote increased awareness of community structure and information exposure*', but also increased the enjoyment and engagement in the social network. Additionally, both Fisher and Dourish (2004) and Viégas et al. (2004) report in their work they found that users were compelled to tell stories around their data when using the visualisations, and even wished to share them with friends in some cases.

While exploring our information behaviours with people can provide insight and awareness of the structure, content, and activity, and even enjoyment – for the most part they tend to be of limited practical use in undertaking tasks involving our collaboration with others. As Whittaker et al (2004) note, automated exploratory interfaces such as these, despite being interesting and rich, do not support people in better performing management or social tasks with their personal data, and are not necessarily always representative of the user's personal view of their relations. While these person-centric interfaces may be interesting to the user and researcher, they are less able to contribute to the needs of the user of an information system in their daily information tasks. This makes it difficult to evaluate these designs in terms of furthering interface design in comparison to other approaches or to measure their contribution.

2.5 Conclusion

There are many common forms of interface which centre on the use of time, keywords, and hierarchical navigation, to locate information. These include email, operating system search tools, and operating system folder directories respectively. However, despite people also being an important attribute linked to infor-

mation, and forming an important part of the wider context the information is situated within, there has been very little work focusing on the use of people as a primary means of interacting with our personal information.

Research in this area has so far largely centred on providing a means of awareness of our behaviours in relation to the people we communicate with, and in aiding information retrieval and sharing. The former strand of research seeks to present rich information about our everyday interactions, to reveal the state of our interactions with those who we communicate with most, show the structure of our communications, and summarise the content of them. These rich pictures, however, are often less practically useful and tend towards providing enjoyable experiences, exploration, and reflection (e.g. Donath et al., 2010).

By contrast, those interfaces which seek to aid us in performing information tasks, such as retrieval and sharing, provide more focused and practical views of our information, but at the expense of the richness of information. Grouping contacts into distinct, although sometimes overlapping, social groups has been one of the primary means of representing people in these interfaces, providing a means to access information about people or groups of people, present information about behaviour with them, and engage in the task of information sharing. Here, identifying representative groups of people efficiently is important for the interface to be of value to the user. In these systems it is desirable to provide a system with the ability to group people quickly and accurately, in a way that makes sense to – or reflects – the user's perception of their contacts. Reflecting the user's perception of groups provides a means of locating and using contacts to perform information retrieval or sharing tasks at hand more efficiently, by reducing time spent searching for information or contacts.

Both manual (Whittaker et al., 2004) and automatic (MacLean et al., 2011) solutions to grouping contacts have been researched, but neither currently provide a completely satisfactory means of grouping large numbers of contacts accurately. Often, the goal of speed and the goal of accuracy are in opposition and compromises must be found between them, either choosing slow but representative manual grouping or automatic quick, but often less accurate grouping (MacLean et al., 2011). In between these extremities lie semi-automated compromises where

identification of automated groupings are supervised by the user (Amershi et al., 2012).

Research into contact grouping has largely focused on questions of what or how groups can be detected, but little research has sought to consider what constitutes a meaningful and useful group from the point of view of the user. While group construction has received some research (Kelley et al., 2011) the different forms of groups, and the role of interface design and decision making processes that determine the groupings for use in information management have not been specifically explored. Research with regards to this aspect has largely been limited to observations about forms and quantities of groups created.

Further examination of how different forms of 'person-centric' interfaces might be employed to support users in performing information re-finding and sharing tasks therefore needs to be conducted. The existing literature has explored a variety of concepts for presenting and utilising links between information and people, but little work has been done to examine how people use interfaces such as these, what is good practice for designing them, and how they might be integrated with our everyday information practices. The research in this thesis seeks to contribute to addressing these research gaps by exploring how users prefer to represent and organise people in interfaces, to provide a basis to inform their design, and by evaluation of the use of different designs in practice to establish the utility of different designs and features for supporting information re-finding and sharing tasks.

Chapter 3

Methodology

This chapter describes the methods, and the reasoning for their selection, applied in this research to establish how information re-finding tasks are performed, to elicit how users wish to order and represent contacts in interfaces for performing information re-finding and sharing tasks, and to evaluate 'person-centric' interfaces. The objectives of this research set out in the introduction (Chapter 1) necessitated – among other needs – examining information behaviours, identifying the relationships they perceive between people, and understanding the decision making processes they undertake in performing information task. Qualitative approaches were primarily chosen as a means of addressing these objectives, as this provided greater scope for eliciting the necessary insights into the participant's experiences, motivations, and behaviours in performing information tasks than quantitative measures provided. Quantitative analysis was, however, used where appropriate to supplement this analysis, for example, in the form of descriptive statistics that illustrated the make-up of datasets in the studies. While other means of addressing the objectives of this thesis – such as measures of task performance e.g. time taken to perform a task – could have also contributed to those goals, this would often have been at the expense of a more appropriate qualitative method.

Three studies were conducted over the course of this research, each of which aimed to address one of the objectives. The methods used in these studies are

described in this chapter and listed here alongside the objectives set out in the introduction (Chapter 1):

1. **Study 1**
    - **Objective:** To identify the role of different factors in the process of information re-finding.
    - **Methods:** Diary study of daily information re-finding events and interviewing.
    - **Justification:** A diary study allowed a cross-section of events to be examined in order to identify what factors play a role in the information re-finding process. Interviewing supported the diary study methodology by allowing greater detail about re-finding events to be captured.
2. **Study 2**
    - **Objective:** Having identified the value of knowledge of an associated person(s) in re-finding search strategies, identify principles to support the representation of people and their networks in person-centric interfaces for information sharing and re-finding.
    - **Methods:** Repertory Grid Technique and free hand diagramming of social networks.
    - **Justification:** By using Repertory Grid Technique to elicit constructs that describe how relationships are viewed between personal contacts these constructs could inform how representations of multiple people might be ordered or filtered. Addressing a different aspect, free hand diagramming of personal contacts elicited the different elements that are important to people when structuring groups of contacts in representations of people.
3. **Study 3**
    - **Objective:** Demonstrate and evaluate 'person-centric' interface approaches in providing novel support for information tasks, and identify

design implications for their use.

- **Methods:** Screen capture, think-aloud protocol, and information task generation by questioning the participant to identify information tasks using their personal information.
- **Justification:** Generating tasks by questioning participants to identify personal information re-finding and sharing tasks using their personal information allowed more natural tasks to be set for participants to perform using different 'person-centric' prototype interfaces. Observations recorded through screen capture and the use of think-aloud protocol enabled the decision making processes involved in performing grouping, re-finding, and sharing to be examined. From this analysis design implications could be identified based on the factors influencing the decision making processes involved in these tasks.

3.1 Methods

3.1.1 Study 1 – Information Re-finding Diary Study

Diary studies involve participants keeping a log over a period of time – typically one to two weeks – in which they record details in the log every time a specific event of interest occurs in that period. At the end of the study period participants are also often interviewed with regards to their diary entries or addressing their behaviour in general (Rieman, 1993).The greatest benefit of the diary study approach is that it can provide a large number and variety of individual and naturalistic instances of events. This was beneficial for the objective of examining the process of re-finding where everyday re-finding behaviours were less understood. In contrast direct observations, for example, would require all day observations, including likely unwelcome out of work hours observations, to achieve a similar level of coverage required in this context. A diary study allowed the collection of data describing many individuals' actions during their daily information behaviour events, over a longer period during both work and home hours, in a straightforward manner for analysis.

When examining behaviours that may occur in mobile situations, a paper diary can be inconvenient to participants and be left in one location (Palen and Salzman, 2002). Rather than using a paper based diary, such as in Dearman et al. (2008)'s study of information needs, a mobile phone diary application was developed (Figure 4.1), based on Google's Android mobile phone platform. By installing the application on the participant's own mobile phone it was intended to make the diary easily accessible to the participants regardless of their varying levels of commitment to the study. However, this required that the majority of the participants already possess Android based phones, to take advantage of integrating the study into the phones they use on a daily basis. Even with these efforts to make the diary easily accessible, as others have noted (Palen and Salzman, 2002; Czerwinski et al., 2004), diary studies are subject to significant caveats in that it can be hard to maintain participant commitment to a study over a period of time, self-reported information can be incomplete, it can interrupt and alter the events it is intended to measure, and the need to report can be forgotten all together or the information considered too mundane to report in many instances (Palen and Salzman, 2002). While these drawbacks are important to consider in study design, a key advantage of diary studies is that they provide a relatively unobtrusive method to observer cross sections of naturalistic behaviour over time.

3.1.1.1 Alternative Methods

A diary study approach was chosen as many interview or observation-based methodologies , such as direct observation or screen capture (Stanton et al., 2013), were considered too intrusive in nature. These methodologies would have potentially greater privacy concerns to address in this context, and be more disruptive to the participant's work.

An alternative approach that was considered, which is similar to the diary study approach, was the experience sampling method (ESM) (Consolvo and Walker, 2003). In ESM, rather than the participant having responsibility for reporting on their re-finding events, participants are instead alerted at various intervals throughout each day and at that time must then complete a report. While this can remove some of the biases participants might have about what they choose

to report on and also provide more consistent reporting, there would be several drawbacks to this approach for this study. One of these was that alerts could occur at times quite distant to when the re-finding events in question occurred, meaning that some specific details of the event might have faded from memory. Retaining as detailed as possible a description of the re-finding events was important to ensuring data was recorded about aspects such as what is recalled about the information prior to re-finding it, in order to understand what the re-finding processes participants were undertaking were. Another issue was that the ESM approach could only result in the most recent event being recorded if participants were asked to record all events during that period. It was therefore decided that a diary study would provide a better sample of re-finding behaviours in this instance.

3.1.2 Study 2 – Contact Grouping: Categorisation and Representation

To examine both the broader groupings participants perceive within their contacts and the specific forms of connections between their contacts, two methods were employed. To investigate the broad structure and groups participants distinguished within their contacts, a free-hand drawing task aimed at representing their groups of contacts was used. To examine the individual connections being made between people, participants were asked to perform a repertory grid task using their most recent contacts. The free hand illustrations represented a summary of the participant's own view of the nature of their groups in terms of how their representations were structured and, with the repertory grid providing a contrasting lower level insight into the individual connections that are reflected within that structure.

Free-hand drawings of social networks have only been used to a limited extent in social network analysis, however they provide a simple method of eliciting personal views of a topic (McCarty et al., 2007). In free-hand drawings of social networks participants are typically provided with paper and a range of drawing media and asked to draw a representation of how they view a personal social network. This lack of constraints on the representations they create captures the

variability in how participants perceive their social networks and produces a wide variety of representations of the same topic, which can be useful in generating new perspectives in areas where little is known about the topic (McCarty et al., 2007; Reyes, 2016). Despite the simplicity of this method, it can have drawbacks, in that expressing views through drawing or the lack of constraints may not be suited to everyone, and some participants will be much more comfortable and able to express the facets of their social networks (Reyes, 2016). The lower-level and more constrained examination of relationships used by Repertory Grid Technique, however compliments the use of this method to address this issue.

Repertory Grid technique (Fransella and Bannister, 2004) is based upon Personal Construct theory, which argues that people develop personal constructs based on their individual past experiences, which they use to understand and interpret the events that occur around them. This aids understanding of their present situation and anticipation of how future events might turn out. Repertory Grid technique provides a means of eliciting these personal constructs. In the Repertory Grid methodology participants are typically shown a series of three images (triads) representing variations of the same concept that is being investigated. From each of these triads the participant identifies the two which are most similar along a specific criteria and different from the third along that same criteria. They then provide a label for the criteria, or 'construct', they used. From these constructs bipolar categories can be created, such as 'exciting' and 'unexciting', with which participants are asked to rate each of the images. From this a picture of how people perceive a specific concept can be built up and how specific items within that domain fare according the constructs used to describe it.The approach adopted in this study, however, was to use only the first half of the Repertory Grid technique to elicit constructs concerning commonalities between personal contacts, by presenting participants with triads of their own personal contacts. In this instance the goal was simply to understand what factors are considered in the domain in question, but not to understand how individual contacts would be rated within the domain. Supplemented by the free hand diagrams, these two approaches provided a means of eliciting how participant perceptions about contacts in order to identify the factors that influence how contacts can be

represented in interfaces.

3.1.2.1 Alternative Methods

The methods used in personal network visualisation research, which seek to elicit the names of members of a network and then evaluate specific aspects of the individual relationships between each of the members by comparison, were also considered for use in this study. McCarty et al. (2007) illustrate how individual connections made between members of a network can be built up into a personal network visualisation that reflects broadly similar structures to those devised through a freestyle drawn diagram, but also provide extra insight into the network. These methods are broadly similar to the repertory grid method (Fransella and Bannister, 2004) that was ultimately used, but rather than build a network diagram the repertory grid approach intended to examine the specific types of connections that participants made between their contacts.

Within personal network visualisation, eliciting members of the network is often achieved via some form of prompt that requires a participant to list people based on a particular criteria relevant to the research. For example, participants might be asked to name people who they would talk about a personal subject with. In an example that is more closely related to the needs of this study, in that it examines weaker ties between people, Lubbers et al. (2007) asked participants to list up to forty-five people they knew – and who reciprocally knew them – by name or sight, who they had been in contact with in the last two years. In this instance the method successfully identified a broad range of both weak and strong ties to people participants had contacted. While the social network visualisation approach provides insights through a similar means, the repertory grid method was chosen as it enabled the research to focus more on eliciting how participants interpreted relationships between their contacts rather than priming participants with cues for how they perceive people.

Visual representations of personal networks are a commonly used approach within counselling psychology and social work(McCarty et al., 2007). Freestyle drawing of personal networks, in which participants are not explicitly directed towards a particular approach can offer a simple approach to understanding the make-

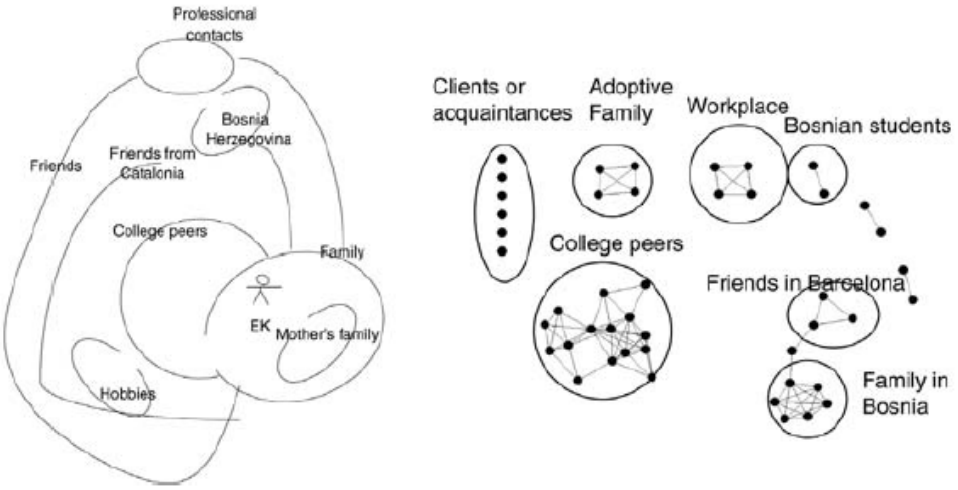


Figure 3.1: Example Drawing and Network Visualization of a Croatian Migrant – Lubbers et al. (2007)

up of personal networks, and provides insight through the variability of how the participants choose to represent their contacts (by how the diagrams were drawn), and possible indications by how they draw the diagrams of what aspects are important when representing them. This contrasts with less free-hand and fixed approaches used in counselling psychology and social work, such as hierarchical mapping technique and Genograms, which dictate the terms and context of how the network is represented (McCarty et al., 2007). This approach is similar in some regards to free-hand drawing, and other expressive techniques, used to elicit the views of users in participatory design (Muller, 2003).

Using a free-hand diagram approach had the advantage of providing the participant's own perception of their personal network. It illustrates their personal view of where co-membership of groups exist, their ordering of the groups, and highlights facets of the form their connection to the participant takes. A drawback to this, however, is that the diagrams can also be a simplification of a variety of aspects that underpin the structure of a social network, and represent only the participants' view and recall of their network, not the reality of the true connections within it (McCarty et al., 2007).

3.1.3 Study 3 – Person-Centric Prototype Evaluation

In order to identify design implications for 'person-centric' interfaces prototype interfaces were designed which accessed information from the user's Facebook and email accounts to identify their contacts and information associated with them. The interfaces allowed user's to group contacts, re-find information, and share information. The purpose of the prototype was used to evaluate concepts identified in previous studies and to identify factors influencing how these three tasks were performed.

To set tasks for participants using the interfaces they were asked a series of questions about their own personal information, such as 'what was the last file somebody sent you?' From the answers to those questions a series of tasks could be generated for participant's to complete based on their own personal information, for example, locating the last file they had been sent. While participant's completed these tasks observations were recorded using screen capture was to record the participant's interactions with the interface.

This approach of eliciting information tasks and performing a qualitative analysis of the data was chosen as it provided the greatest insight into the behaviour of participants, by generating as close as possible to natural information re-finding and sharing tasks with real data. It has been used in a variety of personal information studies due to these benefits((Elsweiler and Ruthven, 2007)), such as Whittaker et al. (2004) in their ContactMap Study and Blanc-Brude and Scapin (2007) in their information re-finding study. It addresses to some extent the issue of not having direct access to the personal data of the participants, due to the importance privacy, that poses difficulties for studying personal information management in a controlled manner (Kelly, 2006).

In addition to screen capture audio was recorded to capture the participant speaking while performing think-aloud protocol (Ericsson and Simon, 1993). Think-aloud protocol requires participants to verbalise their thought processes, requiring them to say what they are thinking at as each new thought occurs. Applying this method allowed elements of the decision making processes regarding grouping, re-finding and sharing to be made explicit, and to understand their interactions

with the interface. As the method was originally published and applied in a range of research no interventions should be made by researchers beyond reminding participants to keep talking. Interventions beyond this were considered to invalidate the data, as it is argued that this can interrupt the flow of information from the participant's short term memory. However, within usability testing interventions have been used in order to probe, elicit and clarify the participant's mental models – so long as the questioning remained neutral – (Tamler, 1998), as it has been argued that 'what is being studied is the system itself, not the cognition of the test participant', and therefore the original theory can be extended to consider contexts of use beyond which verbal protocol was originally envisaged being used (Boren and Ramey, 2000). In this study interventions were used during long periods of silence or hesitation by the participant, during the contact grouping tasks, when they stopped interacting with the interface. The interventions were used to prompt participants keep talking, during what became clear from the participant's comments were moments of decision-making about grouping strategies. Depending on how much information was proffered by the participant about the decision-making process of grouping the contacts, following on from being prompted, one or two follow up questions would be asked to elicit a full description of that decision-making process.

3.1.3.1 Alternative Methods

Capturing performance-based measures, such as time taken to complete a task, was also considered as it could have provided a valuable comparison across each of tasks for each interface type. However, the inconsistency of the personal data made this a difficult measure to use in a balanced manner. Although a uniform experience of the types of sharing and re-finding tasks was provided to ensure participants undertook the same processes, the difficulty of the tasks could potentially vary wildly. Information retrieval studies have often utilised standardised collections (e.g. Text REtrieval Conference – TREC – collections) of data to provide uniformity of the tasks being performed, however, this consistency of data is not available when using personal information. Standardised personal information collections Kelly (2006) have been proposed, which would address issues of

privacy and provide standardisation of tasks. However, the value of equality for comparison in collections has to also be weighed against the unfamiliarity of that collection to participants. Personal information collections are inherently defined by the personal experiences of their owners (Elsweiler and Ruthven, 2007). In order to understand personal information retrieval processes, the data necessarily needs to be personal to the participants as they will need to draw on their own personal experiences with the data in order to retrieve it (see Chapter 4). Additionally, the focus in this study was on uncovering the decision making processes of the participants, particularly through think-aloud protocol, which adds an extra cognitive burden to performing a task (Preece et al., 1994), that could also affect any performance based measures.

Critical Decision Method (CDM) Klein et al. (1989) was considered as an alternative methodology that could be applied to understanding the decision making processes involved in performing information tasks with the 'person-centric' interfaces. It has been used to understand the decision making of fireground commanders, paramedics, and engineers in order to elicit expert knowledge, the cues attended to in decision making, and the strategies employed in making decisions. In it participants are required to recall an event that was particularly memorable to them in the course of performing the work being investigated. The recollection of the memorable event is then used as a probe to identify the decisions being made, and to understand how and why they were made.

As Klein et al describe, much of the value of CDM is in modelling tasks in environments where there is 'high time pressure, high information content, and changing conditions'. By contrast the context in which CDM was being considered for was low time pressure, limited in the extent to which conditions changed, and if not low in information content, quite homogeneous in the forms of informations available. While the absence of the complexity that CDM was designed to describe did not make it inapplicable to the decisions being considered, much of the value of CDM is in the depth of its examination of a few decision making events. The decisions being examined in this study took place in a relatively stable environment that lacked complexity in terms of the range and variability of factors that could potentially influence the decision making processes. A more lightweight

methodology that could be applied to all the decision making events, and therefore be used to identify a broader picture in terms of the types of decision events occurring was therefore determined to be more beneficial overall.

3.2 Summary of the Approach to the Research Questions

In the introduction of this thesis three objectives were set out in order to understand how knowledge of a 'person' can be used to organise and interact with information in order to support users in performing information tasks with their personal information. To address each of these objectives a range of methods were employed as part of the research in this thesis which are summarised here:

1. **To identify the role of factors that enable the process of information re-finding.**

   The first step for this research was to gather data capable of providing an insight into re-finding practices and behaviours. Data that captured a cross-section of how everyday information re-finding tasks were being performed would provide a sample through which common information re-finding processes could be analysed to identify what factors were important to their success or failure. This required a method that would capture a range of natural information re-finding events. A diary study was chosen in this instance due to it's lightweight and less intrusive nature relative to other methods, and it's ability to capture a large number of separate re-finding events from which the important factors in the re-finding process could be discerned. This results of this were followed up with interviewing about each of the captured entries. This allowed participants to expand on their explanations of their re-finding processes, providing a richer dataset than the diary study alone could provide.

   The results of applying this method directed the research that follows it to focus on what role that having interacting with or organising information around people can have for performing information re-finding tasks. This

set the direction of the following study to focus on how people should best be represented in interface design for personal information tasks. In addition to this the results of this study more broadly contributed to the design the of the person-centric prototype interface described in Chapter 6 and the study that evaluated its use in Chapter 7.

2. **Having identified the value of knowledge of an associated person(s) in re-finding search strategies, identify principles to support the representation of people and their networks in person-centric interfaces for information sharing and re-finding.**

   The next step in the research was to understand how best to represent or access potentially large numbers of contacts to support the personal information tasks being performed. In order to exploit connections that are made between people and information it was important therefore to elicit an understanding of how contacts are perceived in terms of the structures deemed most salient for bringing order to those contacts.

   To meet this aim knowledge elicitation methodologies needed to be employed. Two methods were employed in this instance, one to identify the relationships that are perceived between contacts which might be used to organise representations of contacts, and one to identify what form the visual structures those relationships might be expressed through. Repertory Grid Technique was employed in the case of the former to identify relationships between contacts. It was chosen for its ability to elicit the underlying constructs people use to describe a domain. By comparing participant's own contacts a picture could be built up of the types of relationships that are most salient to people in words participant's would choose to describe those facets. Similarly, free-diagramming of how participants would represent their contacts for information tasks, allowed participants to express how they would prefer to represent their contacts, highlighting structures of representation that were most important to them. Analysis of the diagrams and constructs could then be used to address the research aim by identifying common means of ordering and structuring contacts in the data that could guide how 'person-centric' interfaces should be designed. These

results could then be used to guide the design of prototype 'person-centric' interface design in the next study.

3. **Demonstrate and evaluate 'person-centric' interface approaches in providing novel support for information tasks, and identify design implications for their use.**

   The previous studies had identified a means of potentially providing greater support for performing personal information tasks through connections made between people and information, and identified a range of ways of representing people in interfaces. It was next important that validation occurred of the concepts identified in these studies through a practical evaluation that examined the principles and design ideas being put forward. A prototype 'person-centric' system was developed, which incorporated the findings of the previous research into its design in order to facilitate the evaluation of those concepts. For the purposes of this study it was important to make the information tasks being performed with as natural as possible, therefore using the participant's own data, whilst being lab-based so that the value of different aspects of the prototype could be investigated in a methodical and balanced way. A method of setting tasks based on information elicited from questioning participants was chosen as it provided a means of ensuring specific types of tasks were performed by each of the participants, to make their experiences using the system uniform, whilst also setting tasks with the participants own data, which was critical for ensuring the tasks were performed as naturally as possible in a lab setting. A more artificial task without data personal to the participant would not enable them to draw upon their memories to perform the tasks, which had been identified as being critical to the processes being examined.

   Observation by recording screen and audio capture of participants performing the tasks, followed by interviewing about their experiences and preferences using the system, was chosen as a means of evaluating the merit of different elements of the system's design. Think-aloud protocol was also used in conjunction with the observation in order to investigate the processes participants went through in performing the information tasks and group-

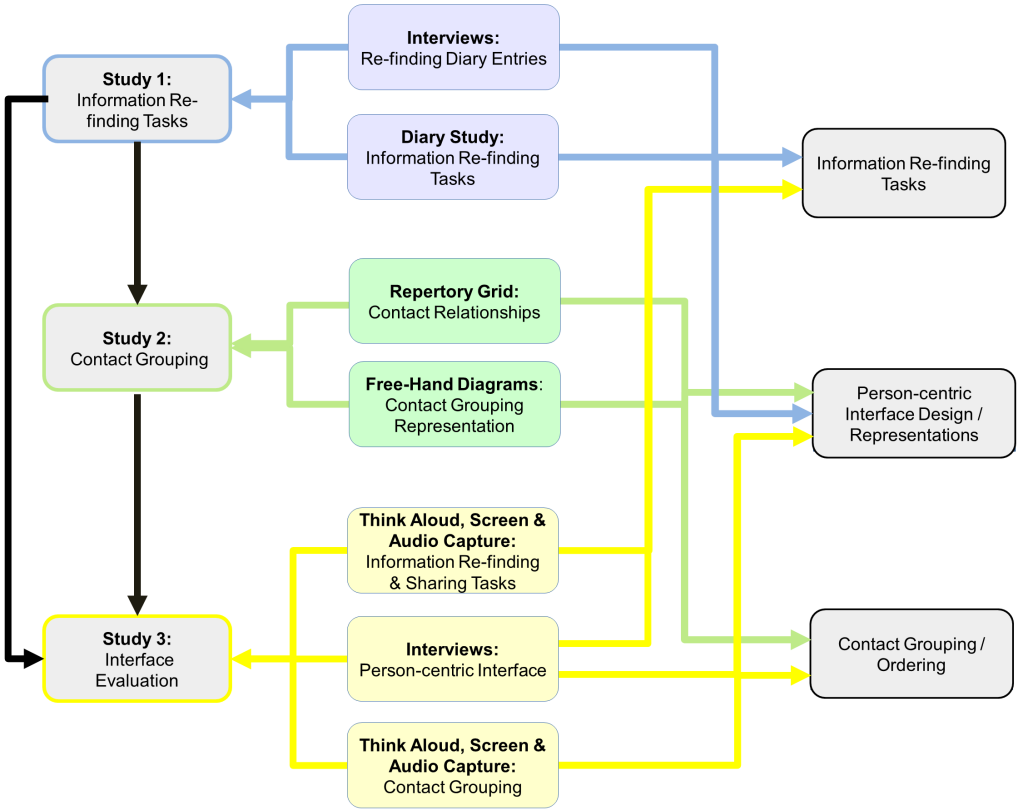


Figure 3.2: Thesis Structure & Methods

ing of their contacts. By having participants make explicit their thoughts as they performed tasks with the system, this allowed a deeper examination of the decision making processes involved in performing the tasks. Analysis of how the tasks were performed alongside analysis of the participant's description of their experiences using the system could then be synthesised to identify broader implications for the design of systems that focus on the use of interaction with people for performing information tasks.

Figure 3.2 summarises how each of the methods described here map onto the thesis. In the centre of the diagram is each of the methods used in the thesis, and to their left is the study they were part of each of. The arrows between the studies illustrate how studies informed design of subsequent studies. To the right of the methods are the three broad objectives that have been outlined in this section.

The arrows between the methods and objective areas represent that the results of applying that method contributed to meeting the connected objectives.

Chapter 4

Re-finding and Information Tasks

4.1 Introduction

Information storage, and access to information across applications and services has become increasingly fragmented (Newman, 2008). Not only has the amount of information individuals generate grown, but the number of different, devices, services and tools used to access that information has also multiplied. This raises the question of whether the increasingly large amount of personal information that people encounter, manage, generate, and receive poses an increasing challenge for managing and, in particular, *re-finding* information. While there is a growing body of research on managing information in the face of this data deluge, there is little to no research on re-finding - even though it is a markedly different process from general information retrieval (Capra et al., 2005), and one that many of the information tasks that are performed consist of (Teevan et al., 2007).

In order to design systems that will provide greater support to people in managing their personal information tasks, this first requires a greater understanding of the nature and difficulty of the processes that people undertake to re-find and reconstruct information. In this chapter the objective outlined in the introduction to this thesis of identifying the role of different factors that enable the process

of information re-finding is addressed. An initial scoping study is described, the intention of which was to identify areas where there may be difficulty in performing information tasks, and where greater support might be provided in completing them. To this end a two-week diary study was conducted using mobile phones to record the diary entries. The study captures a broad cross-section of the information behaviours exhibited by people when performing information re-finding tasks, and enabled the creation of a structured analytical framework to analyse the re-finding process and to codify: the information needs people wish to satisfy; the information tasks those needs satisfy; the methods utilised in the re-finding process; and the types of memories and cues that people use to locate information. Also examined is where people are broadly looking to re-find information, how recently they last saw it, and the difficulty of re-finding it. This contributes to informing a wider view of the characteristics of information re-finding. By subsequently applying this framework to the study's results, areas to support information re-finding tasks in information management systems are highlighted. Analysis of these results identifies the role that the associations made between people and information play in the re-finding process as factor to explore further, which is then considered in greater detail.

4.2 Methodology

4.2.1 Study Design

While many studies that have examined user behaviour in re-finding information, have focused on a single source of information or context of use, particularly web and email behaviour (Capra and Pérez-Quiñones, 2003; Teevan et al., 2004), in this study the aim was to analyse a broad cross-section of re-finding behaviours. To capture and understand re-finding behaviour in general, this was examined across multiple devices, contexts (work and home), as well as in both static and mobile situations. To this end, and in order to elicit a cross section of the information tasks people perform on a daily basis, a two-week diary study (Palen and Salzman, 2002) was conducted that captured all occasions where people attempted to re-find digital information. This approach provided a means to col-

lect descriptions of participant's behaviours concerning a representative sample of naturally occurring daily personal information interactions over the period of the study.

In order to record the re-finding events a mobile phone diary application was developed (Figure 4.1), based on Google's Android mobile phone platform. Participants were asked to record details of any information they searched for that they had already previously seen. When recording a search event, participants were prompted by specific questions in the application to complete their description of the re-finding event. By installing the application on the participant's own mobile phone it was hoped to make the diary easily accessible to the participants regardless of their varying levels of commitment to the study. However, this required that the participants were Android phone users. Of the eighteen participants only three were not Android phone users, and these participants were provided with Android phones that they would substitute as their primary phone for the period of the study.

The diary application was programmed and designed specifically for this study by the author of the thesis. The design of the diary application (see Figure 4.1) took into consideration previous diary studies such as those of Sohn et al. (2008) and Dearman et al. (2008), which examined similar facets of personal information behaviour, related to search and information needs, and the design of mobile phone based diary studies, such as those of Sun et al. (2012) and Palen and Salzman (2002). It was designed for participants to perform two primary tasks. This first task was to provide a simple means to answer a series of questions which constituted a diary entry, this took the form of a questionnaire, with a combination of multiple choice and open ended questions. The second task was to provide a means to support partial diary entries, so that participants could complete the diary entry shortly after the event if they were unable to do so at the time. This was supported through the ability to save partially completed entries and retrieve them, and through note taking facilities that could support their recollection of the event. Note taking took the form of typing, recording audio, photos and videos that were associated with a diary entry. The data from these entries was stored locally in the phone's memory. On completing the study

this data was removed from the phone and stored securely.

Each diary entry consisted of the following questions:

1. What was your information need?
2. Was the information you required part of a work task or for a personal use? (Work, Personal)
3. What sources of information did you use to address your information need?
4. Were you successful in obtaining the information you required? (Yes or No)
5. What prior knowledge about the information you were interacting with did you remember before performing your information task?
6. How easy or difficult was it to find the information you needed? (Rating 1 - 7)
7. How long since you last encountered this information? (Within 1 Week, 1 Month, 3 Months, 6 Months, Year or longer)
8. Please describe how you found the required information. Or if you failed to address your information need, why was this?

The more open ended questions – questions 1, 5, and 8 – aimed to elicit descriptions of the information being sought, the process of finding the information, what information was recalled about the information sought, while the closed questions probed related aspects of behaviour such as difficulty, time since last encountering information, and success. It was intended that investigating difficulty and time would reveal some of the characteristics of circumstances where people struggled to retrieve information.

As in Sun et al. (2012), upon completion of the two week period post-study interviews were conducted with the participants using the diary entries as memory cues to elicit greater detail about the process they went through with each entry, particularly with regards to the open ended diary questions. In addition this also allowed any unclear entries to be clarified Sohn et al. (2008).

Ethics approval to conduct the study (see Appendix C ethics documents) was

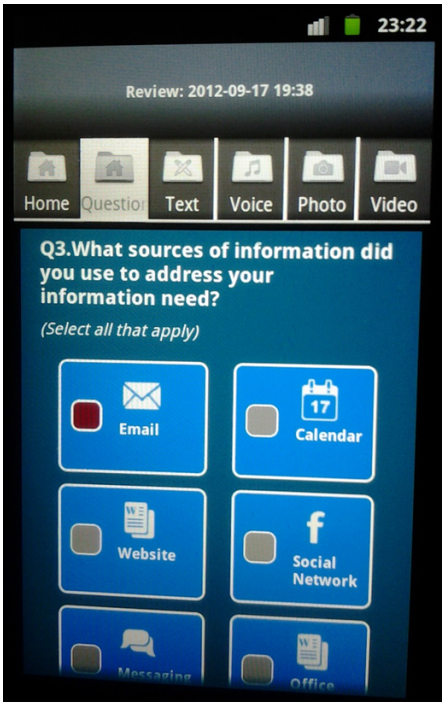


Figure 4.1: Diary Application

granted by the Faculty of Engineering. The application detailed how the data from the study would be collected, used, kept securely, and considerations of its impact on the participants. Before taking part in the study each participant was given time to read an information sheet describing what their participation in the study would entail. This included how data collected about them from the study would be made anonymous, protected, and used, as well as their right to withdraw at any time and for their data to be removed from the study. If they were happy to proceed, participants signed the consent form confirming that they understood the conditions of the study and were happy for the data collected on their mobile phone could be used for the study analysis. Participants were compensated for their time through high street vouchers worth £20.

4.2.2 Participants

Twenty participants were recruited, of which eighteen completed the study (7 female, 11 male) with ages ranging from early twenties to mid forties. Participants

were recruited at the University of Nottingham through email calls for participants and word of mouth. The participants were primarily academics and students from a variety of disciplines. Other professions of participants in the study included an event organiser, an information technology consultant, a project manager and a secretary. The participants were chosen as a convenience sample, it was considered that there was no reason to believe there would not be enough diversity represented among this group, and that most of the forms of processes and types of motivations involved in re-finding information would be present. This sample likely is more representative of people involved in more organisational forms of information retrieval events, possibly influencing the make-up of some of the results, which has been noted in the discussion.

4.2.3 Procedure

Participants asked to record all occasions when they had to re-find information, on any electronic device, and the details of what these tasks consisted of for a period of two weeks. Re-finding was described as any occasion where participants needed to find information that they had already previously seen, and were provided with a broad range of examples of the types of tasks which might constitute this. For example: locating a webpage you have previously seen, navigating to the file you wrote your shopping list in, and locating an old email where you only know the sender's name or the subject line. Additionally, longer multi-part examples were described, where users may find information from multiple sources, such as finding the log of an old instant message that mentions a nice restaurant in an area unfamiliar to them, that they then search for on the Internet to re-find its web page. Participants were also told that the recording of these details would preferably be done shortly after finishing a re-finding task, but they may alternatively be completed at a later more convenient time as in Sohn et al. (2008). Following this explanation the diary application was installed on their phone and participants were shown how to create diary entries with it. Each of the diary questions were explained to the participants and any questions clarified. Additionally each of the note taking facilities were also demonstrated to the participants. At the end of the two week period the diary data was collected

from the participants' phones. Each of the diary entries were then examined with the participant as part of an interview. Each question in a diary entry was examined with the interviewer, in which the researcher asked for more details related to each diary event.

4.3 Analytical Frameworks of Information Re-finding

From both the diary entries collected by the mobile phone application and the subsequent interviewing, three analytical frameworks were produced from the data. As in Vogl et al. (2018) these analytical frameworks constituted the development of a re-usable method for systematically analysing large and complex data sets in a specific context of application, in this case information re-finding. This provided a means of extracting insights into the make-up of the data, by categorising components of different dimensions of the information re-finding process.

The frameworks describe three different aspects of information re-finding search behaviours and quantified the make-up of those behaviours over different search events. Framework 1 describes the different types of information needs that people want to satisfy when re-finding information. Framework 2 describe the tasks, the purpose, that participants stated they were seeking the information that it is that the information needs were satisfying. Framework 3 identified and analysed what people remember about information when re-finding it, to aid understanding of the process of re-finding.

The prior work of Dearman et al. (2008) provided an existing method to describe the information needs expressed in the diary entries and was applied in the first instance to the information needs described by participants. Subsequent to that, any diary entries which did not fall cleanly into any of the groupings described by Dearman et al. were then categorised afterwards into new categories extending the framework based on the common themes between the remaining entries. The information tasks and memory frameworks were elicited by the-

matically analysing both the diary entries and interview data pertaining to each entry without any initial starting point. In the case of the information tasks this was done on the basis that the categories should be mutually exclusive and were revised until all of the data could only be accurately described by one of the categories. In the case of the third framework this looked to identify categories of meta-information expressed in each entry, of which there might be multiple instances in an entry.

4.3.1 Information Needs & Tasks

Each of the diary entries were coded into categories representing the information need that is being addressed in that entry by re-finding the piece of information sought. These information needs were coded and categorised according to a categorisation of information needs by Dearman et al. (2008), which is summarised here:

- Person - Information about a specific person or oneself.
- Establishment and Organisations - Information about a specific business, association, society, civil department, etc.
- Offerings - Information about a specific product or service provided by an establishment and organization.
- Events - Information about a specific activity, function or event of importance that involves people.
- Environment Conditions - Conditions of the physical environment that currently or in the future will impact the inquirer.
- Trivia and Pop Culture - Seemingly random, inquisitive thoughts, factual knowledge, and pop culture references that permeate our everyday life.
- Finding - Information required to identifying and locating a person; establishment and organization; offering; event; and asset.
- Availability - Information identifying when a person is available; an organization, establishment and event are accessible; and an offering or asset is

obtainable.

- Guidance - The knowledge to perform an action and inform a decision.

The intention here was to use Dearman's categorisation and compare the difference in information needs when re-finding to finding. However, the application of the categorisation did not fully describe the results of this study, a finding which further underlines the differences between the processes of finding and re-finding information. While 76% of the entries fell neatly into Dearman et al. (2008)'s categorisation of information needs, it could not be used to categorise the the remaining 24%. In this context of re-finding information Dearman's categorisation's focus on questions about people, organisations, and other sources external to the searcher meant it was unable to describe internal questions about the participant themselves that arose in the data. The categorisation therefore needed to be extended to take into consideration internal questions about information related to the participants themselves; questions such as '*What are the items on the shopping list I wrote yesterday?*'. As a result, in addition to Dearman *et al.*'s scheme further categories of need were identified to account for the entries that were unable to be classified. These new categories were identified by examining the remaining entries and thematically coding them (Braun and Clarke, 2006) according to what distinguished them as being unable to be addressed by Dearman's categorisation. These newly created categories are described in greater detail with examples from participants here:

Personally Created – Information that is created by the searcher themselves that they wish to re-find, usually to review what they've written, often these are memory aids such as shopping or to do lists, '*Okay, shopping list. So, I use my phone to put things in to buy. Nothing exciting, just things I need to get*'. But also this might be part of continuing that creative process, '*. . . this was a piece of video editing that I was doing. It was someone's wedding that I'd taken a video camera to film the service, but I'd previously edited together some bits of music that were played at the service. So I was looking for that file I edited for the service to put over the top of the video*' (Participant D).

About Yourself – Information written specifically about the searcher them-

self or a group of people you belong to by another person or establishment & organisation. Again this was often for purposes of reviewing, '*This is an email telling me what to bring in for my dissertation project meeting with my supervisor*'(Participant P), but not always as a memory aid, sometimes this was to assess the state of relations between the person and others, '*This one was just I wanted to re-read an email for some specific information. I got rejected from a play, and I knew it was in my inbox, I knew I'd only received it the day before. I wanted to go read the kind of wording of the email again [laughs], so I did sort of torture myself*'(Participant K).

Publications – Where a publication as a whole or a section of it is sought, and the value lies in it being a piece of original work, such as photographs, academic papers & films. '*A paper I needed... I think I wanted a certain quote. I was writing something and knew I'd seen that quote, or reference, or comment made somewhere before*' (Participant D).

Applying Dearman et al.'s categorisation defines information needs such that they could be clearly and explicitly stated, but a number of participants in the study expressed more nebulous, less well-defined needs. These generally fell outside of Dearman's categorisation, and were often tied up with a task that the participant was performing. These were needs such as '*This paper won best award at CHI, so I was using his paper as a reference of how to write an introduction*' (Participant C) where here re-finding the document satisfies the search for information, but whether it satisfies the information need is subjectively answered by the participants rather than being an explicit fact to be located. Resolving this information need is arguably a part of the process of the task they might be performing (in this case writing an academic paper). The extra categories identified here are quite broad descriptors of a variety of types of needs, and with more data more specific themes are likely to emerge, allowing the extra categories could be broken down into more specific sub-categories.

In addition to clarifying what information needs participants wished to resolve, interviewing also revealed in most cases why they were looking for the information, i.e. the tasks they were attempting to perform. These tasks were categorised to represent common themes concerning the goals that participants were attempting

to achieve with the information they found.

Checking – Looking for information to confirm the specific details as part of a personal situational awareness. This often pertained to a need to be aware of times and dates of specific events e.g. meeting times, bus timetables. These entries broadly had two different motivations for the Checking; instances where the information was broadly remembered and often being checked primarily for a sense of confidence, and instances where the information acted as an external memory aid reducing the need to remember the details. Thus the category of checking can be divided into two sub-categories:

Confirmation – The person suspects that they more or less know the information – or have it in a form of external memory (see below) – but wish to confirm they recalled it correctly. '*One of the physios I've got sends you text messages, and I got a text reminder telling me when it is. Although I put them in the calendar, sometimes when the text comes through I'll check it against the calendar just to make sure. I have been known to turn up to appointments a week early...* (Participant I)'.

External Memory – The information is effectively used as an external memory aid, reducing the need to retain all of the details of a piece of information. Specific details of the information needed are instead retrieved as required from the information source, '*Checking date of nieces' birthdays. At the start of September there are five or six birthdays and I was trying to remember whose they were*'(Participant I).

Planning – Information needed to aid the planning of a future activity. Participants would search for information they had previously seen to aid them in planning a specific future action or activity. '*In the next few weeks we're off to Liverpool, London, Glasgow, Manchester, Guildford... to look at various drama schools. And this was kind of I wanted to review, which ones where and when*' (Participant A).

Communicating – Communication information, required in order to contact a person or organisation. Participants needed to go through a process of re-obtaining contact details before they could contact a specific person or organisa-

tion. '*Looking for somebody's phone number, and I found it from an SMS they sent me, way back when*' (Participant F).

Creation – Information required as part of a process of creating something new, e.g. a document. Participants needed to create new forms of information by drawing on information they had previously seen. '*This was a piece of video editing that I was doing. It was someone's wedding that I'd taken a video camera to film the service, but I'd previously edited together some bits of music that were played at the service. So I was looking for that file I edited for the service to put over the top of the video*' (Participant A).

Decision Making – Collating information to inform decision making e.g. purchasing. Participants would need to go through a retrieval process to locate information that could potentially influence the outcome of a decision they needed to make. '*This was an email with links to look at for netbooks... from my dad with links to websites I needed to look at*' (Participant P).

Entertainment – Information sought for the purpose of pleasure e.g. listening to music, watching humorous videos, and reading webpages. '*I think it was the website for 4OD. I wanted to get on the website for 4OD. I was looking for a particular programme, which I'd been on before,* Fresh Meat'(Participant E).

Knowledge – Information that is required to expand the knowledge of the searcher e.g. as academic publications. Participants sought information with the intention of remembering it in the long term, often with the aim to learn a specific skill or expand their knowledge on a specific topic. '*I'd looked through a few Photoshop tutorials. I'm trying to teach myself how to ink using a graphics tablet. And errrm, I'd saved a few videos for later*' (Participant C).

Maintenance – Information needed to maintain or update the file, computer, or service of the searcher e.g. password retrieval, serial numbers or software. 'I *was reinstalling some software and I needed the licence key...*' (Participant F).

Physical Action – Information needed to carry out non-digital physical tasks e.g. recipes to cook with or codes for train ticket collection. '*So this was to find a recipe on the BBC Good Food website that I'd looked at ages ago but never saved.*

And I remembered it because I had particular things in my fridge and I wanted to use up, and remembered about the recipe I'd seen' (Participant E).

Sharing – Information remembered and re-found for the purpose of sharing it with another person or organisation it had suddenly become relevant to. '*I wanted to show my mum and outfit an actress had. I presume it was something I had seen somewhere else, and it had come up in conversation with my mum, and I was showing her the same YouTube video*'(Participant C).

4.3.2 Known Attributes and Associations

The data from question five – 'what prior knowledge about the information you were interacting with did you remember before performing your information task?' – of the diary entries was analysed in conjunction with the interview data and coded according to the Known Attributes categorisation (section 2.4.2). These categories represented what people remembered about the information they were searching for before looking for it, breaking down their responses into the constituent pieces of meta-information people said they recalled when locating the information; both stated and strongly implicit in their descriptions.

Device – The memory of which device the information was either stored on or accessed through.

> *Device On* – The device that was used to access the information where the information is held locally on the device. '*...day visitor was coming, approximate time, name of visitor, that details were in a text message*' (Participant I, Female, 42, Administrator).
>
> *Device Used* – The device that was used to access information where the information is not solely stored on the device e.g. via the Internet. '*...meeting was logged on my Google Calendar, and accessible via my mobile phone*' (Participant A).

File Type – The memory of the type of file that required information is in, or is likely to be contained within. '*...that I had made a notepad document containing passwords to web services I don't often use*' (Participant A).

Icon/Colour/Image – Any form of visual image, icon, or colour that is remembered to be associated with the sought information. '*My calendar is colour coded, so work meetings have a purple hue. Going in to the calendar I know I have that visual*' (Participant A).

Source – A source of information that will often contain many pieces of information in a similar file format, e.g. an email account or a website. The required information is remembered to be contained somewhere within the source's structure. '*I knew that I'd put it on Twitter, I knew what it was, it was to do with maps*' (Participant E).

Time – Typically relates to the time the information was received, the time the information was sent, the time something was due to occur, or when it occurred. These can be broken down into three types of memory:

Point in Time – Knowledge of a specific point in time associated with the information sought e.g. a date or time. - '*I knew what the point in the recording was as I noted several times during the slides'* (Participant J).

Period of Time – A rough approximation of a point in time associated with the information sought. Typically either being a time period the point likely lies between or a broad guestimate of the point in time e.g. a few days ago. '*So I also knew it was within a few weeks or so I knew it probably, likely to be in my inbox*' (Participant D).

Recency – Knowledge of a relative order to other associated information based on time, but not dependant on specific knowledge of a point in time or period associated with the information. '*So I knew the name of the sender who sent me the email with the details in and that was the last email they sent me*' (Participant A).

Location – A physical location associated with the information, most often being the memory that an event was due to take place in a particular location. '*I knew that I'd done it through booking.com and I knew what the hotel was called and that it was in Munich*' (Participant E).

Person – An association made between a person or organisation and the infor-

mation sought.

Association of Known Person – Information or recognition of its known association with a person, for example, knowledge that the information was sent in an email by a particular person. '*That the project manager had sent me a private message through Facebook, containing the project dates, that this was the last message receved via facebook, so would be top of the list*' (Participant A).

Organisation as Person – As in 'association of known person' but the association of the sought information is with an organisation behaving in a manner similar to a real person. '*I knew again that I'd seen an email from the hotel that tells me this information*' (Participant L).

Person or Organisation as Author – The association with the information is of a person in their capacity as the published author of the information sought. '*I was looking for it on Google Scholar, and the words I used were Steve Benford 2012*' (Participant C).

Subject – Relevant keywords or titles pertaining to the subject of the required information are remembered.

Keyword – A known keyword associated with the information sought is searched for. '*I remembered who the sender of the email, and the project name that was likely to be on the subject line*' (Participant J).

Title – A form of title or name associated with the information sought, be it a file name, email subject line, webpage name, or name of a thing. '*I'd remembered that I'd got an email telling me that it had changed. I knew what the email contained and the title of the email*' (Participant J).

As an example of how this coding was used in practice: one participant stated they knew the following about the information they were looking for, '*So I knew who the email was from, it was from my project supervisor. I knew the email address it was sent to, the email account it was in*' (Participant P). Using the coding scheme for this extract the participant implies knowledge of a *Person* associated with the information (the supervisor), a *Source* of the information

(the email account), and a *File Type* (email).

4.4 Results

4.4.1 Information Re-finding Overview

Over the two week period, the eighteen participants (7 Female & 11 Male) recorded 252 diary entries averaging at 14 entries each (min: 10 and max: 32). Of these, 64% were related to personal activities and 34% were related to work activities. Of these activities, 81% were completed using personal computers or laptops, and 17% were completed using mobile phones, with the remaining 2% split among tablets, media players and a voice recorder. Of information sought, 53.9% was re-found within a week of last encountering it (see Figure 4.2) and 76.1% of re-found within one month of the last encounter. The remaining 23.9% of entries represented information that could be considered outside of recent memory, older than a month and possibly more than a year old.

4.4.2 Information Sources

The sources of information that the participants stated that they used to fulfil their information needs are shown in Figure 4.3. In some cases, participants stated more than one source of information as being used to satisfy their information need. These results are therefore are not representative of the location where the information was found, but of the sources used in that process. As might be expected the information sought predominantly involved the use of *Email* (25%) and *Websites* (31%). From these two prominent sources of information there is a considerable gap to the next highest sources of *Office Document* (9%), and *Calendar*(7%) and *Text Message (7%).* The remaining sources are spread thinly at even lower values indicating occasional, or less consistent use by participants of those sources. It is also worth noting the presence of retrieval of information from social information sources, which includes both *Social Networking* (6%) and *Text Messages* (7%) – together accounting for 13% of the information sources recorded – which tend to offer less structured support for information retrieval.

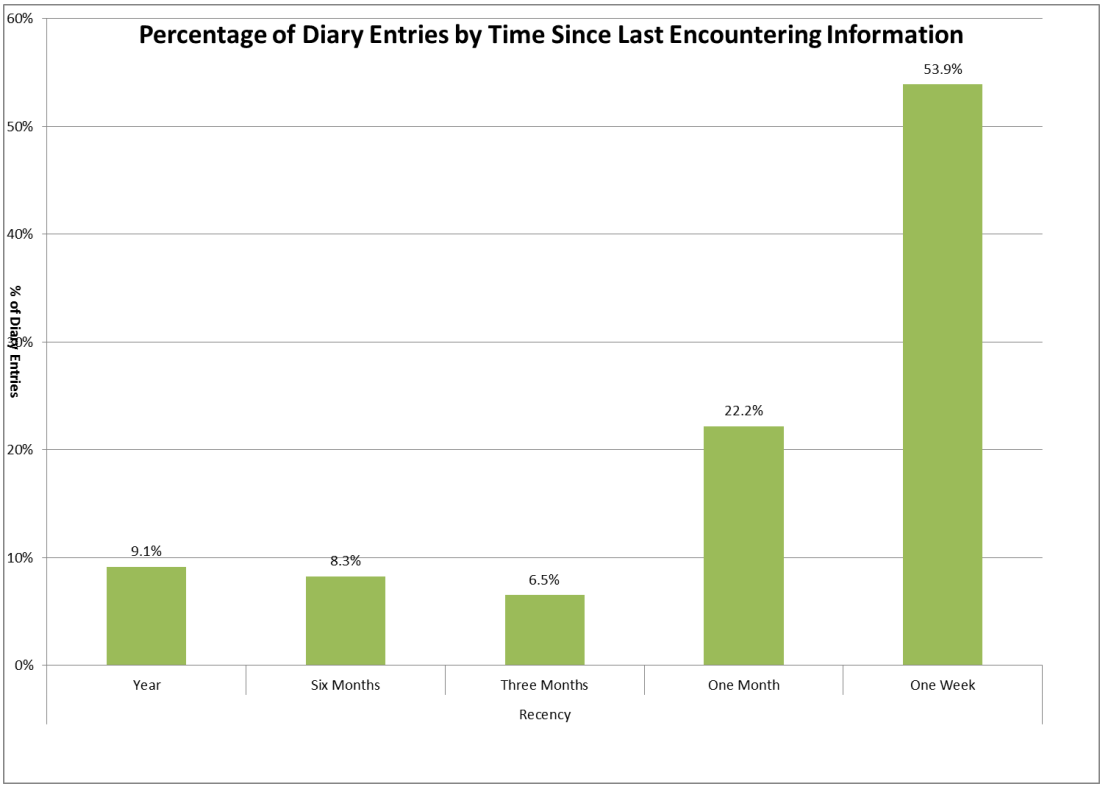


Figure 4.2: Time since last encountering information.

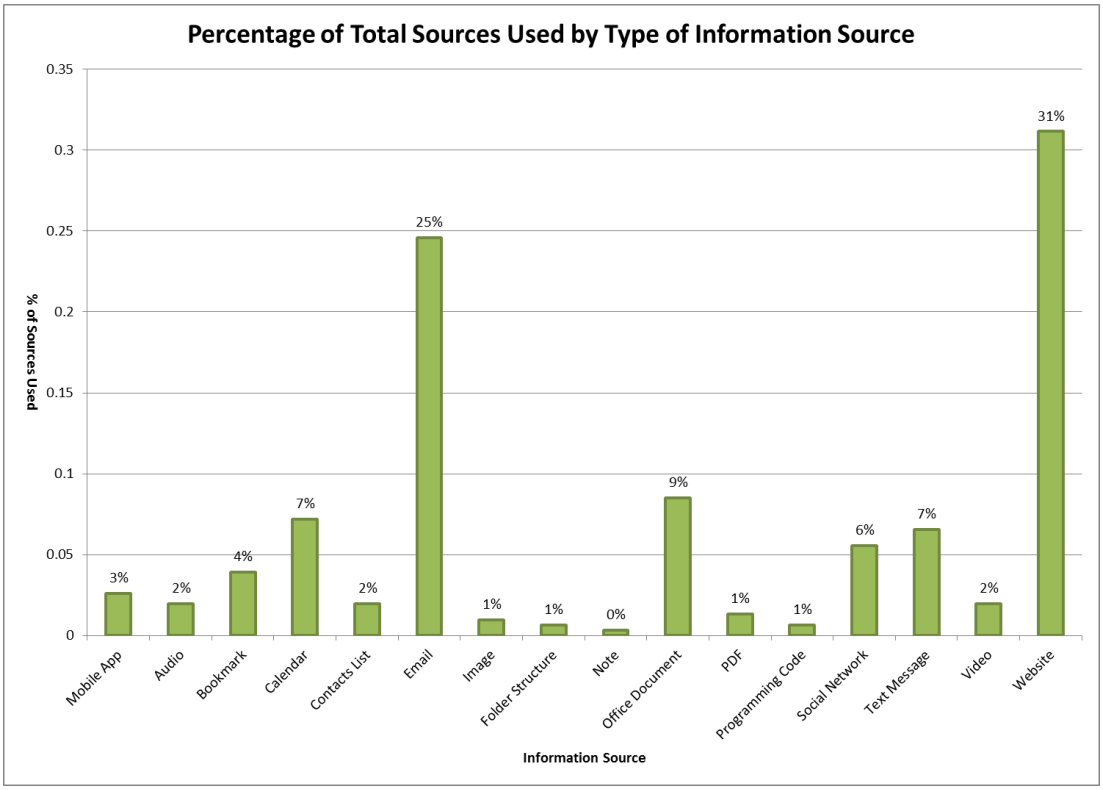


Figure 4.3: Information sources used.

4.4.3 Success and Failure of Retrieval

Participants were successful in retrieving the information they required in 94% of the recorded diary entries over that period. In the 6% of cases where participants were unable to satisfy their information need the reasons given were characterised by the following issues:

- Trade off of the cost of the time spent searching against the perceived benefit of locating the required information.
- The information was not in the location it was expected to be located in, participants came to the conclusion the information they sought had been accidentally deleted or misplaced.
- They had been shown the information by another person in the first instance, making later retrieval difficult due to there being no clear memory of the path taken to locate the information.
- They were unable to remember enough meta-information critical to the process of retrieval in the system being used e.g. not knowing suitable keywords or misspelling suitable keywords when using keyword search tools.
- They were unable to recognise when the information is found e.g. not recognising the name of the file the information is contained within.

Of these instances of failure most were cases where the information sought were infrequently accessed, the majority being last seen more than six months ago.

4.4.4 Difficulty

The entries that were completed successfully were generally completed with ease. The average reported difficulty of completing the tasks was very low, with a median difficulty of 4 on a scale of 1-5 (Very Difficult or Very Easy), with only 10.4% of diary entries being reported as explicitly difficult (see Figure 4.4). Of the entries which were considered explicitly difficult (Difficult or Very Difficult), where the information was found, factors that made retrieval difficult were fairly similar to the cases where it was not found, and included:

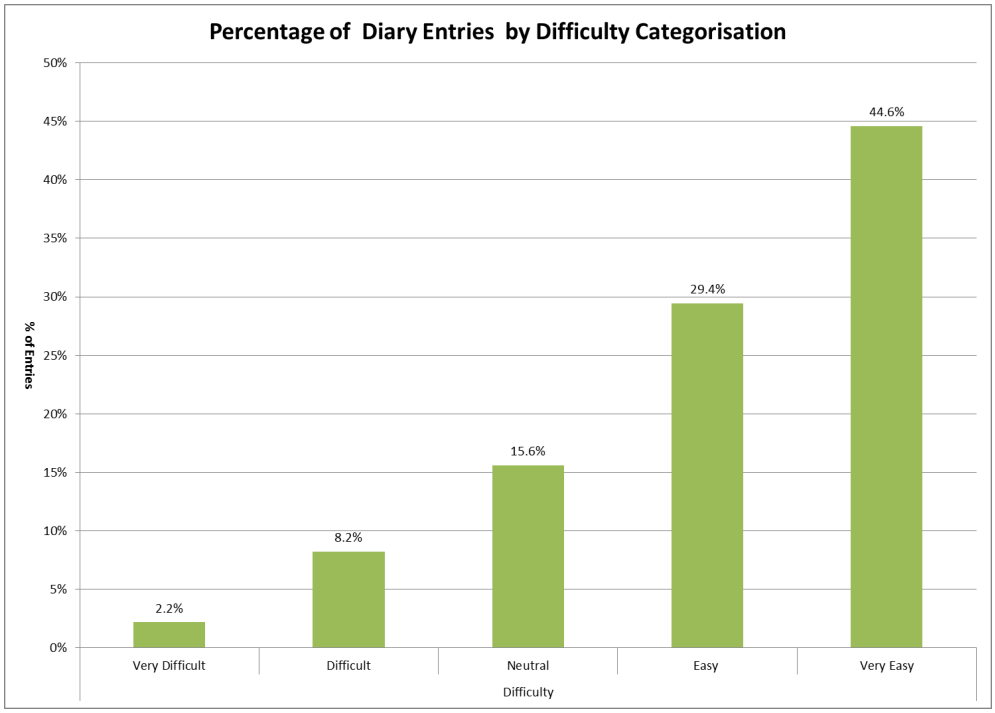


Figure 4.4: Perception of Task Difficulty

- Information is not found within the source it is expected to be within due to misplacing information or misremembering where it should be.
- Not remembering or being unsure of the source of the information.
- Being unable to remember enough meta-information critical to the process of retrieval in the system being used.

These results seem to imply that the vast majority of time people consider re-finding information to not be a challenge, or at least not overtly difficult. However, interviewing revealed that the use of a five-point scale to describe difficulty was perhaps overly simplistic. In some cases participants rated finding information as being easy, but when interviewed, frustration that implied a greater complexity than was suggested by the rating was sometimes evident. One participant, describing an entry that she had rated as easy, indicated that she had used a complex path to the information when interviewed: *'I went to a special folder on my Gmail account where I kept all of them, and I had to go through pretty much each*

one because we'd not been changing our subject lines. But yeah, eventually got it and found it out' (Participant L). While it seems that looking through each email was not viewed as challenging to the participant, her path to the information was not as straightforward as it could have been, and took a while longer to complete than expected. Similarly one participant describes finding information in an entry rated easy as follows, *'I then typed "Alzheimers" into the search bar. I then ensured they were in date order, then scrolled through to find the email I wanted. There was quite a few emails from them so it took me a while to find the one I wanted'* (Participant P). Again, this suggests that it had taken the participant a good deal more time than might have been expected for the task to pinpoint the desired information, but that this did not impact upon their perception of the difficulty level involved.

4.4.5 Information Needs & Tasks

As might be expected, the results of this study (see Figure 4.5) are quite different to that of Dearman et al. (2008). Figure 4.5 shows the percentage of entries in each category of information needs in this study compared to those in Dearman et al.'s 2008, which recorded 858 entries over a four week period. While the two studies are not necessarily directly comparable due to the different intents and final categories used, this comparative chart does provide a view of the contrast in emphasis between what forms of information are needed when people address arising information needs for new information compared to those required when addressing information needs that involve re-find information that has been previously encountered. The most obvious contrast is that between the emphasis on *Finding* in Dearman's study (where it was 35% of entries) versus this study (just 3%), which highlights the difference in focus of the studies. Finding to discover the existence of people, organisations etc. is important when new information needs arise, in the case of the research conducted in this chapter the existence is of course already known, as such very few entries involved locating something new as part of the search undertaken.

By contrast the research in this study showed a much greater emphasis on *Events* (26%). Much of the information sought was for the purpose of remembering the

details of when meetings and other forms of formal and informal events were occurring. This makes more sense when taking into consideration the categorisation of the tasks being performed (Figure 4.6), where Checking was by far the most commonly performed task, accounting for 35.6% of entries. These entries most commonly consisted of participants confirming that they had correctly remembered time-sensitive information. Another notable difference between the two studies is the difference in the *Availability* category (21% for Dearman and 5.9% for this study). Again, this is likely to be down to differences in focus between the studies; re-found information is less likely to relate to time-sensitive circumstantial questions that are about establishing the availability of services or people. Many of the other categories show comparable levels of occurrence between the two studies. *Organisation*, *Guidance*, *Offerings*, and *Trivia & Pop Culture* each represented between 5% and 10% of the queries. The only exception to this being a difference of 8% between the two studies (10% in this study) for *Person*, reflecting a possibly greater emphasis on contact with people when re-finding. Many of these entries involved looking up contact details and birthdays.

4.4.6 Known Associations & Attributes

Source

The attributes people recalled about the information (Figure 4.7) nearly always involved knowledge of a source. As described by Dearman et al. (2008), the behaviour of remembering and accessing a root source of information, when a direct path was not remembered, was extensively used. Participants would start their search by remembering and navigating to a particular form of information organisation, a source, containing the information they sought, at which point their search behaviour would then change. As the list of source types below illustrates, what constituted the source point was not necessarily defined by the different types of applications that enclosed the information, such as an email client or web browser, but by the form of information organisation that most closely surrounds it.

Reaching this closest point of information organisation would involve use of mem-

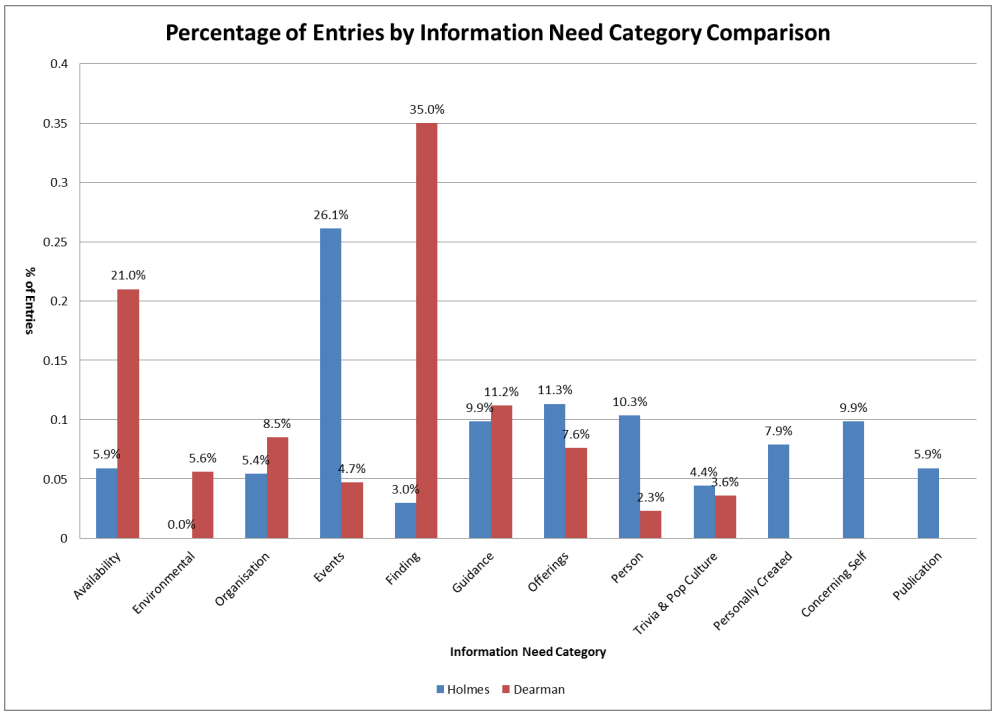


Figure 4.5: Information needs categorisation compared to Dearman et al (2008)

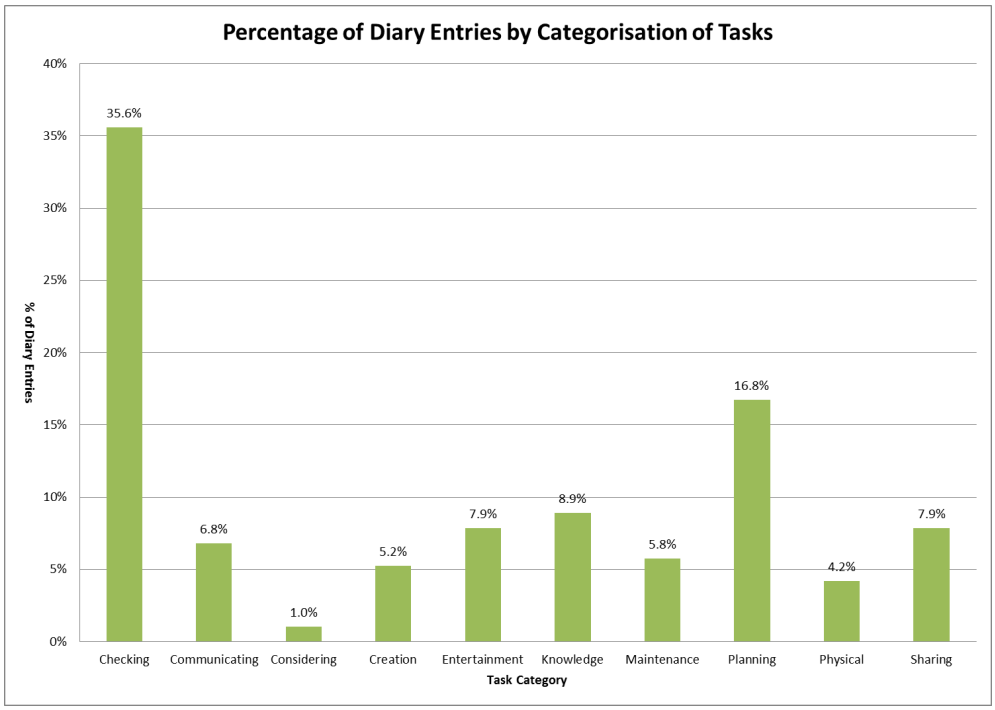


Figure 4.6: Information Tasks Categorisation

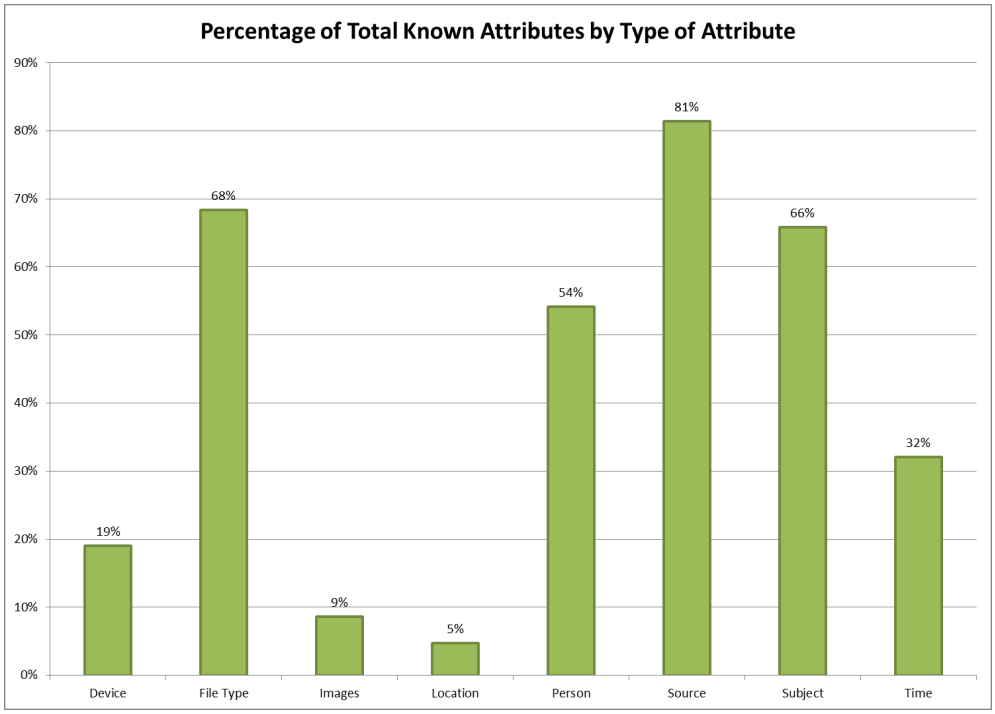


Figure 4.7: Prior knowledge of attributes of the information sought

ories regarding how to access the source. This might require knowledge of which particular *Device* or application the information was stored within, for instance, whether the information needed was only stored on a mobile phone, rather than on a desktop computer. Subsequently, users drew on knowledge such as which particular application would allow them to access the information. During this process of reaching the source participants typically used a familiar path to the source of information e.g. using a web bookmark as a short-cut to reach an email account, or following an often used folder structure.

The root sources that were used in this manner were are as follows:

Mobile Application – '*So I searched on the app "sausage casserole", which threw up loads of different ones, because it's one where people throw in their own recipes*' (Participant J).

Bookmark – '*. . . looking for a section in a paper, and that it was in my bookmarks on Chrome*' (Participant C).

Calendar – '*... birthdays were coming up, that the dates were entered on my mobile phone calendar*'. (Participant I)

Contacts List – '*Just searching through my Blackberry on the home screen because I was looking for a contact*'. (Participant J)

Email Account – '*I used the search button on my phone once I'd selected the correct email account, because obviously it was my university account*'. (Participant J)

Search Engine – '*I was trying to remember how to spell a kind of complicated word, and it's become a very usual way of figuring this, I will just Google it at work, and Google will sort it out for me*'. (Participant F)

Social Network – '*They had been sent to me on Facebook, so I searched through the history to find the video*'. (Participant C).

Text Message Inbox – '*So I had a feeling that she might have mentioned her to me in previous text conversations*'. (Participant D).

Website – ... '*The information was contained in the University of Leeds website, it was actually in her portal*'. (Participant A).

Folder Structure – '*I knew the name of the document, the folder containing the document, the name of the master folder containing folder... I knew there'd be a folder within that folder that contained that document*'. (Participant A)'

It might also be considered that the use of knowledge of a *Source* can involve multiple stages, in that the knowledge of which particular *Device* the information was stored on acted as a prior form of *Source*. For example, in some cases it was clear that there was knowledge that the information needed was only stored on their mobile phone, rather than PC, and that this was in effect the source of the information, but then they also used knowledge of a particular application or folder structure which acted as a Source once the correct *Device* had been selected.

Similar to remembering the *Source*, File type was also nearly always known, perhaps as a result of the participant's knowledge of the source of the information.

The process of getting from the *Source*, to the end point of a file matching the correct file type, was where participants most often lacked some amount of knowledge preventing them from directly navigating to the information. Indeed *File Type* would likely have been recorded in a significantly higher percentage of cases were it not for instances where the information was stored in a form that was not necessarily a file type that could be explicitly stated, or was a subset of a broader form of file e.g. mobile phone applications, calendar entries, and social network entries, where there is no strong association with file types, but are represented as multiple individual elements.

When participants reached the information *Source* it often marked a point of transition in behaviour. Participants would move from following a mostly recalled path to the information, to being unsure of the location and needing to perform some form of search or scan of the interface at hand. The source would be used as a starting point from which to navigate closer in steps to the location of information being sought. This process would typically utilise the other attributes – such as person, time, keywords – identified here as part of the varied strategies used to locate the information, either using them to order or search, or scanning the screen to identify the information with them.

Use of time-based memories to navigate to information involved two common strategies, which relate to the sub-categorisations of *Period of Time* and *Recency*. When *Recency* was described by participants this indicated they were relying on the knowledge that when the information they sought was ordered by time it should be the most recent, or one of the most recent items. Knowing something was recent and could be ordered in this way seemed to act as a short cut, reducing the amount that needed to be recalled in order to locate the information. This behaviour was distinct from *Period of Time* in which a specific or approximate range of time was described as being remembered. In this case the knowledge might form part of a search query date range to limit the information presented, or it might be used to identify an item that might contain the information.

Person was notable as an attribute, because of the attributes that were remembered it was one of the most prevalent, but also most underutilised, aspects remembered in relation to the information sought. The bulk of the cases where

Person was remembered were, as might be expected, email, text message, and social networking based, where communication with a person was a central component of how the information was encountered. As has been noted in other research on email, use of the sender /receiver is a useful filtering technique for locating information (Fisher et al., 2007). When communications (mainly email) were known to be recent, as described previously, participants described scrolling until they saw the first email containing the person's name, *'Because it was so recent I didn't do a search or anything, scrolled a few down, saw the name of the sender, opened it up.'* (Participant A). While for many people *Person* might have been a primary attribute they looked for, it seems likely that often recognising the correct information would involve a combination of attributes to confirm it was the correct message, *'So I knew this was quite a recent one. So I knew it was in my inbox and I literally scrolled down until I saw, errm, not sure if I'd have done it by person or subject, a combination of both sender and subject matter.'* (Participant D).

When the communication was not recent or part of an obvious filing system, then search was commonly used. In many cases this was a search for the person who had sent the email, filtering out other emails and then scanning emails. Thus, the behaviour of participants who had filtered emails by contact mirrored the behaviour of those who were looking for a recent email in an unfiltered inbox. A keyword search rarely occurred first when a person was remembered in association with the information, with other cues such as *Subject* or *Time* more typically being used after using person as a filter point, particularly knowledge that it would be the most recent or a recent message from a person.

Images and spatial memory were another area which may have been of greater prevalence than is recorded here. Visual cues are undoubtedly important as feedback in many interface designs, but participants in most cases did not explicitly state that they remembered visual cues about the information. For example, where the use of images indicating that attachments were contained in emails, this was occasionally stated as being a useful cue in identifying the desired email. This low level of explicit reference to visual cues was perhaps due to these being only recognised and used as a support, as part of the contextual process

of recognising the information, rather than having an explicit association with the information. Although references to visual cues were rare, when they were mentioned it was clear that the association with the information was strong, in that there were specific reasons stated for recalling the association. These were instances such as associations with types of events and colours in a calendar.

It is also interesting to note that several participants indicated that their behaviours would change depending on the device being used, *'So it was finding the email. Now which computer was I on, so yes, I was on my computer. So yes, generally if I'm on my phone I tend to, I don't know why, I'll scroll down my emails until I find who I want. If I'm on a desktop computer, and I know that it's not been in the last few days, then I will do a name search. Occasionally I'll do a name search on the phone, but not as often.'* (Participant A). When probed further on this issue the participant revealed that they felt the difference was due to his mobile phone feeling more of a '*faff*' to search on, despite him not being sure it would objectively take longer.

4.5 Discussion

4.5.1 Limitations of the Research Method

The use of the mobile phone diary study application had mixed results. While the diary study did not force diary entries to be immediately completed, and the note taking functionality was designed into the application and explained, some participants described during interviews how they made a note of the entries on paper before recording them on the phone. The implication of this is that manual note taking was still perceived as a less taxing route than direct input into the application in the first instance, despite attempts to make direct input of diary entries a simple, accessible process. More importantly for the results of this study, four of the participants explained that there were one or two tasks that they had performed several times, but had recorded only once (giving an indication later of how often these events occurred). Participants cited a desire for brevity and to avoid repetition as their reason for recording tasks in this way. These repeated tasks were behaviours such as a researcher looking up papers

for a literature review they were writing or a person who was repeatedly finding their 'to-do' list. This omission of some repeated tasks is important to note, as it suggests that some of the facets of the information behaviours described here may have been under-reported.

4.5.2 Information Needs and Tasks

In general, the results indicate that typically re-finding tasks consist of checking the details of events, offerings, other people, and themselves to confirm that information is correctly remembered or to aid in future planning. Within these broad categories, it was possible to identify a wide variety of differing tasks and needs that led to re-finding of information. The majority of these re-finding events occurred within a month of last encountering the information, and with a large proportion of those occurring within the last week of encountering the information.

Similarly, the sources of information used in performing these tasks were dominated by web and email, but were otherwise scattered over a variety of sources. The tasks identified here in some cases broadly correspond to those identified by Kellar et al. (2007) and Brown et al. (2000). *Knowledge* in this case corresponds to the task of 'information gathering' where the aim is to research some broad topic, *Entertainment* is in some respects similar to 'browsing' where the aim is to be entertained or informed, *Maintenance* has common ground with 'housekeeping /maintenance' which is about routine monitoring or updating resources or profiles, and *Communication*, which corresponds almost directly with 'communicating', which involves keeping in touch with or monitoring contact with people.

The information sources that participants reported using to locate the information they required were broadly similar to the results found in Dumais et al. (2003)'s study of their Stuff I've Seen personal information retrieval system, with regard to the types of information sought – in terms of sources used – despite their study only examining retrieval at workplace desktops. As in this study email messages (76%) and web pages (14%) were the most common sources of information, and

then other file types made up the remainder through many low percentage results, with Microsoft Office documents being the most common of these. It would seem that despite the passage of a decade that brought with it a number of technological changes, workplace information tasks are completed in largely the same manner.

4.5.3 Known Attributes and Information Re-Finding

Most of the information needs in this study were satisfied in a manner similar to the strategy of *Orienteering* described by Teevan et al. (2004). Participants would navigate to the information sought in a series of steps, typically remembering known paths to follow to the information organisation (source), which most closely surrounded it. From there they would recall a variety of meta-information, around which particular pieces of information would be used as cues to locate the information, using a variety of strategies, such as search, scanning, and ordering to navigate closer to the information.

As was evident in Teevan et al.'s (2004, p.417) re-finding study, being able to *'... associate information with a source was critical in helping participants Orienteer to their information target, as participants often remembered a lot about the source'.* They report that people would often remember the source of the information, and from this point be able to recall a substantial amount of meta-information about the required information. This study demonstrated that the source was remembered in 81% of entries, supporting Teevan et al's conclusions. Capra and Pérez-Quiñones (2003) also note this strategy, describing it as an iterative two stage process where people first attempt to locate the information source, and then begin a process to find the information they want within that source. Capra and Pérez-Quiñones (2003, p.8) suggest, at least in the case of the web, that information tends not to be '*remembered by the addressability of the information, that is, not by the path needed to reach the information, but instead by a process of how to reach the information*'. In this study, in instances where participants were shown information by others, but did not go through the process of locating it themselves, those participants then later struggled to, or could not re-find that information. This is illustrative of this principle. It highlights

the absence of memories that can be critical when re-finding information.

The coding of the participant's prior knowledge of the information they sought (Figure 4.7), describes the facets of related information which were essential in the second stage process of utilising contextual data to navigate to a goal. The coding provides an insight into the types of cues, and their prevalence, that people use and focus on as part of their tactics to locate information as part of wider strategies of navigation, and suggests an emphasis on the use of time, person, file type, and subject, once the source of the information has been located.

While the intention in this study was to take a broad look at people's information behaviour, it is important to note that the figures presented here, concerning what information people re-called before beginning the process of retrieving their information (see Figure 4.7), are biased in their emphasis by the types of information being retrieved (see Figure 4.3). As is revealed by the results concerning the sources of information utilised in satisfying the information need, there is a clear bias towards certain forms of information in the results. With email and calendar based information being two of the most commonly used information sources, it is then not surprising to see that many of the queries focused on utilising people and time to locate information, as these are forms of information that are naturally tied to time and lend themselves to those search strategies. However, despite that caveat, these attributes were often present across many other sources, and clearly important in a variety of contexts.

Indeed the identification of these as central cues in locating information corresponds to other studies, which also recorded their use. Elsweiler and Ruthven (2007) identified when studying email that people would usually remember four pieces of meta-information about an email they were looking for. These included the reason it was sent, the topic of the email, the person who sent it, and the approximate date. Similarly Blanc-Brude and Scapin's (2007) study on document retrieval demonstrated that people will most often remember the time, keywords, associated events (context), location and format of documents they are searching for. As has been demonstrated (Elsweiler and Ruthven, 2007) in memory research more generally, people remember what the context surrounding information is much more accurately than they do the content itself, often only

retaining the *'gist'* of what it was about. For example, in a similar, but different, vein that Jaimes et al. (2004) find in a study of work meetings that people would find it difficult to recall much of the content of meetings beyond knowing the major topics discussed, after some time had passed, but that they could recall rather well the persons present, the layout of the room, seating positions, and the roles of people there.

As Elsweiler and Ruthven (2007) and Teevan et al. (2004) suggest, these orienteering and cue based strategies are likely relied upon as they reduce the cognitive load on a person when remembering and retrieving information. Teevan et al. (2004, p. 418) suggests that orienteering saves people *'from having to articulate exactly what they were looking for and by allowing them to rely on established habits for getting within the vicinity of their information need, thus narrowing the space they needed to explore*'. It is clear from this study and a variety of other studies discussed here that the context in which information is encountered and then stored in memory is much easier for people to recall than the actual location of the information and its details. People instead seem to rely on knowing the '*gist*' of the information they are trying to re-find, and utilise the stored contextual meta-data information systems store about it to locate it. Indeed within psychology there has been a wide variety of research demonstrating the importance of this relationship between context and memory in the process of storage and retrieval of memories by people (Elsweiler and Ruthven, 2007) .

4.5.4 Re-Finding and Person

Analysis of the data showed that participants commonly recalled connections between information and people, and made connections between the two in a variety of forms. These results suggest that the memories of people in relation to information are prominent and can be important in order to perform re-finding and possibly other informations tasks. It was at the same time notable that participants were often unable to make use of their recall of people for the purpose of re-finding information. Primarily participants were only able to utilise their recall of links between information and people in relation to communication based information. In these cases participants were able to use recorded meta-information

about people during the process of re-finding that were searchable and had clear roles as a sender, a recipient, or author.

Beyond these three roles, further links were also made between people and information in a variety of other instances throughout the study. These included where information needs arose requiring information about people, when information was desired to be shared with someone, when information was shared with the participants, when information was required about or concerning participants, when events and meetings occurred, when people were seen as a source of information, and when connections between people and information were noted as being known by participants, but not recorded as meta-data.

In a study with comparable results to this study, regarding the information forms used, Dumais et al. (2003, p. 32) suggest that the importance of *People* is *'highlighted by the fact that 25% of the queries involved people's names, suggesting that people are a powerful memory cue for personal content'*. Yet arguably the use of people is under-exploited as a connection, with few systems lending themselves to utilising that information outside of communication systems. Although there might be a person connected to a piece of information that is recalled, this is often not digitally recorded as meta-data that can be presented to, or utilised by, the user (Sawyer et al., 2012). The reasons for this are obvious: social relations can be complex and context dependent, and difficult to capture in any case. Even within the context of email use, where a primary strategy in many re-finding instances was to search for a person and navigate to the information from that point, using this connection was not always well supported by the results of a search. The use of relationships between people and information therefore have the potential to be utilised to a greater extent.

Knowledge of various connections between information and people were still valuable, however, even when data recording a direct link between the information sought and a person connected to it was not available. Knowledge such as this could be utilised in less direct manners, for example, in one instance a participant found it was difficult to locate a piece of information, because they had been physically shown that information by a another person, and had not directly accessed it themselves. They were able to resolve this by going to a source of information

related to that person that they had knowledge of – a blog that person maintained – in the hope they might have published the information sought there. And in this case was able to successfully locate the information sought by utilising this another connection made between a person and the information.

While there are many interfaces which focus on the use of dates, keywords, and hierarchical navigation to locate information, there has been very little work focusing on the use of people as a primary means of accessing or understanding information. In an attempt to address this lack of data about our social interactions and exploit the associations between people Sawyer et al. (2012, p. 886) developed a system that recorded the physical social interactions associated with the creation of information, such as who was in the room when a document was recorded, which demonstrated to some extent the value of this information for later retrieval. They suggest that, '*As the volume of information that we use daily grows, our social interactions may act as an intuitive handle to support how we think of and access information*'.

4.5.5 Difficulty and Recency

Many of the difficulties that participants encountered when re-finding information related to either the *Source* of the information or meta-information critical to the process of identifying it. The broad picture of information re-finding provided by the diary entries highlights how in many cases when difficulty was encountered it derived from instances where participants struggled to remember or locate the source of the information. Commonly this was because they could not recall the *Source* or they had never followed the process of locating the information in the first instance. Another primary reason was lacking memory of a piece of meta-information critical to the process of searching for and / or identifying the information. If considered in terms of (Capra and Pérez-Quiñones, 2003)'s two stage description of information re-finding, difficulties related either to a memory deficit in the first stage – locating the source of the information – or in the second stage – navigating within a *Source* using the recall of the cues, such as those identified in this study. Being unable to recall either of these components would make it difficult for the participant to progress as they intended. As an example

one participant struggled to locate information because of the following:

> *I couldn't remember what the topic of the forum post was, so I tried searching for the, what the, what the actual website that contained the link was. Because that wasn't the title I couldn't find it... I did find it eventually, but pretty much by chance... So I kind of searched the topics surrounding what I thought it might be to do with and eventually found it. But yeah that was a little bit convoluted because I couldn't remember enough information that surrounded the link, so what the forum post was about, that kind of thing.* (Participant E).

In this instance a participant knows the source of the information, the website the information was originally encountered through, but cannot recall enough associated information to browse from there. Critically this included the title of the forum post, suitable keywords to search with, and the website name the information resided within.

However, the results indicated that the majority of personal information tasks people perform are relatively simple in nature and that most of the time these are satisfactorily completed even when more complex to complete than might be desired. In the few cases where an information need was not met the reasons were mostly due to having never personally accessed the information, the information not being available at all or not being accessible due to an external constraints such having a limited amount of time to dedicate to the search.

In contrast a number of studies (Blanc-Brude and Scapin, 2007; Teevan et al., 2004; Boardman and Sasse, 2004) have reported that people often have difficulty in retrieving their documents, when they are asked to find them. In Blanc-Brude and Scapin's (2007) study, 8 out of 25 documents sought posed difficulties for people when they were asked to locate them, often because people were looking in the wrong location for the required document, a difficulty also encountered in this study. Despite studies suggesting that difficulties are common some studies also concur that finding documents is not perceived as a particularly challenging problem. In the Keeping Found Things Found studies Jones and Anderson (2011) report that people were generally good at retrieving the information items they

desired, but struggled much more with remembering what relevant information they had previously seen in order to be able to search for information in the first place. This difference in outcomes regarding difficulty may be down to studies taking a broader cross-section of information re-finding tasks or more naturalistic approaches to querying people about their information needs, as was the case in the study described in this chapter.

This study that revealed many of the types of queries were repeated and simple, to very recently accessed information, and therefore more quickly and easily retrieved by participants. In some studies where greater difficulties have been found in re-finding information a more artificial selection was used, such as in Blanc-Brude and Scapin's (2007) study, where tasks were set to infrequently accessed information, which would imply a greater inherent difficulty of the tasks set, due to the time passed and greater likelihood of memory loss. In Elsweiler and Ruthven's (2007) study of information re-finding on web and email a relatively similar distribution of time between accessing information was found. Using the metaphors of Hot (within a week), Warm (within a month), and Cold (older than a month) information he identified that instances of re-finding were 39.32% Hot, 23.30% Warm, and 18.69% Cold. This compares with this study where using his breakdown 53.9% was re-found within a week of last encountering it, 22.1% of within one month, and 23.9% within a month or more.

As in the study described in this chapter Elsweiler identifies that the majority of instances of re-finding occur within a short period of the last time they were encountered. Yukun et al. (2010) suggest that people need specific support when re-finding older information. They attributed this need for greater support to a result of the increasing difficulty due to information falling out of recent memory as time elapses since last encountering it. They go on to suggest that this difficulty might arise from, in many instances, of people not recalling the information they are searching for based on date, but having few other ways of expressing the need for information not seen for a long period of time. As was identified in this study and the literature review (Chapter 2) time based retrieval can be a valuable method of retrieval, but it is only one of several commonly recalled attributes that might be used. As this thesis argues, providing increased support

for retrieving information by improving the range alternative means of accessing information, such as in terms of people could therefore be a valuable for supporting information re-finding tasks. Yukun et al. (2010) further identify a lack of computing experience, propensity to file information, and not having knowledge of who an email was sent from as leading to a perception of greater difficulty in retrieval.

The results of this study suggest that the actual perceived difficulty of performing information re-finding tasks is low overall, often even when the information is not immediately located. However, the 10% of tasks that are considered outright difficult, that involve retrieving information made complex due to factors such as infrequent access and important navigational cues being forgotten or unusable, are not negligible. This suggests there is still plenty of room for beneficial improvements in aiding people to perform these tasks. Further to this the results suggest there is a need to assess and address the more subtle variations in the difficulty people experience with locating their information. More complex measures of difficulty that address different elements that contribute to difficulty and an understanding of how perceived difficulty differs from task difficulty would enable this. This might be identified via a more detailed multi-dimensional measure of task difficulty, taking into account subjective aspects such as frustration and complexity, similar to the NASA-TLX (Hart et al., 1988). A greater understanding of difficulty would provide a means to identify and address some of the less overt issues people have when re-finding information.

4.6 Conclusion

Over two weeks a mobile phone diary study was conducted that recorded a broad cross-section of daily information behaviours that occurred across all of the devices, applications, files, and Internet services used by the participants to re-find previously encountered information. The study aimed to provide a broader level of knowledge about the make-up of re-finding tasks. To this end, the diary entries and follow-up interview data were used to develop three analytical frameworks, describing the make-up of different facets of information re-finding, and which

in turn were used to categorise the diary entries. The analytical frameworks addressed: what information was being sought (information needs), why it was being sought (information tasks), and how the information was sought (information re-finding process).

Some studies report that information re-finding tasks can be difficult, but in this instance it was found that the reported difficulty of performing a large majority of these information tasks was perceived as being very low to neutral. Tasks that were considered out right difficult were less common, representing 10% of tasks. That said, difficult information re-finding tasks are not a negligible percentage and providing further support to addressing these tasks would be valuable. Further to that, at times the participants' descriptions of their experiences performing tasks was at odds with the low difficulty ratings they assigned. This highlights that more wide-ranging measures of difficulty - taking into account factors such as frustration - may need to be considered to address issues people have when re-finding information more precisely

From the results of the categorisation, the make-up of information re-finding tasks is identified as being distinctly different from those of finding. This difference manifests itself in terms of both the process undertaken and the types of question being addressed. Information re-finding is dependant on recall of memories relating to the source of information found, and then attributes such as time, keywords, and people in order to navigate to a piece of information. This contrasts with finding new information, which is often considered in terms of iterative search processes until an information need is considered fulfilled by locating new information (Belkin et al., 1993). The information needs being satisfied when re-finding information were most commonly details of events, such as meetings, which equated to a quarter of all needs. More broadly, information about offerings, guidance, other people and theirselves was also commonly sought, equating to roughly a tenth of all needs each. This contrasts with information needs when (first) finding information, which generally relate to discovering new information, often about the availability of services or people (Dearman et al., 2008).

The focus on re-finding event based information can be explained by examining the types of tasks participants were performing when the information need arose.

When re-finding occurred the goal of the task was most commonly related to future planning or checking information was being correctly remembered. This contributed to an emphasis on recalling time-sensitive information relating to upcoming events. It also highlighted that it was common for much of the content of the information being sought to be remembered and that re-finding it provides a form of protection against mistakes due to memory lapses. Other tasks that were commonly performed included extending their knowledge e.g. research, creating new information e.g. writing a paper, and sharing information.

What information supported participants in the process of re-finding information was explored by prompting them to describe what they had recalled before re-finding a piece of information and then the process they undertook to locate it. As in Teevan et al (2004), remembering the wider source the information was encountered in was a critical part of the process of locating the required information, and as demonstrated by the study undertaken in this chapter, it was the most frequently recalled attribute about the information participants were looking for. Lacking a memory of the information source was one of the main reasons some information tasks were considered difficult. Participants rarely went directly to the information they sought ('teleporting'), and in most cases would navigate first to the form of information organisation that they believed the information was contained within. From there four commonly remembered attributes (person, subject, file type and time) were remembered and used to filter, navigate to, search for, or identify information in a series of steps, activities Teevan et al. (2004) describe as 'orienteering'. These attributes are clearly important both as contextual cues and as part of search strategies in re-finding information. This memory-dependant process of navigation, when re-finding, is very different from the exploratory forms of iterative search refinement often associated with searches for new information in response to an information need.

Of the attributes primarily recalled by participants about the information they sought, that of person, as the literature review (Chapter 2) discusses, is highlighted as being a potentially under exploited cue for recall, as well as having value for other information tasks such as information sharing. Information is typically situated within a variety of complex and simple social interactions, and

participants had a strong memory of people related to a piece of information. This study has highlighted this complexity, not only illustrating the salience of recall of people in relation to information, but also the range of instances where relations between people and information occur. This includes information to be shared with other people, people as sources of information, information authored by other people, information about the user theirselves, as well as people in relation to specific events such as meetings. While attributes such as subject and time are commonly exploited by interfaces as a method of organising or re-finding information; beyond email, text messaging, and to an extent social media, the links between people and information are utilised to a much lesser extent. Providing increased and improved support to using this attribute could contribute to reducing difficulty and frustration in performing information tasks. Indeed, with a few notable exceptions little research has utilised people as a focal point in interface design to access information, and even less has tested whether it is beneficial in this role. Considering the prevalence with which people were recalled in relation to information being sought, and the range of different possible relationships between information and people that occur, integrating some focus on the person, and the relationships between people, may provide another useful perspective with which to retrieve and manage information. It is clear that there is a role to play for integrating a focus on links between information an people and the relationships between people into future interface designs in order to provide better support for personal information tasks.

In order to demonstrate what value interfaces that use people as a focal point of organisation of information and interaction may have in supporting information tasks, further research into how to represent and organise people in those interfaces is required. This requires understanding of how people or contacts are thought about in order to identify suitable organising criteria of personal contacts, consideration of the role of contact groups in organising and making contacts manageable, and investigation into how contacts and contact groups should be represented to facilitate information tasks. In the next chapter (Chapter 5) research is conducted in order to begin to address these challenges for 'person-centric' interface design.

Chapter 5

Contact Grouping: Categorisation and Representation

5.1 Introduction

The increasing ubiquity of a variety of forms of digital communication – from the now familiar text message and email, to an increasing diversity of social media platforms – for many people, means they now maintain a wide variety of digital contacts from all areas of our lives. Research into both offline and online social networks suggests people are capable of maintaining relationships with a mean of around 150 people (Dunbar, 2016), while people have many additional contacts which come and go over time with whom they have information that may need to be managed or retrieved (Fisher and Dourish, 2004; Viégas et al., 2004). From a sample of 149 million U.S. Facebook users Ugander et al. (2011) put the average number of 'friends' or contacts for each user at an average of 214. Providing some order to the wide variety, and often large numbers, of digital contacts that people retain can therefore be beneficial when managing or re-accessing information associated with them (Sawyer et al., 2012) or share information with them (Amershi et al., 2012). This need might be well served

through alphabetical ordering or by search in many cases, but depending on the task, it may also be beneficial for our contacts to be organised in other ways, such as facilitating the user's management of privacy settings on social media websites by reflecting the users mental model of groups of their contacts(Mazzia et al., 2012). As the literature review (Chapter 2) highlighted, two common approaches to this are to order contacts according to our behaviour with them (Fisher et al., 2007), or as a reflection of the structure of our social networks (Jones and O'Neill, 2010).

In the previous chapter the process of information re-finding was examined with the aim of understanding what re-finding tasks are difficult to perform – to identify where greater support might be provided to that process – and to understand what factors are critical to performing re-finding tasks. From this the role of recalling people who are connected to the information being sought was identified as being a prominent part of the process that was often unable to be exploited as part of the task being performed. Once the source of the information had been accessed (e.g. an email client) participants most commonly recalled people linked with the information, the subject of the information or important keywords, and the time the information was first encountered (often in terms of relative events). However, only the latter two could typically be exploited as part of that process outside of email clients. As is discussed in the literature review in Chapter 2 strategies for re-finding that utilise keyword and identifying information through its relative position in time have been explored to a greater extent than the role of links between people and information. And as highlighted in the information re-finding study in Chapter 4 providing additional means of accessing and interacting with information could provide improved support in performing personal information tasks. The next chapters focus on examining how information re-finding and sharing tasks can be supported by interfaces which exploit the connections that are made between information and people.

If the associations people make between information and people are to be exploited to a greater degree through interfaces which focus on this aspect then consideration needs to be given to how contacts should be represented and interacted with within these interfaces. The social structures people form have

been widely studied within a range of areas of research, but how our personal social structures and relationships might be utilised to manage our contacts and personal information is an aspect which is yet to be extensively explored. To understand more about how to design for circumstances where interfaces might reflect social structures, this study identifies connections people commonly perceive between their contacts, and explores how they wish to group those contacts. By examining the natural social grouping behaviours of people, this chapter informs the design of user interfaces which manage access to personal information through people as a central component of the interface.

One of the primary means of representing contacts in interface design has been through the creation of contact groups. Grouping contacts together allows for representation of the everyday social structures to be used as means of ordering and managing interaction with contacts. People often communicate to groups of people, need to contact members belonging to particular groups of people, and also often recall information in relation to particular people and groups of people, as found in the diary study (see Chapter 4). As Fisher and Dourish (2004) point out, many tasks carried out by single user tools are collaborative in nature. Documents are created in relation to the activities and information provided by others, designed with others in mind, or to coordinate others. This is what they call 'everyday collaboration'. As a result of 'everyday collaboration' there is also often information that is associated with and holds shared relevance to all members of a group. Representing people in groups also provides another dimension through which contacts can be associated, and used to recall, order, and perform information tasks related to people. Indeed there is a strong argument (Mazzia et al., 2012) that people may wish to group their contacts to manage privacy on social networks (Jones and O'Neill, 2010). While others have highlighted the benefits of being reminded of our contacts and mining them for information (Whittaker et al., 2004), and demonstrated how recording links between people and information can be beneficial when re-accessing information based on particular groups of contacts (Sawyer et al., 2012).

This chapter seeks to address lack of consideration in the literature of how contacts should be represented for performing personal information tasks. In order

to identify how interfaces can be designed that take advantage of the recall of associations between information and people in 'person-centric' interfaces this chapter examines how people perceive relationships and structures between their contacts in order to identify principles for their representation in the design of interfaces for information sharing and re-finding.

5.2 Methodology

5.2.1 Study Design

This study applied two complimentary approaches to eliciting how people think about social structures in order to inform interface design. To examine what forms of meaningful groups of contacts interfaces might incorporate and how they should be visually represented, this study uses free-hand diagramming (McCarty et al., 2007) was used to elicit the groupings participants' believe to be most representative of all their contacts. In this method participants were asked to represent all of their contacts in a diagram, in whatever form they wished.

The free-hand diagrams provided a high level view of how participants model their contacts, helping to demonstrate the types of groupings and orderings of those groupings that people naturally make. The diagrams demonstrate the participant's own perception of their personal network, illustrating their personal view of where co-membership of groups exist, their ordering of groups, and highlighting facets of the form of their connection to the participant. However, these diagrams can also be a simplification of a variety of aspects that underpin the structure of a social network, and represent only the participants' view and recall of their network, not the reality of the true connections within it (McCarty et al., 2007).

This was similar in approach to the social network diagrams described by McCarty et al. (2007). Diagrams such as these in other research have often been akin to Venn Diagrams, taking the form of labelled circles or curved shapes (e.g. McCarty et al., (2007)) that overlap, representing the groupings of people and the overlap between people within these groupings (see Figure 5.2). However,

participants were not restricted, and could represent their contacts however they wished, using styles of their own devising. These diagrams were then analysed using a similar process to Sluka (1996). Detailed descriptions of each of the diagrams were written out, then treated as texts to be analysed, and thematically coded into emerging themes (Braun and Clarke, 2006).

In the second approach, this higher level approach to grouping was complemented by focusing on the specific connections that are made between people. Repertory grid triads (Fransella and Bannister, 2004) were used on the participant's own lists of contacts to elicit constructs that described the specific facets of the relationships that people had with their contacts. Participants were asked for the ten most recent distinct contacts they had communicated with via text message, social media, and email. This generated a total of 30 contacts, which Lubbers et al. (2007) identify as being enough to outline the structure of a personal network. The contacts were then compared in sets of three (triads) by the participant, who was asked to identify what is common between two of them (constructs). This was used to both confirm findings from the diagrams and to explore some of the more tacit or less obvious associations made between people, which people may consider in the process of grouping others, but which would not be used as a broad classification in the contact diagramming exercise.

5.2.2 Participants

Twenty participants were recruited, 12 Male & 8 Female, between the ages of 21 and 45, of which four had taken part in the first study. Participants had a range of occupations, including an artist, a waitress, and a small IT business owner, but the majority of participants were PhD Students, research assistants, and lecturers at the researcher's university.

Prior to taking part in the study, participants were given time to read an information sheet describing what their participation in the study would entail and how data collected about them during the study would be used, made anonymous, and protected. If they were happy to proceed, participants signed the consent form confirming that they understood the conditions of the study and were happy

for the data collected to be used in the research. Participants were compensated for their time with high street vouchers worth £10. The ethics application that was approved by the Faculty of Engineering to carry out this research and the forms participants were provided with are contained in Appendix C.

5.2.3 Procedure

Participants were asked to list the names of the last 36 people with whom they had communicated with in order to generate a list of individual items – in this case social network members – to use in the repertory grid activity. This list of 36 was determined by participants examining the names of the last twelve unique people who contacted them via email, Facebook (or alternative social networking website), and text message. A list of 36 was chosen in order to generate a comparable number of contacts to those used in similar methodologies such as Lubbers et al. (2007). If a participant felt they did not use one of these mediums to an extent that they felt it would be representative of people they contacted recently, they were questioned until an alternative form of communication they used was found, or failing that they were prompted to increase the number of people selected from the other two sources. The objective was not to obtain contacts representative of their use of particular mediums, but rather to gather a cross-section of the participants' contacts from all areas of their lives.

Following this selection of names, participants made 36 comparisons between those contacts using repertory grid construct elicitation triads (Fransella and Bannister, 2004). For each comparison that was made, three contacts were chosen via random selection out of the participant's list of contacts. The random selection was determined through numbering of the contacts, and use of a random number generator to generate three integers between 1 and 36, representing the selected contacts. The three randomly selected contacts were then compared by the participant, who then selected the two contacts whom she or he perceived to most share something in common, and wrote down what this commonality was. Participants were instructed that these selections, the constructs, should relate to their relationship with or perception of the contacts, but not physical traits, such as eye colour, unless that was important to how they viewed the association

between the two contacts . For example, an association between two contacts might be that they met at a festival for people with red hair. Participants were also instructed that they could only make the same observation once, and that any further comparisons that might share the same most prominent commonality would require an alternative more specific or vague point of similarity to be stated. This was done to avoid repetition and force participants to think of more ways in which they might view their contacts.

Having completed the repertory grid method participants were then asked to create a free-hand drawn map, representing the groupings of contacts that they perceive as part of their own social network. Participants were provided with a large piece of paper and marker pens, and then instructed to draw a representation of how they would organise their contacts, the context being these are groups of people they would share information with or who would share information with them.

5.3 Results

5.3.1 Repertory Grid Results

5.3.1.1 Affinity Analysis

To analyse the large number of constructs that the repertory grid methodology elicited, an affinity analysis was performed on the data. Affinity analysis (Beyer and Holtzblatt, 1997) can be useful in identifying patterns and organising them into categories in qualitative data, and establishing relationships between categories. This approach has been used in a variety of systems research to help understand people's experiences and behaviours with technology, including customer service (Beyer and Holtzblatt, 1997), evaluating novel mobile technologies (Ballagas and Walz, 2008), and identifying information needs (Dearman et al., 2008).

The process involved organising the data into themes from the bottom up, drawing on grounded theory (Dearman et al., 2008), to allow concepts to emerge from the data. Five-hundred and one constructs were elicited after removal of

inappropriate constructs and in some cases lack of recent social media communication reducing the number of comparisons made. The constructs were printed out, cut up, and combined into a single pile of constructs. This pile was then gradually sorted into a range of different named piles that represented the theme of those constructs. This part of the process was performed by two researchers cooperating to organise the constructs according to their similarities. As part of this structuring process the two researchers could challenge any of the groupings that had been made, and discussion and restructuring was required to take place until agreement could be reached on the categorisation. As in Dearman et al. (2008) the categories that emerged were then challenged by a third researcher who was shown the categorisation of the constructs after the first attempt to order it was completed. From the input of the third researcher the categories were discussed and restructured and challenged again until a final consensus was reached between all three researchers on the most representative categorisation of the different themes within the constructs.

- **Characteristics** — Constructs that described aspects which characterised a person:
  - Physical Characteristics – Relating to the appearance of the person. Short, Both male.
  - Same Age – Same age or age group. '*Both over 50 years old.*'
  - Shared Name – Noted as having the same or similar names. '*Same first name*'.
  - Personality – Describing aspects of a person's personality. – '*Extremely intelligent.*', '*Both creative and arty.*'
  - Ethnicity or Cultural Identity – Describing the ethnicity or a cultural identity. – '*Jewish Friends*', '*Northern English*'.
  - How Others Make the Participant Feel – Description of how the participant is made to feel by the person. – '*Both make me feel uncomfortable one to one*', '*Both make pleasant company*'.
- **Event** — Constructs describing events engaged in with others:

  - Recurring Event – An event which occurs or is implied to occur on a regular basis or at least more than once with specific person or group of people. – '*Both pop round for tea/coffee*', '*Both give me lifts*', '*Thursday Night Trips Into Town*'.
  - One Off Event – A particular event in the past that does not or is not implied to occur more than once with a specific person or group of people. – '*Watched 2nd Lions Test Rugby Together*', '*Went to Olga's BBQ.*'
  - General Event – Ambiguous in terms of whether it is a repeating event or not.

- **Expertise** – Constructs concerning particular skills held by others:
  - Qualified Skills – Skills or expertise indicated by profession, degrees, or qualifications – '*Engineering degrees*', '*Know how to programme.*'
  - Perception of Skill or Achievement -– Subjective perceptions of another's skill or achievements in a particular area – '*Both mathematically gifted*', '*Stats Gurus*'.
  - Field of Work or Study – Having a particular or similar field of work or study – '*Both do freelance photography.*' '*Rail related work*'.
  - Social Skills — Being able to provide emotional support or advice – '*Help Me to Organise My Life!*', '*Always there for me.*'
- **Frequency of Contact** -– Constructs relating to the frequency of contact between participant and others either physically or digitally – '*Talk Regularly*', '*Saw them last week*'.
- **Interests** -– Constructs relating to particular areas of interest shared between people:
  - Likes or Loves – The extent to which a particular thing is enjoyed or disliked – '*Both enjoy* Deadliest Catch', '*Like being outdoors*'.
  - Shared Interest or Hobby – Particular hobbies or topics an interest

is shared in – '*Both are interested in gallery curation.*' '*Both very politically active*'

- **Location** – Constructs which centre on a shared location:
  - From Same Place – People originating from the same location – '*Both from Oxford*', '*People from Nottingham.*'
  - Live -– People living in the same place:
    - Live In – Living in a city/town/region – '*Both live in Penryn*'.
    - Lived In – Previously lived in a city/town/region – *'Lived in Spain'*, '*Lived in Nottingham*'.
    - Living Together – Living together within the same building – '*Housemates*', '*Live Together*'.
  - Physical Proximity – Being close to another person in a context outside of the home, usually work related – '*Their office desks are next to each other*'. '*In My Office/Room*'.
- **Meeting** – Constructs relating to when, how, where a contact was first met:
  - Know Through Another Person – Meeting a contact through association with another person – '*Met via Rabbi*', '*Know through [Partner]*'.
  - When I Met Them – The occasion or location of first meeting a person – '*Met at Uni*', '*Recently met*'.
- **Organisational** – Constructs relating to involvement in an organisation:
  - Organisation – High level organisation such as the University of Nottingham – '*Work at The University of Nottingham*', '*Colleagues at Redhill Academy*'.
  - Sub-Organisational Groups – Lower level organisation, such as a research group – '*Both work for County Youth Arts – Nottinghamshire County Council*'.

  - Past Work or Study – An organisation where someone worked or studied in the past. '*Both worked at [Education Organisation]*'.
  - Project Together – A project being collaborated on – '*Both have worked on the "Showcase" project*', '*Worked on a project together*'.
  - Role – The role of a person within an organisation or relative to a person – '*Both are band leaders*', '*Human Factors Professors*'.
  - Work or Worked With Together – Working with someone or having worked with someone in the past, in the same organisation – '*I am currently working with both*', '*Worked/studied with all of these*'.

- **Relationship** – Constructs concerning the type or intimacy of a relationship:
  - Friends in... – Friends defined by their being in a location, involved in an interest, in an organisation, and from a particular time in the participant's life – '*Acquaintance in Nottingham*', '*Friends from conventions*'.
  - Intimacy of Relationship – A type of relationship and an implied or explicit level of intimacy of that relationship, from relationships of low intimacy (acquaintances), to high levels of intimacy (close friends) – '*Less "intimate" friends*', '*Close Friends*'.
  - Mutual Friends – Friends who are mutual to particular individuals – '*Both are friends of Andy*', '*Have mutual friends that we all see*'.
  - Positive or Negative Relationship – Relationships that are described in positive or negative terms – '*One really hates the other*', '*Like each other*'.
  - Family – Constructs concerning family relations:
    * Children – The child of a particular person.
    * Relationships – People in a romantic relationship with each other – '*Married*'.

* Relatives – The relationship between two relatives (e.g. Mother – Daughter) – '*Close family*'.

5.3.1.2 Construct Categorisation Results

As a result of performing the affinity analysis, described in the previous section, each of the constructs was categorised according to the type of relationship it described between two contacts. The categorisation (see Section 5.3.1.1) describes the construct in terms of both low level and high level themes.

The high level themes represent the different types of connections that are made between people's contacts, and the low level themes identify a variety of forms through which those relationships can be expressed. Considering both is important, as looking at the data only from a high level would risk oversimplifying the results. Understanding the make up of some of the less common low level themes is as revealing about the results as the overall tallies of the high level themes. For example, the distinction between Roles and Organisational Groups is important to consider when looking at the organisation related constructs, while they are both about organisational relationships, they tell you quite different things about the contacts.

Examining this data first at the lowest level of themes from the categorisation, Table 5.2 shows an extract of the most commonly occurring constructs from the complete data set in Table 5.1. Table 5.2 shows the ten most common construct themes, the percentage of the total constructs each type represented, and the percentage of participant datasets that type of construct occurred within. Reporting the prevalence of construct types occurring among participants, as well as the total percentage of occurrence, illustrates how commonly these construct types were perceived by participants. This is valuable for understanding whether the prevalence of a type of construct was limited to a few participants or was perceived by most. The most common were those constructs referring to Recurring Events (6.03%), which participants held with particular groups of people on a regular basis. However, this was only by a slim margin, with the occurrence of the low level constructs being quite closely spread across the highest results.

Category	% Participants	% Occurrences
Characteristics	**80.00%**	**11.22%**
Physical Characteristics	15.00%	0.84%
Same Age	35.00%	1.34%
Shared Name	20.00%	0.84%
Personality	60.00%	5.70%
Ethnicity or Cultural Identity	35.00%	1.34%
How Others Make the Participant Feel	15.00%	1.17%
Event	**85.00%**	**12.90%**
Recurring Event	65.00%	6.03%
One Off Event	70.00%	5.70%
General Event	15.00%	1.17%
Expertise	**85.00%**	**9.55%**
Qualified Skills	35.00%	1.17%
Perception of Skill	25.00%	1.51%
Field of Work or Study	65.00%	5.86%
Social Skills	30.00%	1.01%
Frequency of Meeting	**35.00%**	**1.84%**
Interests	**75.00%**	**8.04%**
Likes, Loves	50.00%	3.18%
Shared Interest or Hobby	65.00%	4.86%
Location	**85.00%**	**13.07%**
From Same Place	20.00%	0.84%
Live	80.00%	5.53%
Live In	75.00%	4.02%
Lived In	15.00%	0.50%
Living Together	25.00%	1.01%
Physical Proximity	30.00%	1.17%
Meeting	**65.00%**	**4.02%**

Category	% Participants	% Occurrences
Know Through Another Person	20.00%	0.67%
When I Met Them	50.00%	3.35%
Organisational	**95.00%**	**19.93%**
Broad Organisation	80.00%	5.53%
Sub-Organisational Groups	65.00%	4.52%
Past Work or Study	30.00%	1.34%
Project Together	40.00%	1.84%
Role	60.00%	5.53%
Work or Worked With Together	35.00%	1.17%
Relationship	**80.00%**	**14.07%**
Friends in...	70.00%	5.70%
Intimacy of Relationship	60.00%	5.53%
Mutual Friends	30.00%	1.84%
Positive or Negative Relationship	25.00%	1.01%
Family	**65.00%**	**5.36%**
Children	15.00%	0.50%
Relationships	40.00%	1.51%
Relatives	50.00%	3.35%

Table 5.1: Repertory Grid Construct Results

These most common themes of constructs, with the exception of one, all belong to different distinct higher level themes that emerged during the affinity analysis. These were also the most common higher level constructs (see Table 5.1), 20% Organisational, 14% Forms of Relationship, 13% Location, 13% Event, 11% Characteristics, 10% Expertise – all of which were present in the data of at least 80% of the participants.

By far the most reported and common high level theme of construct reported by participants were those of an Organisational nature. These were reported by 95% of participants and constituted 19.93% of the total number of constructs.

Category	% Participants	% Occurrences
Recurring Event	65%	6.03%
Field of Work or Study	65%	5.86%
One Off Event	70%	5.70%
Personality	60%	5.70%
Friends in...	70%	5.70%
Live	80%	5.53%
Organisation	80%	5.53%
Role	60%	5.53%
Intimacy of Relationship	60%	5.53%
Shared Interest or Hobby	65%	4.86%

Table 5.2: Repertory Grid Most Common Themes

The majority of this is constituted by three primary themes of Broad Organisation (5.53%, 80%), Sub-Organisational (4.52%, 65%), and Roles (5.53%, 60%). Arguably Broad Organisation and Sub-Organisational could be combined into a single category, but it was considered during the affinity analysis that the difference in specificity of organisational grouping between affiliation with an organisation as a whole compared to specific groups within a broader organisation was a clear distinction within the data. The same participant would sometimes make broader, or more specific organisational links between contacts. In either case, particular affiliations and the roles people perform within an organisation were among the most salient connections that people made.

The Expertise themes, many of which related to work, were perceived as being separate to the Organisational themes by researchers during the affinity analysis. It was considered that while many of the constructs were relevant to work, they were independent of organisational structures. However, those themes do seem to share a similar basis, in that many describe a form of affiliation, albeit not with an organisation, but with a field of work or ability related to a particular field of work.

The Forms of Relationship theme represented a variety of factors in describing the relationships between people, with both the level of intimacy, type of relationship, and context clearly being significant to how participants perceived connections

between people. Common types of relationship described included Colleague, Friend, and Acquaintance. Friend and Acquaintance would appear to largely imply opposing levels of intimacy of relationship, perhaps with relationships like Colleague lying somewhere between, describing both the type of relationship and possibly level of intimacy. But while it might be simple to perceive that Acquaintance and Friend form part of a scale of intimacy with a person, this was also entwined with the type of relationship with a person.

Colleague, Friend, and Acquaintance can imply both the intimacy of a relationship, and also the context or form of relationship. Colleague usually denoted a specific professional context to a relationship: the term colleague could imply different levels of intimacy, from a purely professional relationship to a friendship with someone whom the participant also worked with. Friend implies a social non-work aspect to a relationship in addition to the implication of some level of intimacy to the relationship. The use of Acquaintance, to imply a lack of intimacy or frequency of association with a person, was slightly different in that it does not provide a context for what activities might be engaged in with that person.

Equally, ambiguity over the intimacy and the nature of the relationship was present in other forms of relationship descriptions. One participant described two people as 'social friends' to describe the relationship they had with people they met when they went out drinking with their friends, whom they were not particularly close to, but who were part of their friendship group. The variety of other words used as synonyms for friend, such as 'mate', serves to create further ambiguity about the nature of the relationships described. Different words for friend could imply different forms of relationship, but the use of these synonyms seemed to be idiosyncratic within this data set, there was not a single casual synonym for friend that was used across multiple participants.

Additionally, the descriptions of the type of relationship would very often be modified by qualifiers to denote more specific forms or levels of social intimacy of the relationship. Words such as 'friend' would be used as a baseline from which to modify either up or down from the perceived intimacy of that form of relationship. The qualifiers were typically words such as 'very close' or 'good', or

even the length of friendship with 'old', but downward revisions were also made using words like 'less'.

Another dimension to their perception of a relationship were various phrases with which to ground the relationship, and provide context to it. These were often related to the 'Friend In...' theme, though not exclusively, providing context with aspects such as location '*Friends from Sweden*', shared interests '*Close Friends in Re-enactment*', and activities '*drinking mates*' .

The Interests theme, while not the most prominent in terms of number of occurrences, was common among the large majority of participants, with 75% of them perceiving some connection based on an interest between their contacts. This would suggest that connections made between people's interests are an important connection where they exist between people. A distinction which would be useful to understand more about the form of relationship, but could not be made from the data due to the ambiguity of the constructs, was which interests were shared by the participant as well as between the two contacts.

One of the obvious drawbacks to the repertory grid methodology is that while the number of occurrences of each type of connection are salient, this does not necessarily relate to the importance of those connections. Constructs relating to Familial relationships were much less common, despite the common presence of 'family' in the diagramming task, but intuitively these are important connections people make between particular people. This can largely be explained due to the lower chance of this occurring within the methodology due to the typically small number of close family members appearing in the 36 contacts.

Other interesting, but less common, constructs include the Frequency of Contact theme. Despite its low level of occurrence (1.84%), and low number of participants describing it (35%), it stands out as being unique among the themes identified in that it describes the participant's own behaviour with their contacts, as opposed to their connection with them or the commonalities between the two compared contacts. Also reported at a similarly low level, How Others Make the Participant Feel, is different to the other constructs in that it describes the emotions of the participant in relation to their contacts. The Perception of Skill (1.51%, 25%)

and Personality (5.70%, 60%) categories similarly share a subjective judgement, which contrasts with the majority of other forms of constructs, which largely deal with factual information. These themes highlight three broadly different forms of connections that people made, which were less prevalent within the data, of personal behaviour, personal emotions, and personal judgements.

When or where participants met others might also be seen as interesting in that it relates solely to a unique memory of the first instance of meeting someone, and is for some people evidently a significant connection. As with the Friends In... theme there was again an element of contextual grounding, and these constructs were often seen tied together with descriptions of Events, Locations, or Organisations.

5.3.1.3 Multiplex Constructs

During the affinity analysis it quickly became apparent when grouping together similar constructs that a large portion of them, 187 constructs (37.32%), were describing relationships that were multiplex in nature, and implied links to multiple overlapping social structures and their temporal nature e.g. 'old friends from Sheffield'. This led to some of the constructs being difficult to categorise, as they would contain multiple facets of the relationships between people. To resolve this issue these constructs were categorised according to what the primary subject the construct was describing. However, the multifaceted nature of many of the constructs was clearly a significant feature of how people think about their relationships. The construct '*Nottingham Artist*' is primarily about the shared role two people have as artists and would have been categorised to reflect that, but the location appended to this implies that their location or origin as an artist is significant to how the participant would group this person. To examine this more thoroughly, each of the constructs were re-coded in a separate analysis (Table 5.3). In this case, each of the constructs were coded to reflect each aspect that constituted the intent of the construct. For example, the construct '*Close Female Friends from Work*' would be reflected in the coding that this related to work, gender, and friends as qualified by their intimacy.

Of these multiplex constructs the most common among participants was com-

Category	% Participants	% Occurrences
Activity	50.00%	19.25%
Characteristic	30.00%	4.28%
Ethnicity	5.00%	0.53%
Event	30.00%	4.28%
Expertise	15.00%	2.14%
Interest	20.00%	2.67%
Intimacy of Relationship	75.00%	29.41%
Length of Time	25.00%	3.74%
Location	45.00%	13.90%
Meeting	20.00%	3.74%
Family	10.00%	1.07%
Role	30.00%	6.95%
Gender	25.00%	7.49%
Skill	5.00%	0.53%
Total Combination Constructs		187
Percentage of Combination Constructs		31.32%

Table 5.3: Table of Multiplex Constructs

bining the intimacy of the relationship (75%), followed by activity performed together e.g. work or an interest, location (45%), and role (30%). Similarly, of the total number of multidimensional constructs included 29% intimacy of relationship, 20% activity, 14% location, and 7% gender. Whilst these were the most common forms of constructs that were combined, a wide variety of other construct types described in the categorisation were combined on a less commonly, to describe the less common forms of social connections that occurred.

5.3.1.4 General Themes

In addition to the forms of constructs that were commonly encountered there were also some less common but interesting features in the data. In a number of cases, participants would use negation of one group to define another group. For example, one participant wrote '*close non-work friends*', defining the connection as being the negation of people they work with from the group of people they consider to be close friends. Here, the grouping is defined by subtracting one

group from another larger group, by defining which connections and context is unwanted.

Another feature of the data was the context of past, present, and future. The majority of the constructs spoke to the present, but in areas such as Events there was an evident variation in the use of all three. The constructs reported in One Off Events would for example describe single events in the past, '*I went to Alton Towers with them both*', or in the future '*Going to ladies day at Southwell Racecourse with both*'. In contrast, the Recurring Events constructs described continuing current relations with people in the present, '*People I go to the pub with*'. Whilst this was most evident in the Events related constructs, this was also present elsewhere, such as describing membership of past organisations, *'worked with them on a project'*.

The methodology did not specify that the construct the participant identified had to be between the participant and the two people being compared. As a result the categories of constructs identified generally belonged to one of two distinct forms of relationship: those relating to the relationship between the participant and the people compared, and those relating just to the two people compared. For example, a relationship solely between the two compared might be '*Both live In Dundee*'. This may or may not be true of the participant as well, but the emphasis is on the relationship between the two people being compared. A relationship between the people being compared and the participant might be '*They live with me*'. A third form of construct was also arguably present, at least as far as this distinction is concerned, in which a construct's intent was neither obviously referring to just contacts or contacts and participant. It might refer solely to the two people in the comparison, or it might relate the participant and the people being compared. For example some participants simply wrote '*University of Nottingham*', making it unclear which people are involved in this connection.

The categorisations of the constructs tended to lean towards either one or the other of these relationships and therefore some of the categorisations reflected what would seem to be a repetition of the same idea, but were inherently different expressions of it. For example, themes emerged distinguishing between two people

belonging to the same organisation, and to two people working with or having worked with the participant in the same organisation.

5.3.2 Contact Diagrams

The process of creating the free-hand drawn diagrams varied between participants. Some participants would immediately start drawing out groups or potential groups to be included, whereas other participants felt the need to list some of the people they knew first to prompt them and properly consider the types of groups that would be necessary. Seven of the twenty participants created multiple versions of their diagram before they were satisfied with it. When deciding to redraw their diagrams participants were questioned as to why they were beginning again. Reasons given for redrawing the diagrams were realising that they had left out an important group of people and would not be able to position that group in the structure where they wanted it to be, being unhappy with the approach they took to organising contacts, and wanting to neaten the diagram or redraw it with greater clarity, and learning how to improve the diagram from their previous attempt.

- Left out people (N = 2)
- Revised approach (N = 4)
- Redrawn for clarity (N = 6)
- Learning to improve diagram (N = 2)

The two participants who created the most versions, four diagrams in total, said they would restart because they were adamant about exactly representing every grouping and facet of the structure of their contacts. The four participants who created more than two versions, also corresponded to the participants who produced the most detailed diagrams. For other participants this was clearly less important, and establishing the main overarching groups of people was a greater concern than modelling their contacts in extensive detail.

As the emphasis was put on the groups of people participants perceived within their contacts, the majority of the diagrams represented only groups and not

individual people. Three of the participants, however, did take the approach of representing individual people by placing names around grouping labels. In two of these diagrams the relative positioning of the names held significance for its interpretation.

5.3.2.1 Diagram Types

The diagrams that participants created in the contact mapping task were clearly divisible into five primary approaches to ordering groups of contacts.The styles encountered and the occurrence of these different styles of diagram were as follows:

- Axis Diagram (N = 4)
- Venn Diagram (N = 7)
- Physical Location (N = 2)
- Spider Diagram (N = 6)
- Grid (N = 1)
- Mixed (N = 7) – Of these diagrams 7 were of a Mixed nature, containing prominent elements of multiple styles.

Axis

A one dimensional – but sometimes two dimensional – axis is central to this style of diagram (see Figure 5.1). Typically each axis denoted the strength of a particular factor the participants perceived, relative to the groups as a whole. These factors mostly stemmed from the participant's behaviour with the groups and were used to highlight important groups, representing factors such as their frequency of contact with groups or closeness of relationship to people in the groups. In instances where the axis did not represent their behaviour, or in cases where a second axis existed, then time was used to denote how long the group had existed, or the currency of the group in terms of its immediate utility to the participant. Groups were represented within the axis typically as ovals, their relative position denoting the strength of the dimensions the axis represented.

Although Figure 5.1 is quite a simplistic series of groupings, it clearly illustrates the features of an axis diagram. In the top left and right are the labels 'High' and 'Low' separated by a double headed arrow, which represent the central axis element of the diagram. In this instance the axis indicates the 'Importance of Communication' of different groups in the diagram. The groups are represented by the red labels which sit along the continuum of the axis. Green lines have also been used in this case to denote the movement of people from their work groupings into the their friends grouping as work had been a source of friends for them.

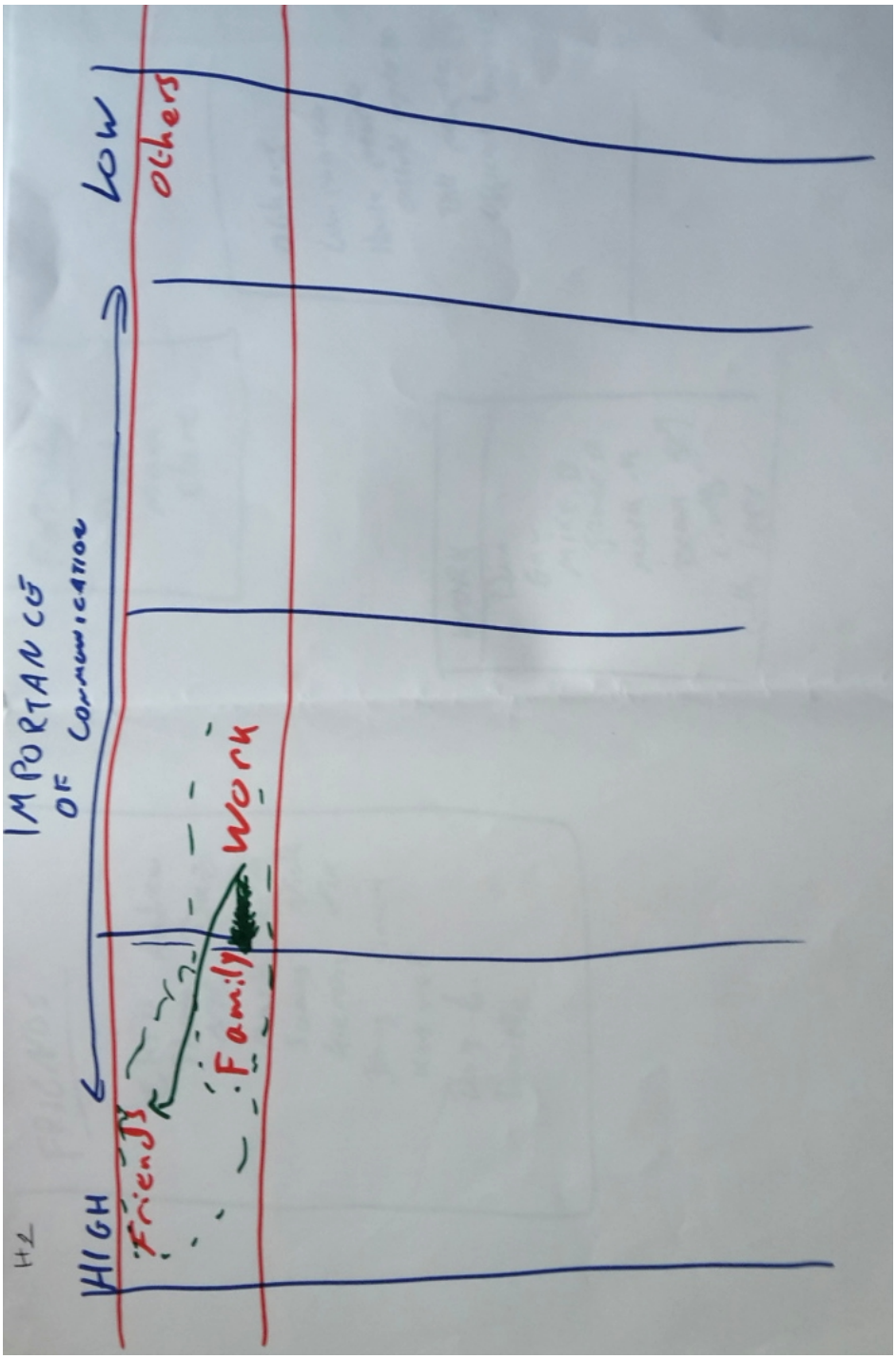


Figure 5.1: Axis Diagram

Venn Diagram

The primary focus is on groups represented as either circles, or ovals, or polygons and the overlaps of contacts between the groups as represented by layers of overlaps. This type of diagram is essentially a Venn diagram of groups of social relations. Within these diagrams it was common for the groups to contain a hierarchy of sub-groups, where each shape might break down into further specific groups through the inclusion of groups drawn entirely within the outer shape, one or more times.

Figure 5.2 provides an example of this type of diagram, where there is a large amount of overlap between groups and many nested sub-groups. The outermost circles denote contacts who are 'Personal' and those who are 'Work' contacts, and the overlap between those two types of contact. The shapes within each of those two groups denote more specific sub-groups within that grouping and overlaps between those sub-groupings. For example within the grouping of 'Personal' there is a grouping 'Close Friends' which contains further groups representing specific types of close friend e.g. those from uni.

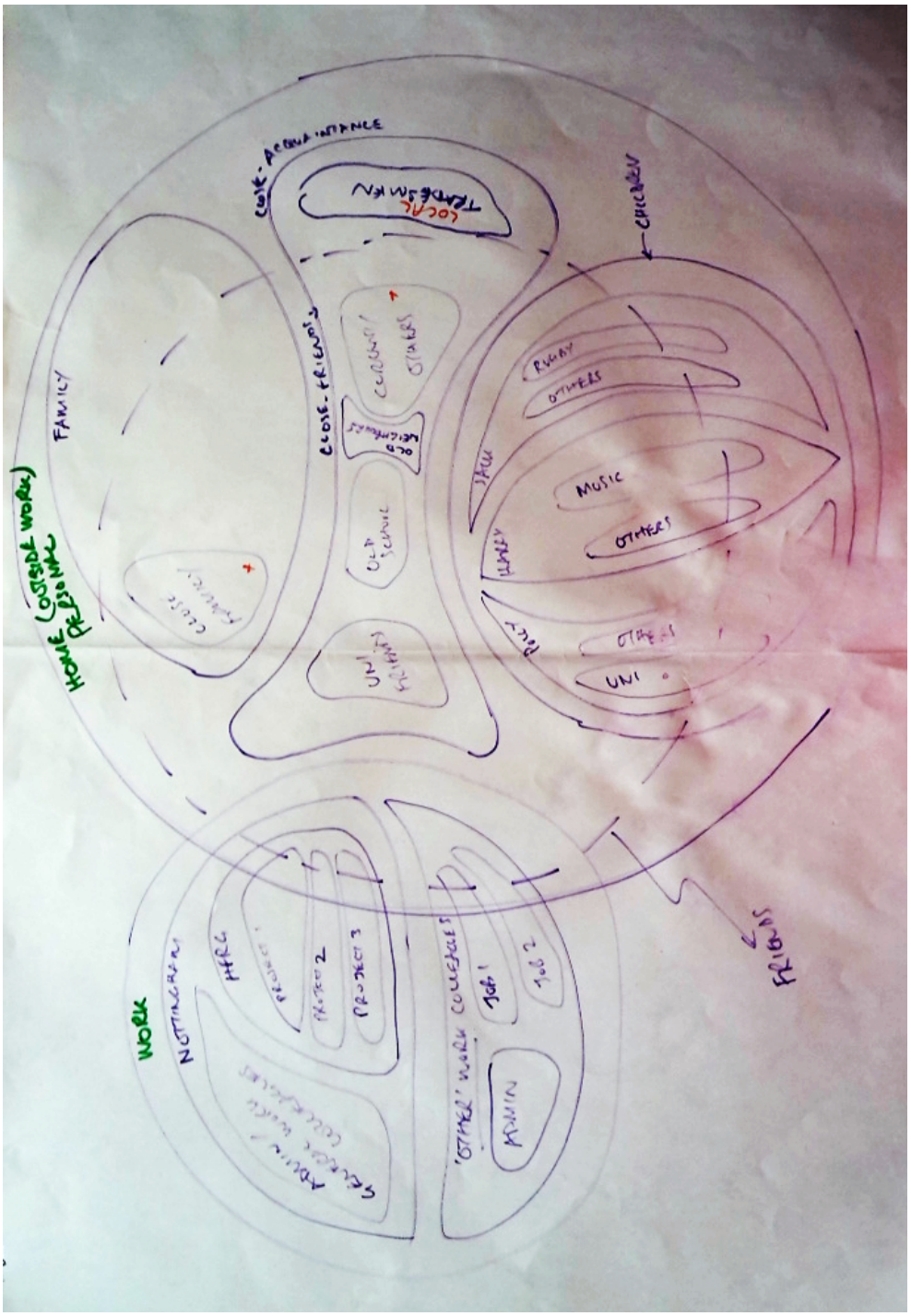


Figure 5.2: Venn Diagram

Physical Location

Labels denoting particular geographic regions – usually directly, sometimes by implication – are drawn at various points on the diagram. Groups are then positioned relative to the labels, their proximity to a label denoting the geographic region to which a group pertains.

In the example of Figure 5.3 primary locations have been represented by large orange text in each of the corners – 'Notts', 'Elsewhere', 'Family', 'Ellis', and 'North East'. 'Notts' (Nottingham/Nottinghamshire) and the 'North East' represent specific geographic areas, the sub-groups of which are primarily various social associations within those areas e.g. 'Creative Writers' or 'Rocksoc'. Elsewhere is a grouping of location in a more abstract sense of being a grouping of all those places not explicitly stated in the diagram. 'Family' contains locations as sub-groups making it a container for locations even if it is not a location itself. Similarly, 'Ellis' represents a single important person to the participant who, Ellis is a point of connection between specific locations and social groups.

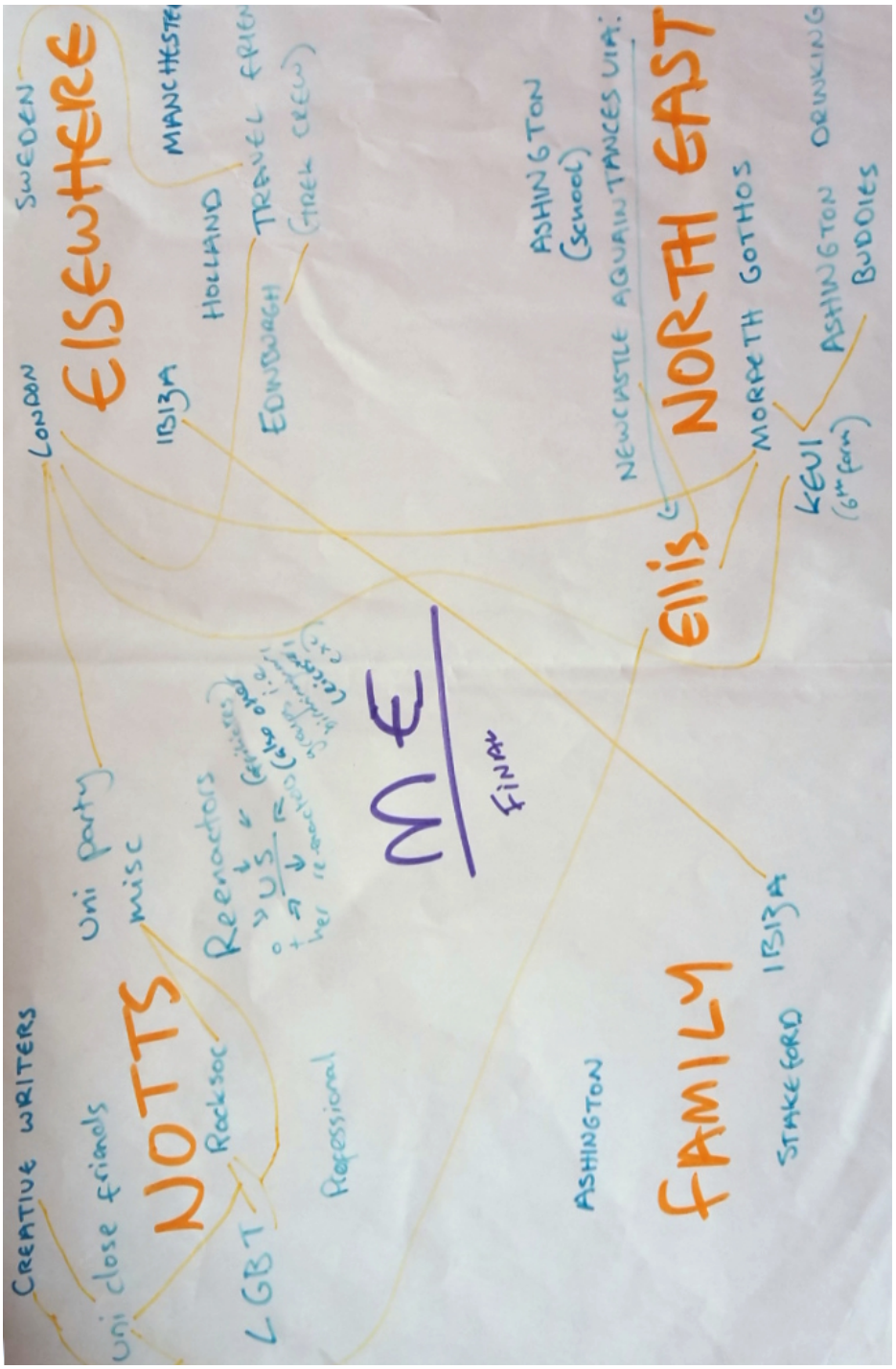
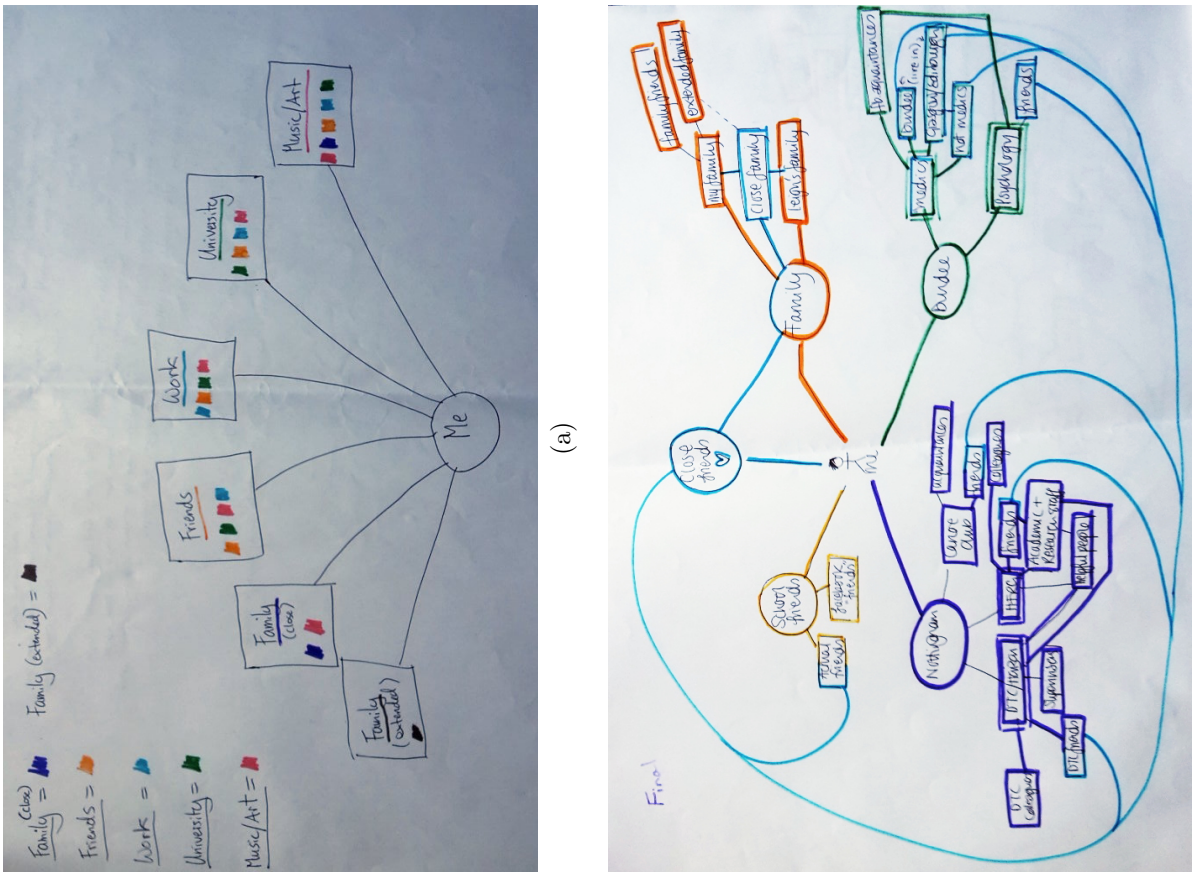


Figure 5.3: Location Diagram

Spider Diagram

Egocentric in style, the word 'Me' – or some other form of self-representative symbol – is in the centre of the diagram representing the participant. From that central point multiple lines span outwards to high level groupings of contacts, which in turn each have further lines extending from them to further sub-groupings. It is common that connections and overlaps between various groups are represented in this diagram by both uni-directional and bi-directional arrows between groups.

The example spider diagrams in Figure 5.4 illustrate a single level (5.4a) and a multi-level diagram (5.4b), which were both common approaches to making this form of diagram. In both cases, groups were represented by enclosing the name of the group in some form of box. In both cases here colour coding has been used to denote multiple memberships of groups akin to the overlaps in the Venn diagram approach. Colour coding has also been used. Each major group has a colour and when this colour was used outside of that major grouping it denoted that there was co-membership of that group with the groups own primary grouping. In the case of the multi-level diagram this could also be achieved through sub-grouping into more specific groups which share a higher level group i.e. a superset of the sub-groups.



(a) (b)

Figure 5.4: Spider Diagrams

Grid

This style is denoted by a grid structure being imposed on the groupings to construct groups with multiple common forms of attributes. The grid is constructed either by outlining a variety of distinct types of groups across a column or down a row, then labelling a second type of attribute across the unlabelled row or column. This might be, for example, forms of organisations down the rows and different locations across the columns, thereby creating groups in the grid representative of a particular organisation in a particular location. In one example this was created by combining 'friends', 'colleagues' and 'acquaintances', with 'PhD students', 'Academics', and 'Administrators'.

In Figure 5.5 the participant has created a grid, but used it in an unconventional manner. The different coloured columns represent different areas of the participant's life – 'Cornwall', 'Home', and 'None Social'. Groups within these columns are then defined by different hatch marks indicating a specific type of group within that area, for example 'Photography' is represented by a right diagonal hatching in the Cornwall colour. In addition to this the diagram creates a more traditional row across the grid for 'VIP' contacts across each of the three grouping areas identified by the column. This row represents groups from each area containing contacts who are particularly important to the participant in terms of how much they speak to them.

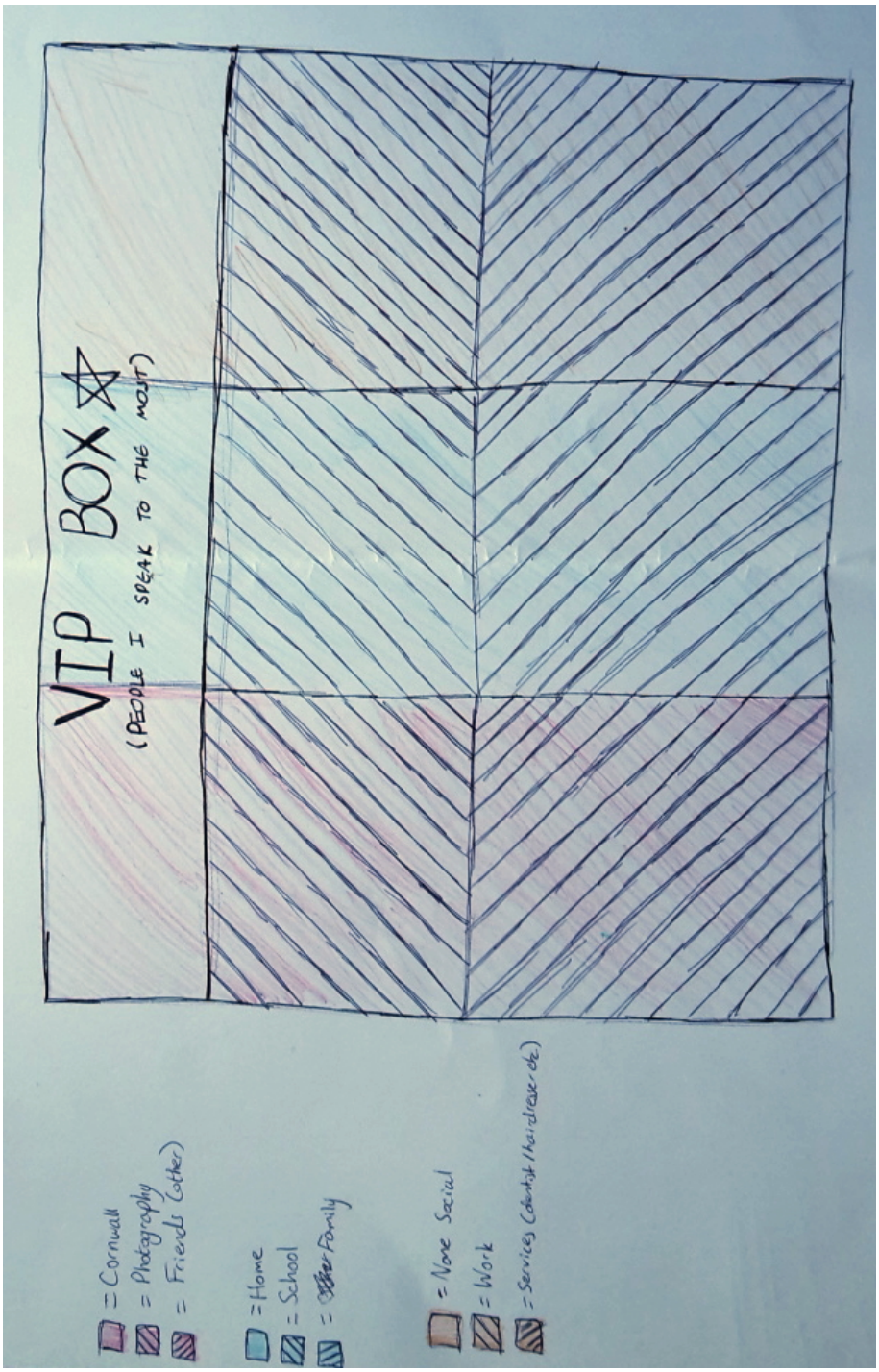


Figure 5.5: Grid Diagram

Mixed

A proportion of the diagrams encountered contained a mixture of one or more of the previously encountered styles. These diagrams all had a clear primary style they followed, and are counted as such in the previous categories, but large elements of the style used by the participants followed central elements of other styles. This allowed participants to reflect more clearly certain factors that were important to them. It also gave participants the flexibility to reflect particular group structures that would not be obvious immediately within their primary style. With the amount of data available, no particular pattern could be discerned in this behaviour to group these mixed diagrams further.

As previously described, these diagrams had a clear primary style meaning that they retained most of the features of a particular category of diagram. The features that were different to the primary style were considered part of a secondary diagram style, which borrowed some of the features of another diagram type. To illustrate how these mixed diagrams broke down they are listed below categorised with the frequency that each type of diagram occurred as the primary design, alongside the frequency of categories that described those diagram's secondary influences. The primary classifications of the Mixed diagrams with their secondary classifications in brackets were as follows:

- **Primary:** 2 Axis **Secondary:** 2 Venn
- **Primary:** 2 Physical Location **Secondary:** 1 Venn, 1 Spider
- **Primary:** 3 Venn **Secondary:** 2 Grid, 1 Location.

Figure 5.6 is an example of a mixed diagram. In this case the participant has combined an Axis diagram with a Grid diagram. The Y Axis is the length of time that the participant had known the group members, while the X Axis is the type of relationship. The X axis, however, also serves as part of the grid element of the diagram defining categories for columns of relationship type. The names and widths of the groupings create rows, and form the basis of a grid structure, along with the Y axis.

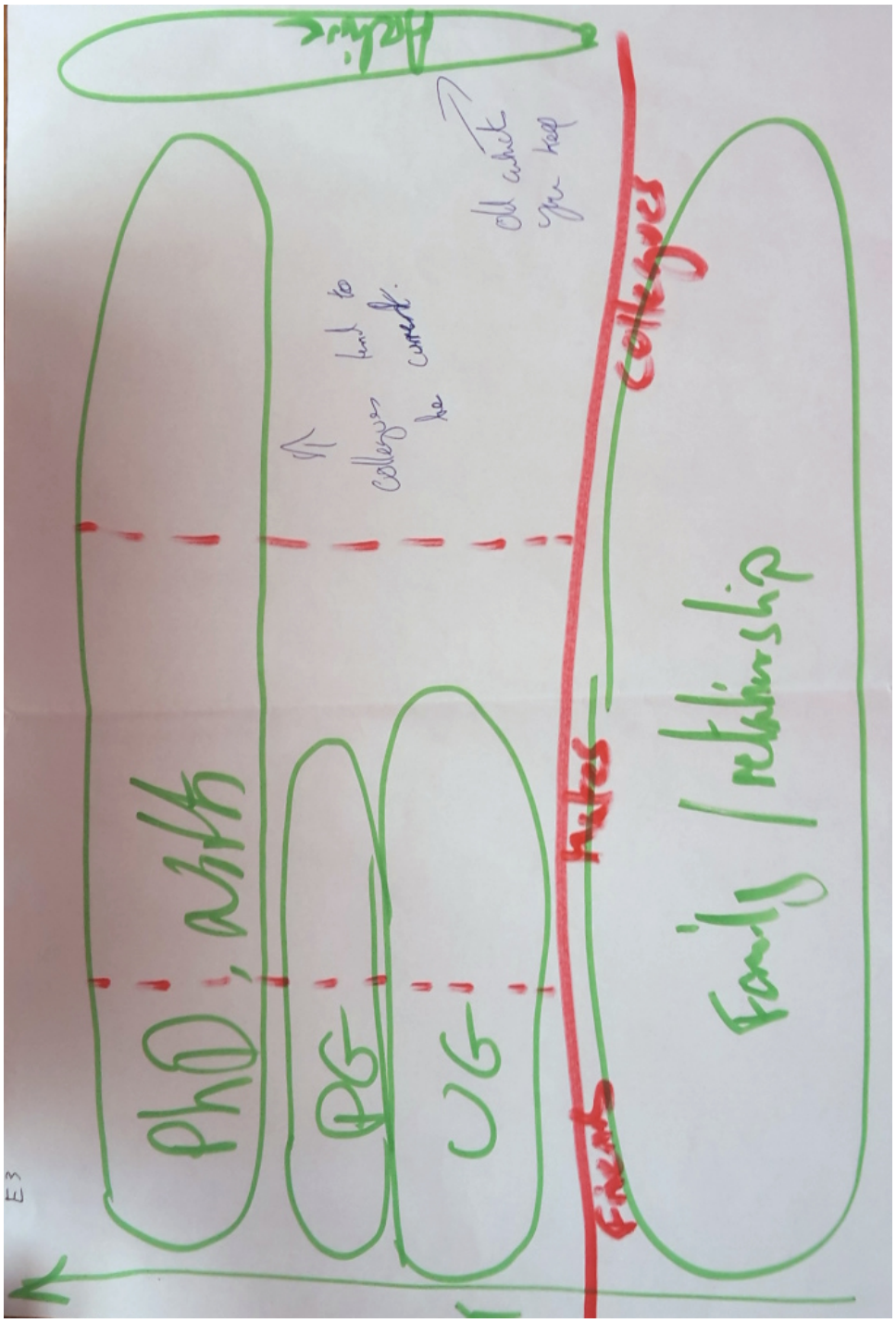


Figure 5.6: Mixed Diagram

5.3.2.2 Groupings

In order to examine the groupings of the diagrams further, each of the diagrams was coded to describe how many levels of groups were depicted within it. A level was considered to be any distinct hierarchy of groups, where one group was clearly a subset of another larger group. In the Spider diagrams this is indicated clearly through the structure of connected sub-groups that are inherently hierarchical in their layout, while for other diagrams this involves a degree of interpretation. For example, the diagram depicted in Figure 5.5 was interpreted as having two levels. Across the columns of the diagram there are dominant large groupings that contain the rows that create more specific sub-groups within the superset of the column groupings (for instance – 'Cornwall photography', 'none social VIP'). In addition to this, each group in the diagram being analysed was coded to describe the type of connection the group was describing, to determine what forms of group were common within the diagrams. The coding for this was determined by inductively coding groups into themes that emerged from the data (Braun and Clarke, 2006). The results of the coding of the number of levels of contained in each diagram and the coding of the types of groups that were identified are summarised in this section.

In the coding of groups there was an evident split that existed between the diagrams, of those which were a broad flat representation of the participant's contacts (N = 8), and those which delineated a hierarchy of sub-groups (N = 12) stemming from broader groupings. Of those which represented sub-groupings the following numbers of levels were found: two levels (N = 7), three levels (N = 2), four levels (N = 2), five levels (N = 1).

The first level of broad groupings present in all of the participants' diagrams often shared many similarities, suggesting a similar conceptual view of the high level kinds of groups that contacts can be broadly divided into was shared among the participants. There was quite a large amount of variation in the number of groups created representing those high level concepts, however. The average number of first level groups was 6.61 with standard deviation of 3.16, and a minimum value of 3 and a maximum of 14. The common divisions of groups encountered and their occurrence of each group type are described below.

Education (N = 11) – A grouping of people met at specific, or a variety of points, during the participant's education, whether it be school, undergraduate degree, or postgraduate studies. While school friends were common, the added prevalence of university contacts reflects a bias due to the number of people in the study with a university education.

Family (N = 16) – Family members were prominent as a top level group for most of the participants. The distinction of 'close family' and 'extended family' as separate groups was commonly made, sometimes as a major grouping, and sometimes as a sub-grouping of Family.

Friends (N = 18) – While a grouping describing friends was prevalent in nearly every diagram, the manner in which the groupings were described were varied. Ten of the diagrams described friends in a purely general sense, while the remainder specified aspects such as close friends (N= 2), old friends (N = 2), friends of friends (N = 3), and friends in a location (N = 1).

Interest (N = 9) – For some of the participants particular hobbies or areas of interest were significant groupings of people. Typically these would be named for the interest, e.g. 'Photograph' or 'Comics People', only one participant labelled their interest as 'Hobbies'.

Location (N = 21) – Many of the participants chose to group contacts in terms of specific physical locations. These were typically cities or fairly broad geographical regions, such as the North East, but in several instances were specific to a building. In three instances the concept of "Home" was also used, indicating a particular, but undefined location. While it was one of the most commonly occurring types of group, this grouping was used by only twelve participants, with some participants breaking down their contacts locationally and some specifying several locations for groupings.

Work (N = 29) – This was by far the most common grouping used and denoted the distinction between work and other areas of the participants' lives. This was further divided into five different forms of work based group. These were those broadly titled 'work' (N = 6), those described broadly as 'colleagues' (N = 10), specific organisations (N = 5), occupations, roles and expertise related groups (N

= 6), and specific project based groups (N = 2).

Other (N = 14) – These were groups that people used to collect together contacts that they no longer contacted, but once did, or contacts whom they only infrequently contact. These groups had names such as 'Archive – Old Contacts You Keep', "Random People", 'Acquaintances', and 'Others, Companies Etc'.

Those diagrams that contained multiple levels represented the further intricacies of their network structure in a range of ways. Some participants created a hierarchy based on a particular subject, expanding upon the relationships within it through sub-groups. An example of this would be for a participant to state an area of activity such as work, or an interest, such as canoe club, and then go on to split that into the various relationships that composed that higher level grouping. This might be groups of friends and acquaintances, or within an organisational grouping this might be particular projects or roles filled by individuals. This could also be the reverse of the area of interest / activity then relationship sub-division pattern, starting with a higher level grouping of a relationship like friends and then dividing it into divisions such as interest, location, organisations etc. Similarly, several participants organised groups based on their relationships with particular individuals, their partners or children. This involved a group named after an individual with sub-levels of associated groups. For instance, some participants used their partner as a group and named sub-groups of friends associated with their partner, and another participant had their child as a group, and associated subgroups that related to their child's hobbies.

5.3.2.3 Annotations

Participants made a variety of annotations to their diagrams, which served to provide a range of additional information where the structural design of their diagrams was unable to convey what participants perceived as important social information about their groupings or the relationship of the groupings to each other.

Some participants chose to overlay distinct circles cutting across group representations, to illustrate additional information about less formal forms of grouping

or social distinction. This particularly applied to the Venn diagrams, with these grouping cutting across multiple areas within the diagram, and were indicated, for example, with a dashed circular line as opposed the solid lines typically used to indicate groupings. This method made a distinction between people belonging to personal and work circles. Similarly, one participant chose to use a key of coloured symbols to indicate whether contacts in the groups they drew were friends, family, colleagues, musicians, or extended family (see Figure 5.4). Colours were quite prominently used by some participants to denote different areas of their lives. By drawing groups in distinct colours, the participant could quickly establish the area of their life to which the groups pertained. This might relate to a particular location, or to a form of relationship e.g. friendship, family etc. Use of colour also enabled participants to represent groups of contacts which bridged or rubbed up against the wider context of the social structures and not just co-membership of individual groups.

Arrow links were the feature most common within the diagrams, and largely used to indicate links between distinct and separate groups. Arrow links were most common within the Spider Diagrams, and were a response to the need to indicate membership links between groups in different hierarchies. However, they also sometimes indicated other wider social links between groups. Some participants chose to use arrow links to describe the trajectories towards friendship between their groups. This might show how people can move from belonging to a colleague group to a friendship group, to a close friendship, demonstrating the progression people have taken in the past within their social circles. And in a variation on this one participant used arrows to indicate the direction a person was moving within a social group, closer towards the centre, or further away to the periphery of the group, with closeness to the centre representing the relative importance of that person within the group.

Other annotations included the use of stars to indicate important contact groups and cross hatchings to indicate what type of group it was, as is described in the Grid representation in Figure 5.5. The use of a central point in diagrams labelled 'Me' – or a self-representative symbol – was also a common feature among all of the Spider Diagrams, but also occurred in several times in other categories of dia-

gram, for example the Physical Location diagram in Figure 5.3. This represented the participant at the centre of the diagram, with their groups and connections radiating out from them as the central point of the diagram.

5.4 Discussion

Predictably, perhaps, the results illustrate a complexity of human relationships when grouping people. The diffuse nature of the repertory grid construct results, over the many themes of connections identified and the highly individual diagrams, demonstrates a wide variety of factors people consider when defining relationships between their contacts.

5.4.1 Constructs and Diagrams

The broader grouping behaviours found in the diagrams are similar to those in Kelley et al.'s (2011) privacy contact sorting study, where they encountered a range of groups from 2-30 and a mean of 6.5 groups, against the 6.61 mean of this study and the range of 3-14 groups. Kelley et al. (2011) similarly note the multidimensional nature of how contacts could be grouped according to many or a few of the contact relationships they identified. Jones and O'Neill's (2010) privacy contact sorting study found a similar range of 4-20 groups, but a much higher mean number of groups with 16.7. Leskovec and Mcauley (2012) asked participants to identify all the social circles their friends belonged to and found a similarly high mean of 19 groups. While the results of this study are comparable to these previous studies, one reason this study may reflect a lower mean than some studies is that the contact groups created were not done so methodically considering each contact, as was the case in these other studies, but rather from memory, considering what the groups might be rather than going through the process of creating them.

Contrasting the diagrams with the repertory grid constructs reveals that the diagrams contained broadly similar, though less varied forms of groupings (see Section 5.3.2.2) – on an individual basis – to those identified in the repertory grid constructs (see Table 5.1). The literature would suggest that in many instances

these broader groups seen in the diagrams tend to be created in practical examples of manual contact group creation (e.g. Kelley et al., 2011). While groups more akin to the specific constructs of the repertory grid are harnessed more by systems, such as ReGroup (Amershi et al., 2012), that identify groups that fit specific criteria, suggesting a variety of groups with along different dimensions for people to be grouped by.

Less immediately obvious, but also present in the design of the diagrams, was the tendency to create constructs that contained multiple social dimensions. This was evident in a more diffuse nature throughout the entirety of the structure of the diagrams. All of the different types of diagram demonstrated different ways of organising and making explicit the various social layers concurrently at work. While each diagram might focus on a particular facet of their social network to emphasise, such as the importance of contacts or co-membership between groups, they each worked to make the multiplex social layers visible within the chosen structure of their diagram. A simple example of the ways in which diagrams reflected multidimensional social constructs is in the form the Spider Diagrams took. An arm of a Spider Diagram might be labelled 'Work', with a sub-group labelled below this named 'Friends'. Here, the result was the formation of a multidimensional group, 'work friends', through the layout of the diagram.

The methodologies were also complimentary in some regards. In the diagrams there would often be a distinction made between family and extended family that was not often present in the constructs. This could be due to fewer opportunities for this occur within the construct elicitation approach. Conversely, the constructs provided an opportunity for some participants to describe attributes, and make emotional and behavioural judgements, such as making a distinction of people who are family and those they considered 'family' such was the closeness of their relationship.

Some of the constructs, including some of the low-level categories of constructs identified may be overly specific in the relationships they represent to provide a useful basis to represent contacts generally for performing information tasks, reflecting more the specific connections participants perceived between several people rather than groups of people. However, identifying groups of this kind may

be useful when creating more adhoc groups, such as those supported in Amershi et al. (2012) in their semi-automated group detection approach. For example, in the case where a construct represents a shared interest in a football team, which the participant may have no interest in themselves. This might represent a grouping of people the participant knows to be interested in a particular subject or may be able to advise them on a subject should the need arise. Knowledge such as this can be useful when performing what Whittaker et al. (2004) call outeractions, where a person's social network is mined for useful information from their contacts.

On the whole all of the constructs could be divided into two general forms: those which the participant was involved in and those they were not involved with. Those they were involved in tended to relate to the participant's perception of specific social relations e.g. social skills, intimacy of relationship, people they work with. Those constructs they were not involved in tended to be factual attributes about the people in question e.g. a qualified skill or a place they lived. This suggests two distinct forms of criteria for use in identification or creation of contact groupings, those based on the participant's behaviour with contacts and those based on factual information about contacts.

Asking participants to draw the diagrams of how the they would group their contacts created a conflict or tension between the multidimensional nature of human relationships and the participant's need to represent this as a flat representation of their thoughts. As a result, the different diagram types that emerge are not just representative of the participant's own preferences in representation, but also reflect the need to represent the variety of complex relations at work, by focusing on a particular aspect of how these relations might be structured. The Spider Diagram focuses on representing multiple facets through an increasingly specific structure, the Venn Diagram focuses on representing overlap in relations between groups and to some extent structure, the Location Diagram focuses on the spatial positioning of people relative to groups, the Axis diagram focuses on the participant's behaviour or intimacy with people and the Grid Diagram ignores the representation of overlap for the benefit of representing groups of people in relation to a variety of attributes or other social layers. Most obviously the Mixed

diagrams demonstrate the tension of reconciling the 2D diagrammatical structure they had tried to impose on their multidimensional social network by combining different ideas.

The annotations, most commonly in the form of connecting lines that participants made to the diagrams are further evidence of this tension. Often, these annotations would be exceptions to the design structure within a diagram, and represent special instances of a relationship, or an attempt to remedy their inability to represent the multidimensional nature of their relationships.

The structure of the diagram, and the annotations to the diagram, highlighted a particular desire to organise their contacts according to the frequency of contact with a person or group of people, for some participants. The concept of visualising closeness is something that has been explored within other research, and its presence in these results further suggests some people do perceive contacts in this way and that there is merit in pursuing its use as a metric. Within social network research Antonucci (1986) explored this with the Hierarchical Mapping Technique, whereby names are placed within concentric circles, with names in the inner circles representing those with a greater level of closeness. Similarly, within the field of visualisation this is an approach that Viégas et al. (2004) used to represent frequency of email communication in *PostHistory*, whereby those people users exchange emails with most frequently tend towards the centre of a series of concentric circles.

The diagrams and constructs elicited here, however, suggest that this might valuably be expressed in a variety of other manners than the relative spatial closeness of contacts to a point as much of this research has done. Here closeness was also often expressed in a variety of forms through words to indicate intimacy, such as 'close friends' or by other pictorial means such as colour coding of contacts to indicate their importance. This has been explored to some extent in personal information management by Fisher et al. (2007) looking at email triage with their SNARF system, where contacts which were frequently contacted could be filtered out from all their other email messages among a variety of other metrics that might determine importance, such as whether a person was the only recipient of an email.

5.4.2 Methodology Drawbacks

One of the drawbacks of the repertory grid triads approach to construct elicitation was that it was likely to favour eliciting the types of constructs that are shared by many of their contacts, but not necessarily those which are most important to the participant. This is not necessarily negative, given that identifying common forms of grouping that are relevant for ordering a person's contacts was an intended goal of this research. However, it is possible that this approach may potentially hide the relevance of certain useful forms of grouping that are less prevalent, and equally it does not comment on the significance of the groupings it does identify. It is also worth noting that the methodology imposed a restriction on repeating the same connection, which encouraged participants to reveal the connections they perceive in greater depth, but did not favour eliciting the most salient connections between people.

The freestyle diagrams, however, had other drawbacks, related to their potential to describe relationships between contacts too broadly. This largely manifested itself in that the diagrams would sometimes result in participants generalising very broadly about a person's relationship to others, sometimes oversimplifying diverse relationships, and sometimes exaggerating the importance of tenuous relationships between groups of people. In some respects this is reflective of the same apathy with the task which can affect the creation of useful or accurate groups on social network websites (Mazzia et al., 2012). Due to this issue there may be questions over whether the diagrams reflect a complete depiction of all the groupings they perceive. This method also particularly fails to identify isolated people who do not really fit into any generalised group of people, but who can be nevertheless interesting and significant people distinct from particular groups.

Constructs relating to location, while clearly an important part of how contacts are remembered, are potentially less likely to appear in the results, due to the restriction on repetition, as it is quite likely that many of a person's recent contacts to live or have lived in the same places within, particularly the participant's current home area. Equally constructs such as frequency of contact are less likely to be present in any given comparison, but as the mapping exercise illustrated, for some people the Frequency of Contact with people is useful signifier of their

importance as contacts when organising them. Issues such as these are why both these methodologies were employed, so that a more complete picture could be ascertained and to mitigate the individual issues of each methodology.

5.5 Conclusion

The multiplex nature of the relationships that participants described highlights the complexity of the relations and the variety of factors taken into consideration in identifying those relations between the different groups of people that people perceive. The diagrams and repertory grid constructs that people generated were highly individual and multidimensional in nature. Certain commonalities, however, could be discerned in the forms of diagrams, groups, and constructs being created.

From these commonalities a categorisation of repertory grid constructs describing relationships between people was developed. In particular these constructs most commonly related to locations, affiliations or roles in organisations, forms of relationship, and events. Many constructs, however, combined two or more of the categories of identified connections or modifiers to define more specific multiplex relationships. The implication of this is that some of these categories can act as building blocks or dimensions through which people can include or exclude people to describe subsets of contacts. Considering the use of some of dimensions, particularly where they are unrelated to the social structure of the groups, may be able to aid in boosting the accuracy of automatic detection of groups. These dimensions may also be beneficial for the design of interfaces when considering how to present and organise contacts, for example, through the use of filtering contacts.

The visual representation of the groups through the free-hand diagrams highlighted five different forms of diagram, each demonstrating a different aspect of how participants perceived their contacts. This included a focus on:

- Structure.
- Contact importance.

- Location.
- Contact overlap between groups.

The groups generated by these diagrams tended to elicit broader descriptions than the constructs, but also corroborated findings from the repertory grid constructs. At the broadest level, the groups from the diagrams and also the constructs both commonly consisted of locations, organisations and roles, and the nature or intimacy of relationships, but also common were events, how or where people were first met, and the frequency of contact with them. However, at the lowest level the constructs described more nuanced, transitive and specific connections between people, such as events, behaviours, motions, modifiers, and multiplex constructs.

Having identified relationships that were prevalent in linking together contacts and forms of representation to order contact groups, these findings can be used as the basis for interface designs, through which their utility for supporting information tasks through a person centred approach can be evaluated. To this end designing an interface that reflects the characteristics of the representations and orders or filters contacts based on the relationships identified is the next step toward this and is detailed in the following chapter.

Chapter 6

Person-Centric Prototype Design

6.1 Introduction

Having established some of the ways in which people think about their contacts, in terms of groups and relationships, and how they thought about visually representing those groups, a prototype system was designed with incorporating those findings. The aim of this system was to establish to what extent performing information tasks by interacting with people might be beneficial, to establish which forms of representation of contacts are preferable for facilitating personal information tasks in this way, and to explore the process of contact grouping.

The system was designed to use the participants' own personal information from their email and Facebook accounts in order to generate representations of their contacts and then to use those representations to perform information re-finding and sharing tasks with their data. This provided a means through which to evaluate the design elements based on the previous research, and to collect data closer to real world use in a lab setting, as well as an opportunity to establish the validity of the methodology and results of the grouping study in Chapter 4, particularly with regards to the grouping framework.

The design of the interfaces of the prototype system to be used in the evaluation were based on the findings of the diary study (Chapter 4) and the contact grouping and representation study (Chapter 5). In the first study participants' current information practices were recorded and analysed to understand what enables their information re-finding processes. In the second study knowledge elicitation tasks identified different structures for designing person-centric interfaces from participant drawings, and different organising relationships with which contacts could be grouped or ordered. These results provided requirements that formed the basis for the prototype interface designs. For example, interaction with the interfaces needed to support the re-finding process of associating information with people as part of its interaction, and the representations of contacts should reflect the central facets of participants diagrams by illustrating aspects such as links and co-membership. How these findings translated into design is explained throughout this chapter, and their subsequent evaluation is described and discussed in the following chapter.

The design of the prototype system is broken down into three parts that are each described in detail in this chapter. These are as follows:

1. **Contact Extraction** – The function of the first part of the system was to automatically extract and collate relevant data about the user's contacts and provide a means for the user to verify the validity of the results of this process.

2. **Contact Grouping** – The second part of the system provided several interfaces with which users would place their contacts into groups for use in the third part of the system.

3. **Person-Centric Interface for Information Tasks** – In the third part the previously created contact groupings are presented through each of four alternative views, clicking on contacts in the interface allows the user to either share or access information associated with each contact.

6.2 Contact Extraction

As was the case with the repertory grid contact elicitation in Chapter 5, it was desirable for the prototype system to elicit contacts from multiple sources to gain a more complete picture of the user's communication behaviour. To this end, the system would extract contacts from both the user's email account and Facebook account – if both were available (see Figure 6.11). Having entered the login details for their respective accounts, the extraction process would begin identifying contacts. Contacts were defined as being people with whom the user had reciprocated contact during the last six months. This was a fairly crude means of identification, but it served to include as many plausible contacts as possible and then allow for the user to determine which were genuine contacts – in the sense that they would contact them again – at a later point. From the identified plausible contacts, the system downloads the header information of the last six months of correspondence from each of the contacts. This consisted of the *sender*, *recipient(s)*, *time received* and *subject*, but not the content of the messages to avoid unnecessarily downloading potentially personal and sensitive information. URLs sent to the user, however, were identified from the content of the messages and stored in relation to the contact, as were the filenames of any attachments associated with messages. Additionally, in the case of Facebook contacts the profile pictures were downloaded and associated with the contacts where privacy settings permitted it. All of this information, barring the images, was then stored as a JSON file.

Following this process the user would be asked to help clean up the stored data (Figure 6.1). This involved removing any duplicate contacts. Duplicate contacts were those where the same contact appeared more than once, but where this had not been automatically detected by the programme. This was due to either multiple email addresses in use or the presence of the contact in both Facebook and email, but with different associated names. From the identified contacts, those who had highly similar names were presented to the user in turn, and they could confirm whether the similarly named contacts were in fact the same person or not. If they were the same the data was combined, if they were not then no action was taken.

Similarity of the contacts' names was determined using n-gram matching (Ekmekcioglu et al., 1996). In n-gram matching each of the names being compared are each broken down into a series of sub-sections of letters with the amount of matching sub-sections determining the similarity. In this case tri-grams were used, meaning that first three letters of a name were one sub-section, the next three after first letter were another, the next three after the second were yet another and so forth. N-gram matching was chosen over other similar methods due to its ability to consider similarity in small matching sections, as opposed to similarities of individual letters. It was considered that variations of the same person's name would contain many similar sub-sections, and therefore were more likely to be the same name if they shared subsections over merely just containing the same letters.

Having removed any duplicate contacts, participants were then presented with a screen listing all of the remaining identified contacts along with their pictures and email addresses (Figure 6.1). At this point participants were told three things:

1. Participants were given the opportunity to remove any contacts they considered to have been erroneously identified.
2. Participants were told that relevant contacts were those whom they would likely contact again in the future.
3. Participants were asked to rate each relevant contact on a scale of 1-7 – indicated via a drop down box – according to their importance to them as a contact in terms of the value they placed on receiving their messages, not necessarily their importance to them as person. For instance, the user might rate their child's teacher as an important contact, although they may not have a close personal relationship.

Following successful completion of these steps the contact extraction and data cleaning was complete, and participants moved on to grouping contacts.

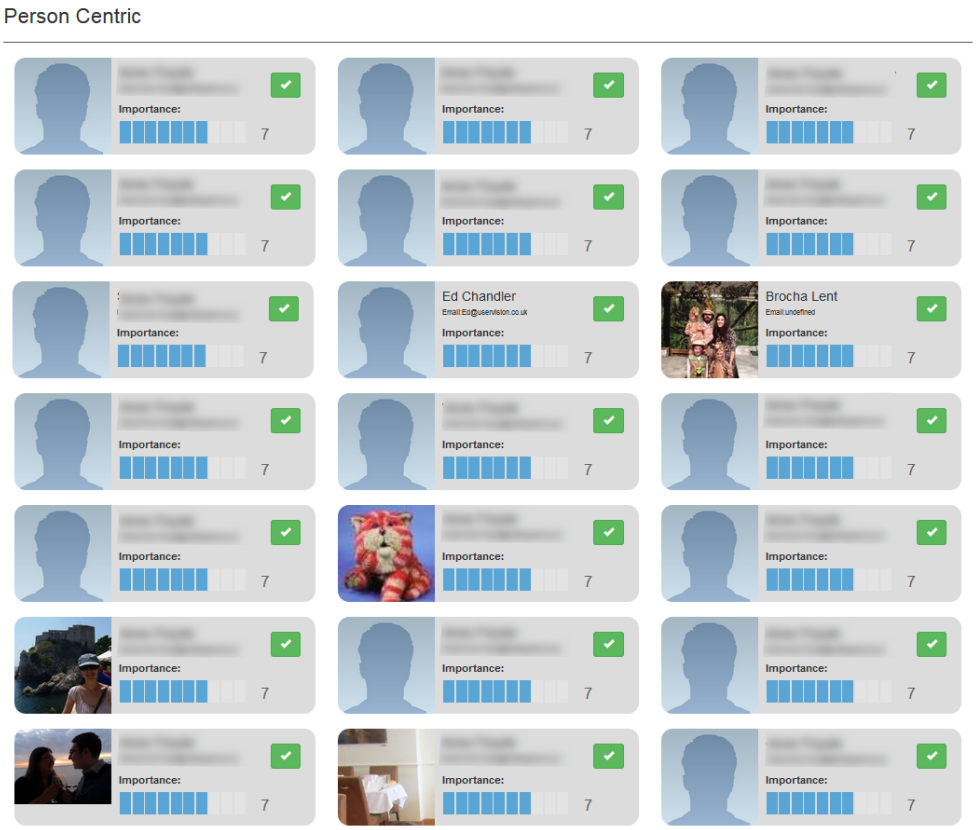


Figure 6.1: Contact Clean-up and Importance Rating

6.3 Contact Grouping

In the next section of the system participants were presented with their final list of identified contacts. The participants were shown two separate interfaces with which they could establish groups of contacts from their finalised the list. The two separate interfaces performed the same function and were devised for the purpose of examining how interface design might influence group construction. The first interface (Figure 6.2) was based on categorising a list of contacts into groups, the second (Figure 6.3) was based on creating an interconnected structure of groups. Although multiple grouping designs were used here for the purposes of understanding how participants construct groups, the grouping interfaces devised here were not proposed as solutions for contact grouping. The interfaces were intended to explore the decision-making processes involved in group construction

in order to identify how design could influence the creation of more useful groups or how useful groups might be more accurately detected by automation.

6.3.1 Grouping Interface A: List Grouping

The intent of the List Grouping interface was to represent the most simple forms of contact grouping interfaces. Those interfaces that involve sorting people by simply selecting them for inclusion within groups, typically represented as a list. Groups were created by typing into a text box and clicking a 'Create Group' button in the top left of the interface. The created group would then appear along the side of the interface. The contacts were presented alphabetically (by first name) in a grid next to the groups created by the user. Contacts could be added to groups by dragging and dropping them onto the desired group. Once grouped the contact would indicate this by fading slightly and changing colour to grey, however, contacts could still be added to multiple other groups if the user wished. It was also possible for the participant to view which contacts had been added to each group. Clicking on the groups themselves turned on or off whether contacts added to it would be presented as faded out, so participants could remind themselves of who they had added to each group or collective groups.

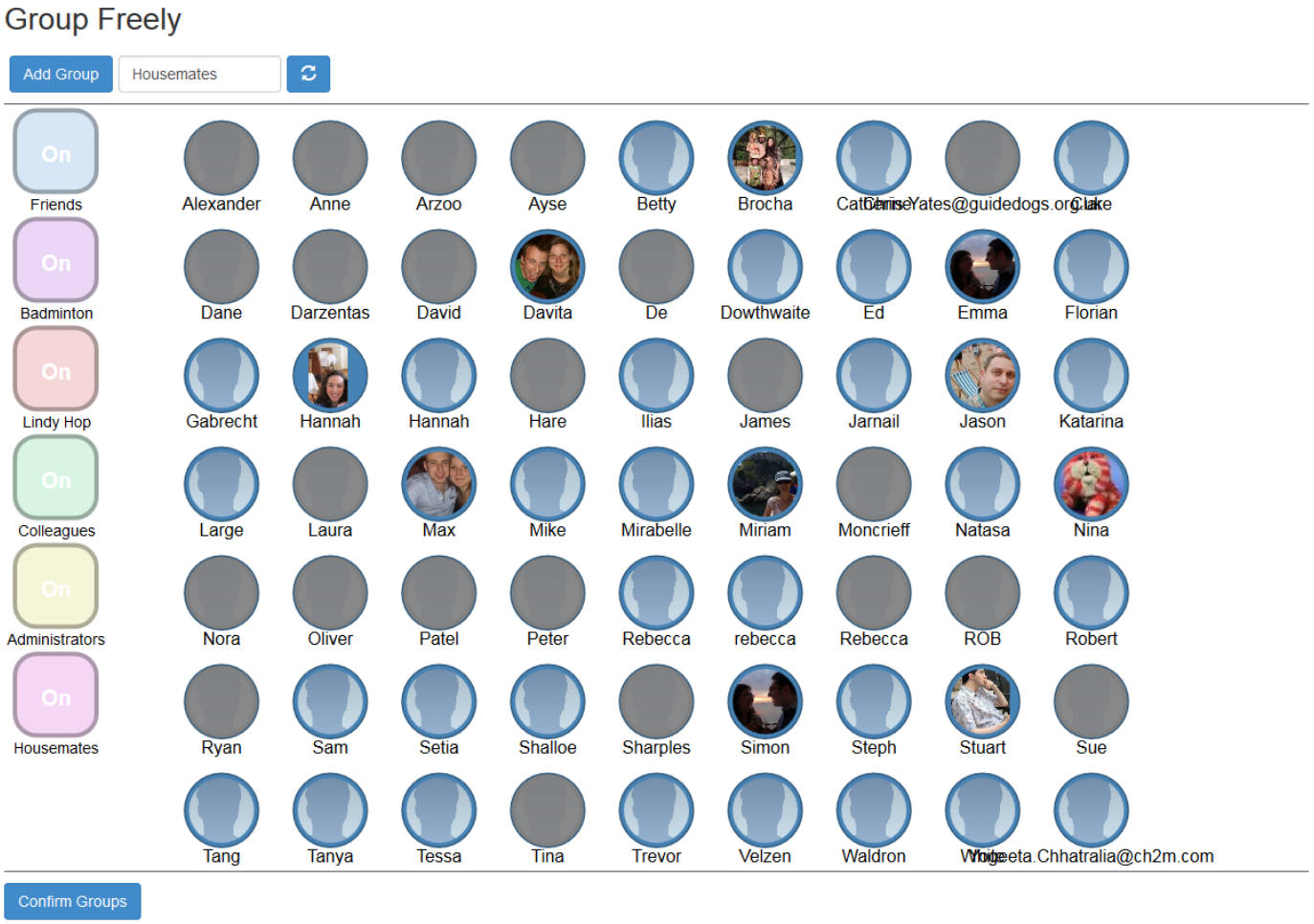


Figure 6.2: List Grouping Interface

6.3.2 Grouping Interface B: Egocentric Grouping

The egocentric grouping interface aimed to provide the means to construct a more complex structure of groups, influenced by the manner in which some participants had created visual representations of their social networks that contained involved hierarchical structures. This interface was based on the use of a force-directed graph (Fruchterman and Reingold, 1991). Force directed graphs provide a means of laying out the structure of graphs (in this case the hierarchy of connected contacts and groups), by assigning forces to the edges and nodes of that graph, such that the resting positions of the elements of the graph when those forces have played out result in an aesthetically pleasing layout. Primarily this approach reduces crossover between edges and produces edges of similar lengths. The results of this produce a representation of groups and contacts very similar to the Spider Diagram drawings participants used to represent their contacts in Chapter 5.

The interface would initialise with a large central circle representing the user and the contacts unconnected and free to move around the periphery. Creating a new group in the same manner as the previous interface – typing its name and pressing 'Add Group' – would create a new, smaller, circle representing the group, which would be connected to the initial circle via a black line. Contacts could then be added to the group by dragging and dropping them over the group, at which point a black connecting line would appear between the group and the contact indicating membership of the group. At this point the user would have a choice as to where to add future groups. Clicking either the central circle or any group would give it a black outline indicating that any future groups added would connect to that specific node, thus potentially allowing the creation of sub-groups. In addition, contacts could be added to more than one group. Doing this would in this case duplicate the representation of the contact (not the record of the contact) in order to keep the structure clearer, preventing the need for connecting lines to criss cross over the interface. As a force-directed graph all of the contacts and group nodes reacted according to set forces, determined such that groups and nodes repelled away from each other, spreading out to create a relatively clear structure.

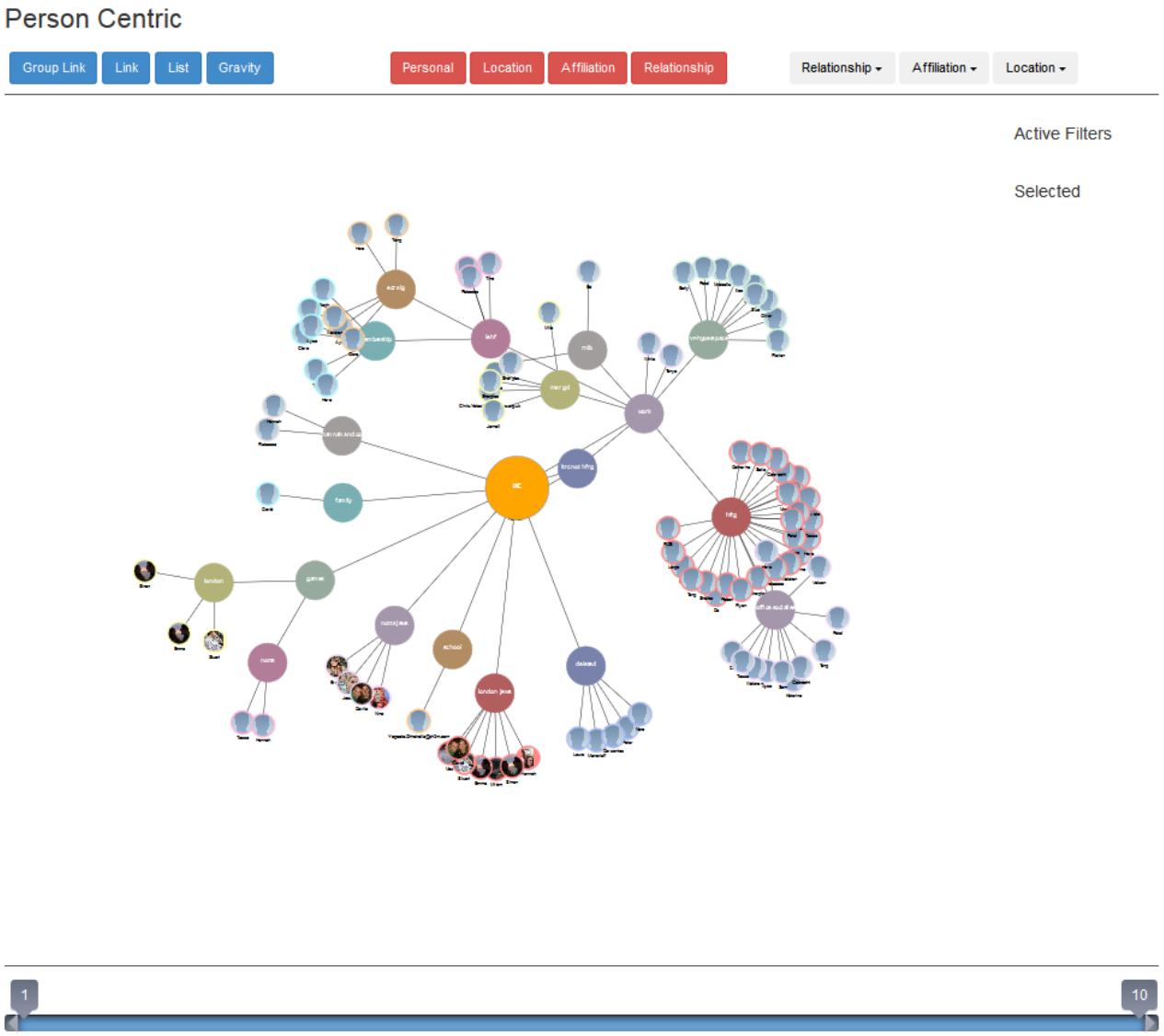


Figure 6.3: Egocentric Grouping Interface

6.4 Person-centric Interface for Information Tasks

Having completed the construction of their groups, participants were presented with a series of variations on an interface with which to perform a series of information sharing and re-finding tasks using their previously created groups. The Person-Centric interface was designed to examine information re-finding and information sharing based on a design centred on people as the primary component of interaction. Its design incorporated ideas identified in Chapters 4 & 5 through the diary study, repertory grid and contact grouping diagrams.

The following elements were identified from those results as warranting inclusion and further investigation within the prototype system:

- Representation of groups – diagram types (e.g. spider diagram), links between groups/contacts, use of colours to distinguish groups – Grouping Diagrams
- Association of information with a person – messages, files, links accessible via people – Diary Study
- Judgements of the relative importance of contacts - representation of contacts according to importance – Grouping Diagrams
- Multiplex nature of groupings – representation of different views/constructions of the same groups – Repertory Grid, Grouping Diagrams
- Grouping Criteria – Organisations, Relationship, and Location ordering were the most commonly used grouping criteria – Repertory Grid, Grouping Diagrams

6.4.1 Representation of Groups

From the results of the group diagramming (see Chapter 5), five forms of grouping representation were identified; Grid, Spider Diagram, Location, Venn and Axis. These were taken as different design ideas which may appeal to users to or fit how users of the prototype system think about grouping. Using these design ideas as a starting point, four different representations of contact groups were

devised.

6.4.1.1 List Presentation

The first representation *List* (Figure 6.4) represented a traditional or common presentation of information/contacts. In this presentation the contacts are ordered in rows and columns in clearly delineated boxes representing groups. If the number of contacts or groups exceeded the screen space, scrolling the screen up and down would provide a means to view all of the contacts. The intention with this presentation was to provide a comparison point for the presentations derived from the diagrams, which represented how people thought about contact groups, against what would be a more typical list-based representation (e.g. see Facebook). From this comparison the goal was to identify factors that are important when representing contacts and ordering contacts in order to support information re-finding and sharing tasks.

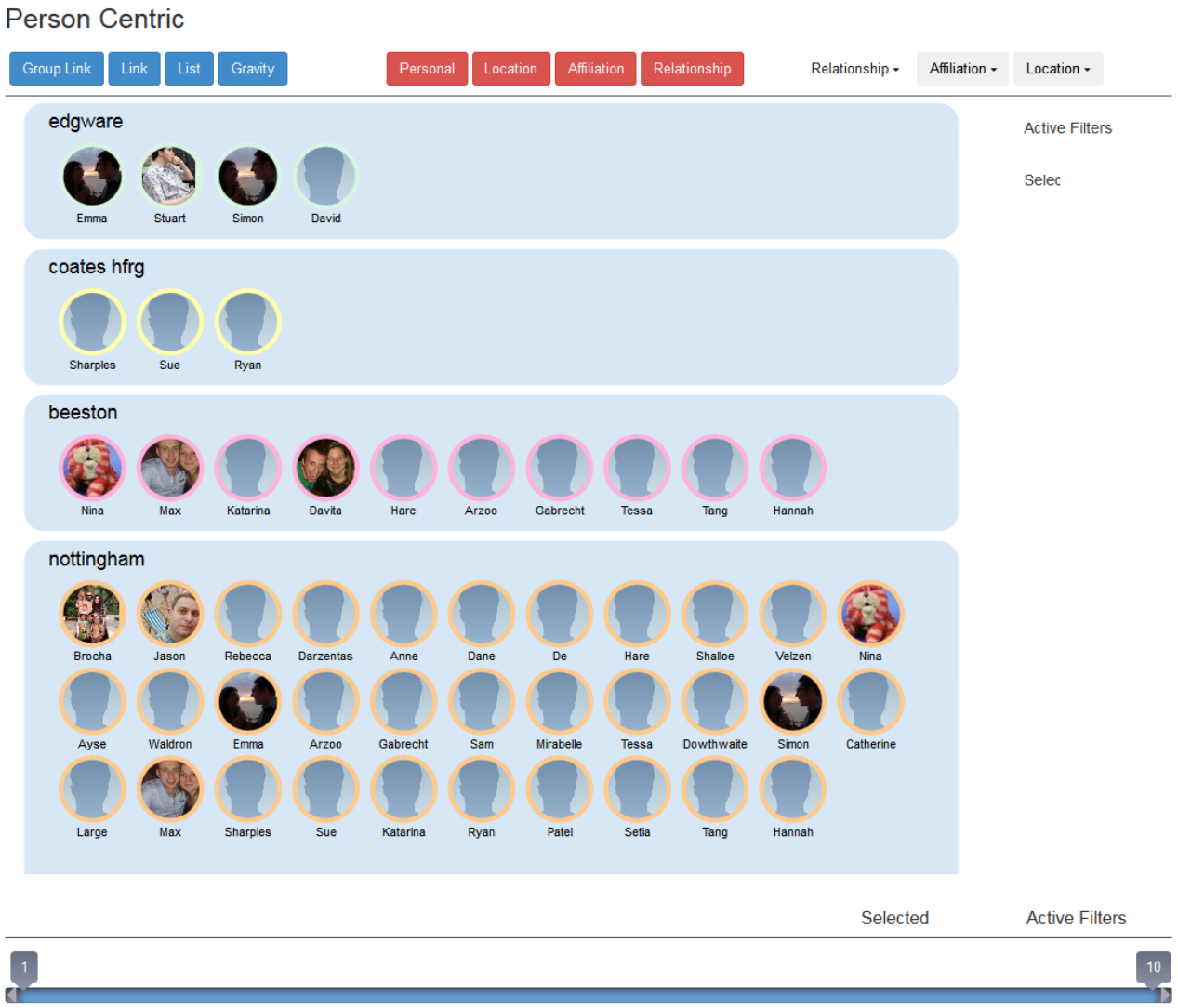


Figure 6.4: List Interface

6.4.1.2 Egocentric Presentation

This presentation of the grouping information derived from the Spider Diagrams participants created (Figure 6.5). It focuses on three features in particular from those diagrams – 'Me' as central locus in the diagram, links between groups, and the use of sub-groups. In this presentation the user is free to pan and zoom their view of the diagram by clicking and dragging, and by using the central wheel on the mouse peripheral that was used for the study. The presentation was generated as a force-directed graph, meaning that the various individual nodes representing people and groups were subjected to multiple forces that variously repelled them from each other and the centre, ideally creating a clear graph akin to those the participants created in the spider diagrams. It should be noted that every time the diagram was created the positioning of the various elements varied, though not how they were linked together. The groupings used in this diagram correlated to those the participant made during the *Egocentric* contact grouping task. In all other presentations the groupings correlated to the groupings constructed via the *List* grouping interface.

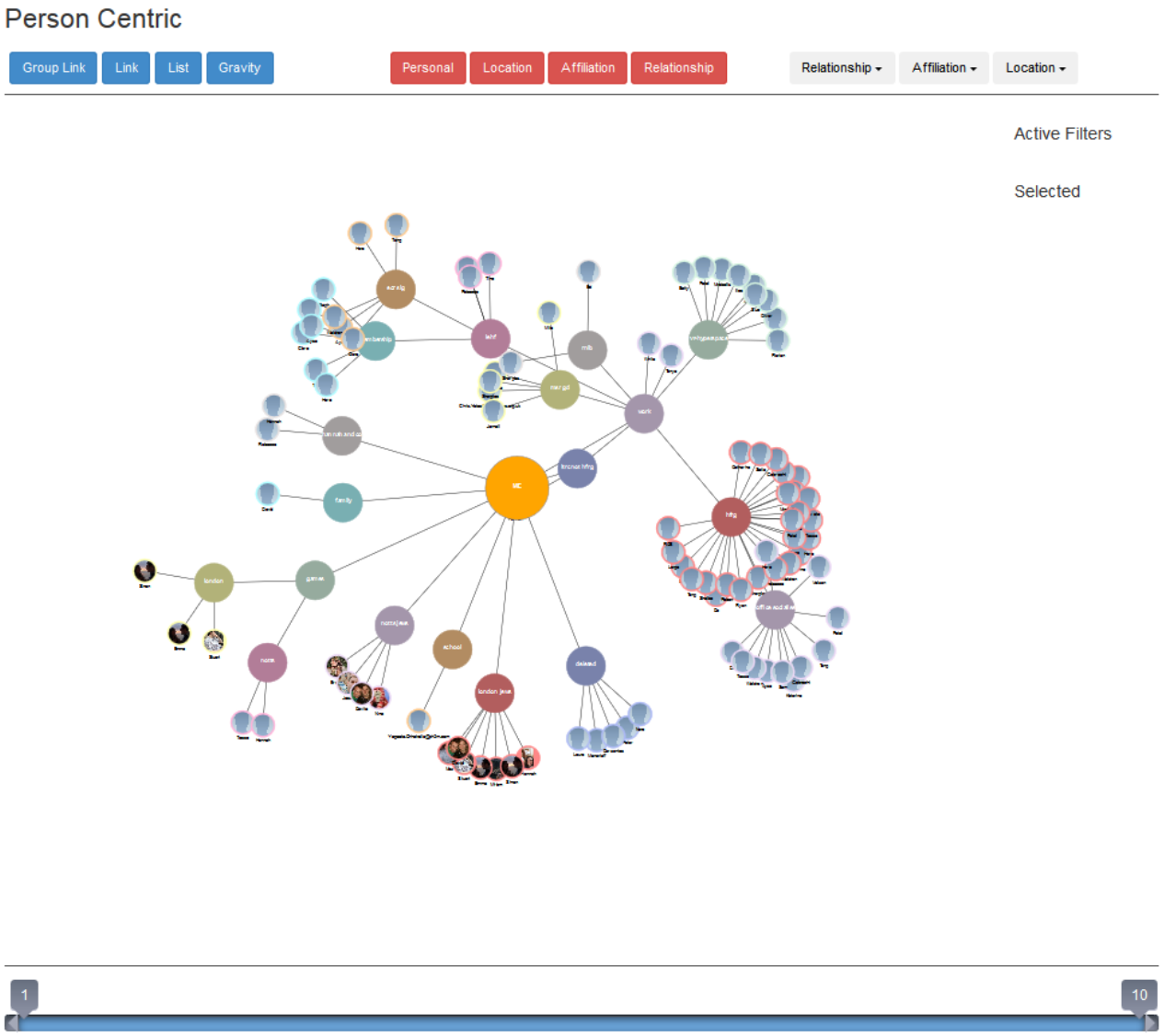


Figure 6.5: Egocentric Interface

6.4.1.3 Gravity and Group Link Presentation

The Group Link presentation (Figure 6.6) aimed to incorporate the Venn diagram idea of showing overlap between groups. Groups in this design were free standing and unconnected to each other, with links only going to those contacts who belonged to the group. However, contacts in this instance could be linked to more than one group, to represent the that the contact belonged to more than one group. This had the effect of highlighting how interconnected groups were in terms of membership.

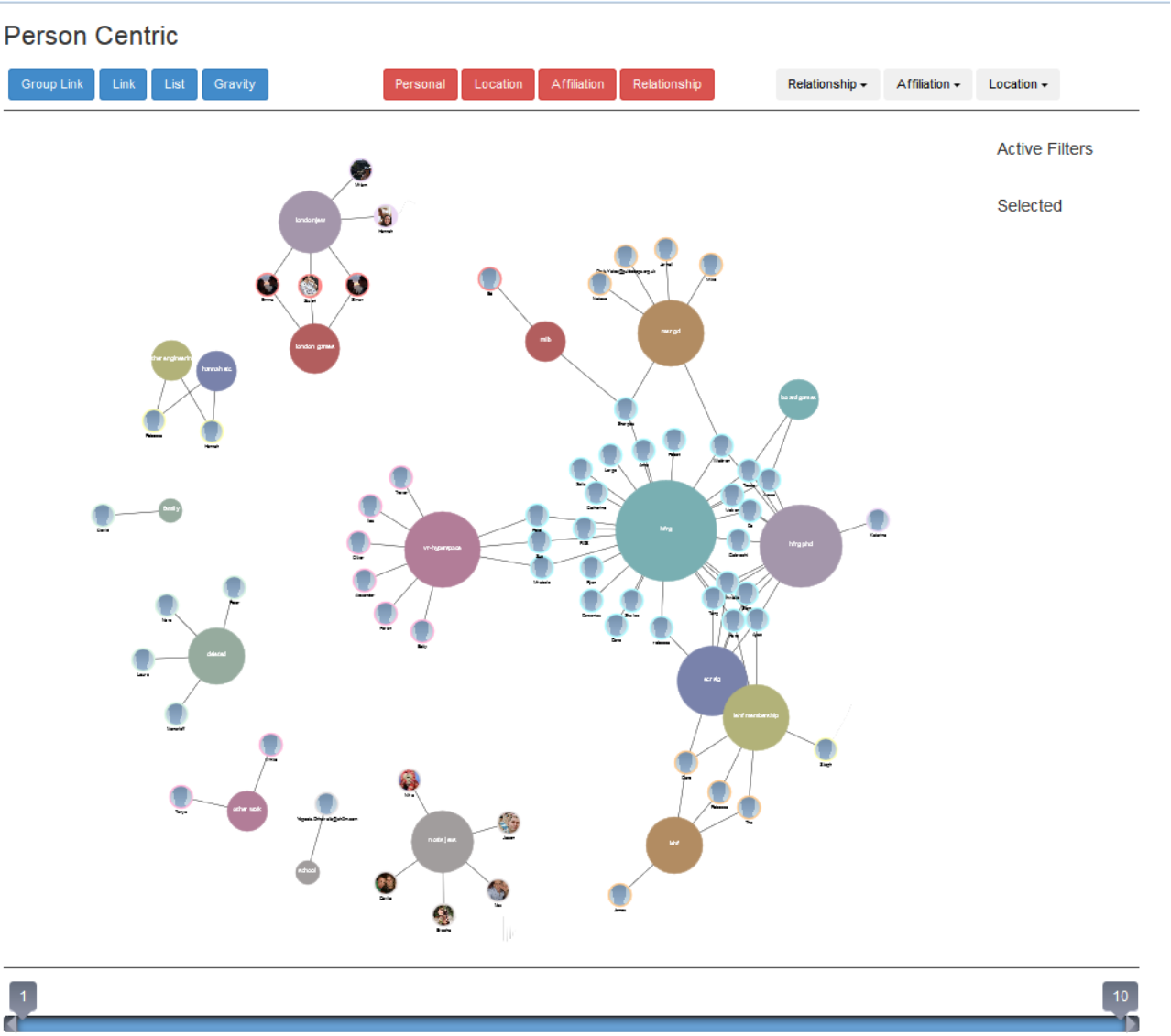


Figure 6.6: Group Link Interface

The Gravity presentation (Figure 6.7) by contrast aimed to show this same presentation, but without any links. Instead of links the individual contacts were pulled by a 'gravity' towards the group they belonged to, their association to it being implied by their closeness to it, matching colouration with the group node, and distance from any other group. Contacts in this case that belonged to more than one group were duplicated. In addition, the size of the group node in this presentation was proportional to the number of contacts belonging to it. The inclusion of the contrasting gravity and group link interfaces in the study was intended to explore the value of links in supporting the information tasks. This had two major aims: first, to contrast a representation dependant upon links to represent connections with a functionally similar interface without links, and second, to contrast a less cluttered interface with the sometimes complex representation of links. This additionally allowed examination of whether the size of group node had any navigational value through the inclusion of the Gravity presentation.

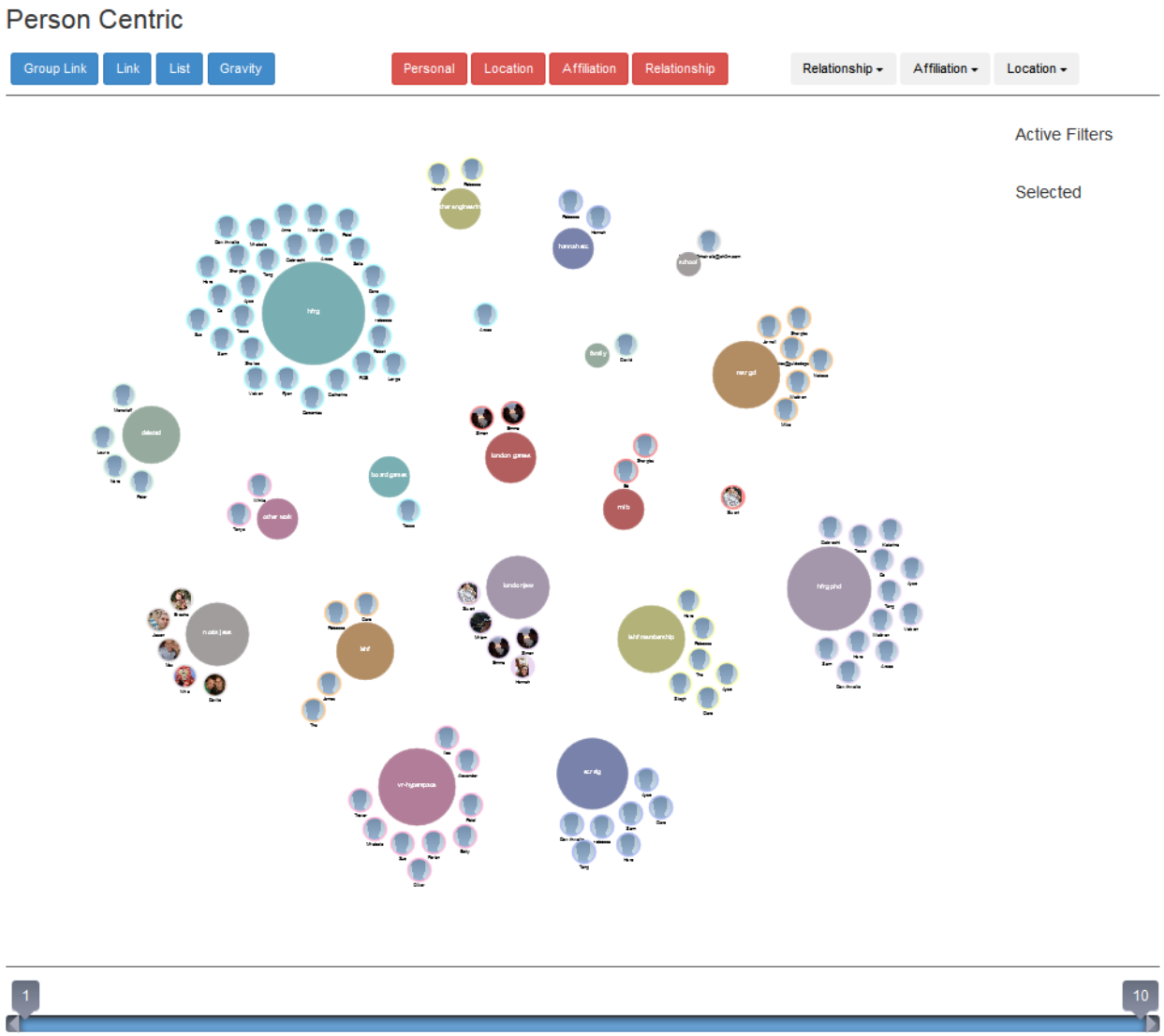


Figure 6.7: Gravity Interface

6.4.2 Association of Information with a Person

Clicking on a contact would reveal an information screen (Figure 6.8). This screen displayed all of the identified information related to a contact, split into three sections – messages, files, and attachments. Each of these sections contained items ordered by date – most recent first – and enabled the user to filter out items based on whether it was sent by, to, or CCed to the user as well as whether the information was sourced from Facebook or email – as indicated to participants using the typical respective symbols for each alongside each information item. Messages displayed header information of sender, date, and subject line, and clicking on this information opened the message. Links displayed the title of the web page and the date, and attachments displayed the file name and the date.

The purpose of this separation of elements was to establish whether providing direct access to information, in the manner that participants described remembering it in relation to people in Chapter 4, was beneficial in practice. Participants previously described a process of remembering a piece of information they required and then remembering that it was related to a particular person before seeking it out. By separating out individual elements of information that participants often sought – links, attachments, and messages – the aim was to provide a comparison between more common browsing methods, where these items might be embedded more deeply within the original source, and direct access to the information sought through the person recalled in relation to it. The source of the information was clearly displayed in the interface as the diary Study in Chapter 4 indicated the knowing the source of the information was an important part of the process of re-finding in nearly every re-finding event. By including this information the aim was to provide feedback about the information that users would look for in conjunction with a person in order to navigate to the information they required.

Additionally, users could right click, either on contacts or entire groups of contacts, to add them to a list of selected people on the right-hand side of the interface. This was feature used to mimic the process of selecting people to share information in the evaluation of the interface (Chapter 7).

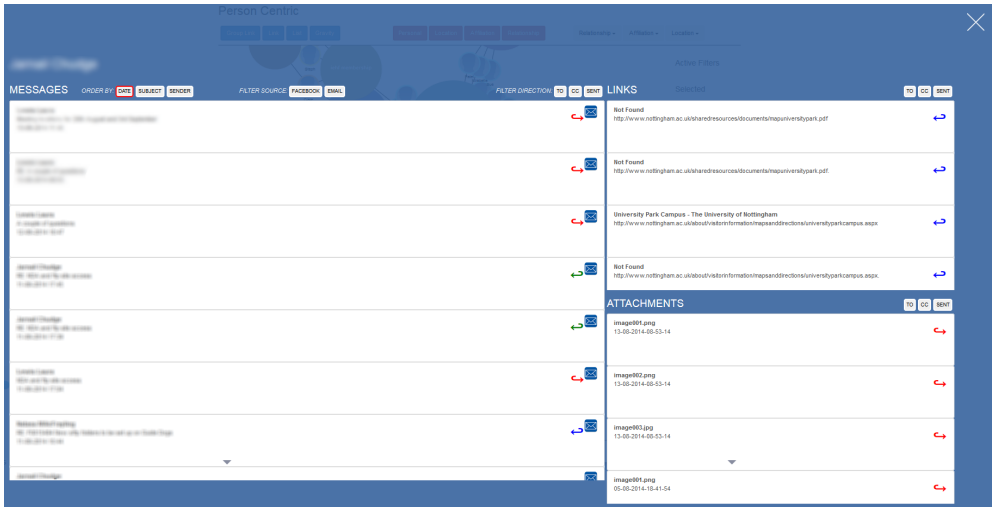


Figure 6.8: Association of Information with a Person

6.4.3 Grouping Criteria & Multiplex Groupings

It was evident from the results of the repertory grid and free hand diagrams that participants used a variety of different criteria to construct the groups and other modifiers to further define groupings. To reflect, to some extent, the variety of forms users potentially perceived in their contacts the prototype attempted to facilitate changing views of their contact groupings to match those perceptions. This was achieved through two design features – re-structuring groups and filtering groups.

As part of the group construction phase of using the prototype, users were also asked to group all of their contacts according to each of three pre-defined criteria. These three criteria were location based groupings, organisation based groupings, and relationship based groupings. These groupings represented the most common forms of constructs participants created, as shown in Table 5.1, and provided the basis of providing different grouping views of the contacts which may be beneficial for completing information tasks with the groupings/contacts. The re-structuring mechanism (Figure 6.9) provided four options for how the user could view their groups – either their own freely created groupings or one of the three defined groupings. Clicking on any of these buttons would re-structure the all of the groupings to reflect that particular set of groupings created by the user. Group

names would change, links would be created (or removed), and contacts would re-locate creating a different view of their contacts. Clicking on the Filtering (Figure 6.9) buttons changed the view by providing alternative criteria with which to restrict contacts in the view. Clicking on the buttons revealed a drop down list of the groupings made according to the button's criteria. Clicking on one of those groups would remove any contacts not associated with that grouping. For example clicking on the 'Location' filter button and then a group named 'Nottingham' would restrict contacts to only those the participant had grouped under 'Nottingham'. If necessary more than one filter could be applied at a time to further specify the viewable contacts. Equally filters could be removed: once a filter was applied, a tag naming that filter would appear on the right-hand side of the screen; clicking this tag would remove the filter and restore any contacts. This feature mimicked the types of multiplex groupings that participants created as part of the Repertory Grid and diagramming exercises in Chapter 5. This allowed the user to filter contacts down using the same constituent parts – such as locations, organisations, and relationships – that the multiplex constructs were often built from (see 5.3) allowing more complex definitions of which contacts are visible to the user.

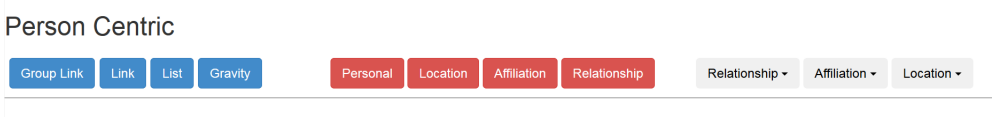


Figure 6.9: Interface Selection Buttons, Restructuring Buttons, and Filter Buttons respectively

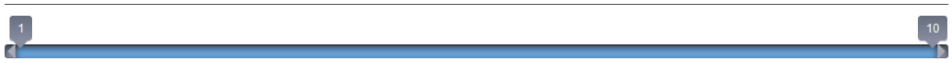


Figure 6.10: Importance Filter

6.4.4 Relative Importance of Contacts

During the creation of the free hand diagrams and constructs in Chapter 5, one of the themes that were expressed was the relative importance of contact groups. This was expressed in multiple forms: it was a central component of the Axis form of diagram where positioning in the diagram indicated the relative importance of

a group, where group names indicated the closeness of the relationship (e.g. close friend or acquaintance) or the group was constructed based on the importance of the contacts (e.g. the idea of VIP contacts). The idea of importance of contacts or contact groups was implemented into the prototype as an extension of the group filtering feature. Users were provided with an importance slider (Figure 6.10). With this users could define a range of importance ratings using the slider, and contacts who fell outside of that range would be filtered out of the presentation. To determine the importance rating of each contact the system used the ratings of importance of communication users provided in the initial contact preparation phase of the prototype system. This provided a crude means for users to highlight only those contacts whom they considered important, those in-between, and those whom they considered unimportant as contacts.

6.5 Prototype Development

The development of the prototype utilised two primary programming languages – Python and JavaScript. Python was used for the contact extraction, contact comparison, and contact name similarity evaluation. This consisted of three primary modules, one which parsed data from the mailbox of a specified IMAP service to identify contacts and related data, one which requested information about friends utilising Facebook's API (Application Programming Interface), and a third which integrated the data retrieved and generated a JSON (JavaScript Object Notation) file.

The interface design for the prototype was developed using HTML and CSS web standards. JavaScript was used to provide the interaction with the visual elements and management of the data imported via the JSON file generated by the contact extraction. Multiple JavaScript libraries were used to aid the development of these aspects, such as JQuery.js to simplify handling events and managing elements within the web pages. Of particular importance was D3.js. D3.js provides a means of managing data for the purpose of visualising it, and provides a scaffolding on which to build complex web visualisations. In this case it was important for managing contact groupings and calculating the position

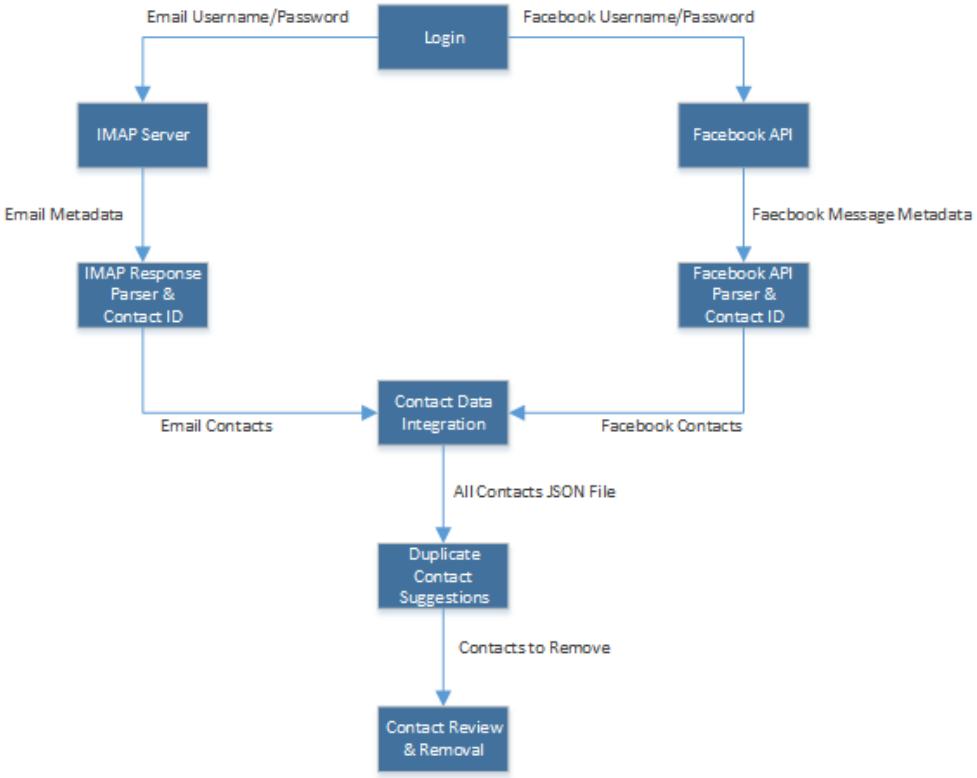


Figure 6.11: Data Collection Process

of contact nodes and groupings, for example managing the fore-directed graph calculations.

The web pages were hosted locally using a Python based web server – Flask – which served the web pages and handled requests from interactions with the pages for the previously described Python-based elements of the prototype e.g. communicating login details to initiate retrieval of contacts in the form of the JSON file.

The development took place over a period of five months, which included the development of several discarded design ideas for parts of the system. For example, as part of the group creation process a tagging system was developed that allowed groups to be generated through a process of tagging individual contacts with words that described the user's relationship with that person. This was done in relation to each of the most commonly created high level group types

(see Chapter 5), Location, Organisational, and Relationship, as well as freely tagging. Contacts with the same tags could then be considered to be part of the same group. This approach was discarded for two primary reasons. It was too labour intensive an approach – typing out tags for each contact even with auto-completion of previously used tags was arduous – and it obscured group membership from the user as part of the group creation process. If the process did not make it obvious to the user that they were grouping people together it was less clear to them that they were in fact engaging in a grouping process as opposed to one of associations with words, and secondly it prevented the user from making social judgements about which contacts belonged together within a grouping.

A phase of testing took place, initially using the personal data of the researchers to validate that contacts were being extracted and combined correctly with the relevant data, that groups were re-arranged correctly, and that the prototype was generally free from bugs. This process involved running the study described in Chapter 7 through as a trial and making comparisons between expected results and actual results. For example, a list of expected contacts was drawn up manually by browsing through Facebook and Email messages to identify all contacts within the planned date range. This was then compared with the results of identifying contacts via the prototype. Any omissions were examined to understand and correct their absence in the prototype. Following this further testing took place, with several other researchers, who were asked if they could see any omissions or errors in the data identified by the prototype given its selection criteria. Data from the testing stage was not recorded, personal data was only stored for the duration of use by the researchers taking part in the testing.

6.6 Conclusion

This chapter has detailed the design and development of a prototype 'person-centric' interface, based on the results of the information re-finding diary study described in Chapter 3 and the contact grouping study described in Chapter 5. These two studies provided a basis for understanding how people will often

recall information in relation to people, and for understanding how they represent contacts and the relationships between them. The prototype system aimed to support information re-finding and information sharing – two tasks identified in the Literature Review Chapter 2, and the diary study Chapter 3, as being tasks that where a 'person-centric' interface design might have a useful role in supporting these information tasks.

Based on the findings of the diary study (Chapter 3), in particular that when people are re-finding information they will often recall a person they associated with that information, the prototype system sought to enable re-finding tasks to be conducted by allowing information to be accessed primarily in relation to the user's contacts. To enable this the user's Facebook and email accounts were analysed, this identified contacts the user reciprocates contact with for use in the prototype system and also allowed related information – links, attachments, and messages – to be associated with the identified contacts.

In order interact with the identified contacts and perform tasks using the interface, multiple design ideas for ordering, structuring, and grouping the contacts were implemented based on research in Chapter 5 into contact grouping and representation. Contacts were structured in the interface based on the different purpose of the different types of contact representation diagrams identified, such as spider diagrams and venn diagrams which emphasised hierarchy and overlap respectively. The prototype system also allowed contacts to be grouped together or filtered out based on prominent relationships identified from the repertory grid exercise that described common types of relationships participants saw between contacts. In addition to this other concepts from the contact grouping research were also incorporated, such as the relative importance of contacts to the user, which could be used as a basis to filter contacts out of the interface.

In the following chapter the prototype development work described here is evaluated in order to examine how suitable the different approaches towards a 'person-centric' interface incorporated in it are able to support the tasks of information sharing and re-finding. Evaluating these features with users will enable greater understanding of the processes involved in performing information re-finding and sharing through 'person-centric' evaluation, and whether the features incorpo-

rated in the prototype are considered beneficial to performing the information tasks or not and how they might be improved, and what support is required with regards to aspects such as grouping contacts, locating contacts, and accessing information.

Chapter 7

Person-Centric Prototype Evaluation

7.1 Introduction

Previous chapters of this thesis have established the role that the associations people make between information and other people can play in performing information tasks by examining information re-finding (Chapter 4), and have elicited how people think about representing their contacts and how they perceive relationships between them (Chapter 5). The results of the studies in those chapters informed the design of the prototype 'person-centric' interface, which was described in detail in Chapter 6. This interface reflected four of the most common forms of representation participants created ((location, organisation, relationship), and provides ordering of contacts via the three most common forms of relationship participants perceived between their contacts. It also allowed participants to access information, such as messages, links, and attachments related to the contacts identified by the interface. In order to meet the objective of demonstrating and evaluating 'person-centric' interface approaches for supporting personal information tasks, and from that identify design implications for their use, this chapter describes the evaluation of the prototype interface. It addresses three key questions, on which there has been only a limited amount of

research, which will help inform the design of 'person-centric' interfaces: What are the decision making processes involved in contact group construction? How should contacts and contact groups be represented in order to support information sharing and re-finding? And what value can person-centred interactions provide for supporting personal information tasks?

To address these questions the study described in this chapter was conducted in two parts. In the first part of the study the process of constructing contact groupings, and the content and description of those groupings is examined. By better understanding the decision-making processes involved in group construction, the role of the interface in this process, and the make-up of how people group their contacts it will be possible to inform how contact groups should be represented and how people or automation can be supported in generating representations of personal contacts. In the second part of the study the contact groupings from the first part of the study are used to perform information re-finding and sharing tasks using four different representations of their personal contacts, informed by the representations and contact ordering criteria identified in Chapter 6. The interactions of participants with the 'person-centric' interface designs in performing these tasks is examined in order to evaluate the value of the representations and ordering criteria and determine what contribution they might make to supporting personal information tasks.

As the study contains two distinct parts, the methodology and results of each part are described in two separate consecutive sections, followed by a discussion of the overlapping results of both. The results are considered together with a view to the design of interfaces and systems which involve the creation of contact groups, and the implications for accessing, managing, and sharing information where people are a central component of the interface's design.

7.2 Group Construction

7.2.1 Group Construction Methodology

7.2.1.1 Study Design

In order to examine the process of group construction the 'person-centric' prototype facilitates participants constructing contact groups from their own personal contacts. This includes two separate interface conditions for contact group construction; one designed to approximate interfaces that enable grouping through lists of people and the other designed to allow users to reflect the structure of groups and contacts visually through links. The aim of comparing these two conditions was to see whether the differences between the two interfaces might influence the groups constructed and the process of group construction, as well as to look at contact grouping as a whole in order to inform how systems that utilise contact groups should be designed.

Prior to performing the grouping tasks it was first necessary to identify digital contacts from the participant's own data. This was done by the prototype accessing their Facebook account and email server to identify contacts to use in the grouping tasks. Both Facebook and email were used as this reflected the same process of contact identification as was used in Chapter 5 to perform the repertory grid task, in which a broad cross-section of contacts was sought across different areas of the participant's life, and to enable contacts to provide a point of access to as much of their data as possible – the objective being that the contacts would become a unifying access point to the information, regardless of the source of the information, although still retaining the data about what the source of the information was.

Following the identification of relevant contacts from the participant's own personal data participants grouped their contacts twice in a balanced order between the two interface conditions and a further three conditions in which a specified criteria was given for the participant to group contacts by. The criteria for these grouping conditions were 'Location', 'Relationships', and 'Organisational'. The purpose of these criteria was primarily to provide data for alternative contact or-

derings and the ability to filter contacts, accurately according to the participants view of their contacts in the second part of the study.

While participants were grouping their contacts they were asked to undertake think-aloud protocol (Ericsson and Simon, 1993), in which participants must constantly vocalise their thoughts as they perform a task. This was used in order to elicit data about the decision making involved in contact grouping. In order to capture the grouping processes participants undertook while using the interfaces and performing think-aloud protocol screen capture and audio recording was used. Whilst performing the think-aloud protocol participants would often pause for relatively long periods of time. They would stop interacting with the interface and cease thinking aloud as fluently or altogether, in contrast to their otherwise near constant interaction with the interface in order to complete the task. From participant comments it became obvious they were pausing to make decisions about how to group, or what groups to create next. By asking participants to explain at greater length these hesitations, the decision making processes that occurred during contact grouping were elicited. Although interventions can interrupt the thought processes of participants performing think-aloud protocol (Ericsson and Simon, 1993), they have been often used within usability testing in order to elicit specific thoughts on decision-making processes, where appropriate, in order to understand interaction with systems (among other aspects) (Tamler, 1998). At these moments of hesitation participants were prompted to continue thinking-aloud, and if participants were not clear about the rationale for the decision they were considering then the researcher would intervene in the protocol and asked further questions about why they had paused their current activity and the decision making they were undertaking.

Analysis of the screen capture and audio recordings of think-aloud protocol were thematically coded in several passes (Woods and Dempster, 2011) to identify different emergent participant behaviours with the interfaces. While analysis of the contact groupings created was performed by applying the grouping framework described in Chapter 5 to categorise the groupings created by the participant. The process of categorising the groupings created in the two free grouping conditions, according to the grouping framework developed, was not exhaustive in describing

all possible forms of grouping. In several cases, where the framework did not logically fit some of the results, the non-categorised groups were split into new themes for inclusion in the framework.

In seeking ethics approval for this study the sensitive nature of data being dealt with was an important issue to address. It was made clear in the ethics application exactly what data was being accessed, how it would be processed, what data would be retained, and how participants would be able to remove data from the study as they saw fit. Prior to beginning the study all participants were provided with an information sheet about what their participation in the study would involve and a consent form to sign indicating they were happy to participate in the study and understood what that entailed. The participation sheet made clear what data would be accessed and potentially recorded during the study, what data would actually be retained, and the previously described process for removing any sensitive information from the collected data. It also made clear their right to leave the study at any point and have their data removed, that their data would be made anonymous, and that their data would be securely stored.

7.2.1.2 Participants

Eighteen participants were recruited (12 Female, 6 Male) for this study, sixteen of whom worked at the university either as PhD students, research staff, or administrative staff. Of those sixteen seven were part of the School of Cultures, Languages, and Area Studies, five were part of the Human Factors department, and three were part of an interdisciplinary digital research centre. The participants were recruited through calls for participants advertised via mailing lists at the University of Nottingham and through word of mouth. No constraints were made on participation in the study, besides use of Facebook and Email for personal communication on a regular basis. It was considered that having a sample which was predominantly students and researchers might potentially influence the emphasis of some types of groups created, in terms of quantity, as Kelley et al. (2011) found, where students created many university related groups. In the case of this study, this effect might skew the number of organisational / soci-

ety groups occurring. Although a similar distribution might also plausibly occur were the sample more reflective of people working in other non-university organisations. While it is important to note the potential effect of this sample, this should not influence the range of types of groups encountered and should not affect the aims of the study which were to validate the categorisation from Chapter 5, and to look at the factors influencing the process of grouping and process of accessing those groups, rather than in determining types of group occurrence that are representative across the whole population.

7.2.1.3 Procedure

In the first part of this study participants were asked to use the prototype developed in Chapter 6, to log in to their primary email account and Facebook account. Following this the system then contacted their email server and the Facebook API to download and analyse the header information (sent to/from, date of message, subject) of their messages from the last six months. From this analysis the prototype automatically extracted the names, and associated message headers of those people whom the participant had reciprocated contact with over a period of a few months. A period of a few months was chosen over a longer period of six months, year long, or even the entirety of a person's inbox, so as to reduce the amount of contacts which were potentially no longer relevant, and instead keep the number of contacts manageable. This also had the additional benefit of significantly reducing the time required for the analysis of the messages to complete, which in the case of particularly large email accounts could take several hours.

Once the email and Facebook accounts had been analysed the participant was instructed through a brief process of cleaning up the data. First the prototype identified people who were potentially the same person from the two different sources (Facebook & Email); this judgement of similarity was based on how closely elements of the names of the contacts name matched. It then prompted the participant to judge whether each contact identified was the same person. If two contacts were judged to be a match by a participant then the data for the two contacts was combined into a single entry. This process compensates for

any discrepancies in names between recurring contacts they have contacted via Facebook and email.

After this process of combination the participant was then presented with a list of all the contacts the prototype had identified and asked to delete the contacts whom they considered to be irrelevant. Irrelevancy was defined as seeing no future need to message a contact, or that a contact represents a business providing them with a service, e.g. a mobile phone provider, rather than being an individual whom they want to be in contact with. As the participant examined the list of identified contacts they were also asked to rate each participant according to how important they perceived that person to be to them as a contact at this point in time. Rating was performed via a provided scale of 1-7 for each contact. Participants were told this was not necessarily a judgement on how they valued the contacts as people, and were rather asked to consider how much value they would attach to that person's correspondence appearing in their inbox or Facebook messages.

After this process of data clean up and rating was completed the participants were given five contact grouping tasks to complete. These were completed in different ordered sequences from participant to participant to counter any task order effects. Four of the tasks used the list interface pictured in Figure 6.2 - and described in greater detail in Chapter 6 - with the intention that the interface be roughly analogous to interfaces where groups are created by selecting people from a list of contacts. The interface utilised a simple drag and drop functionality to allow participants to place contacts into the groups that they had created at the edge of the screens. The groups could be toggled on and off by the participant; this indicated membership of those groups by fading the opacity of the member contacts in and out accordingly. Contacts could be grouped into as many different groups as the participant wished, but no sub groups could be made other than by linguistic choice.

The remaining task was completed using the ego-centric interface pictured in Figure 6.3. In this design, a force-directed graph is constructed by the participant, with unordered contacts. Initially the participant is presented with a single circle in the centre of the screen meant to represent them, and from this circle partici-

pants could attach groups to the central point. Participants could then add the contacts, again dragging and dropping them onto the groups and in this instance connecting lines indicate their membership to groups. Additionally, in this interface, participants were shown that if they wished to they could create sub-groups by connecting new groups to existing groups. Contacts could again be used in more than one group, but to prevent the interface from becoming cluttered with connecting lines, the contact's graphic was duplicated each time it was added to a new group.

The four tasks set with the list interface were Free Grouping, Location Grouping, Organisational Grouping, and Relationship Grouping (described in more detail below). These four tasks aimed to elicit any connections between the contacts significant to the participant, by encouraging them to group comprehensively based on the most common categories groups identified in the second study of this thesis in Chapter 5. The ego-centric interface's task meant that there was always repetition of the Free Grouping task. This repetition aimed to examine whether the ego-centric interface's more visual demonstration of connections, and ability to model connections in greater detail with sub-groups, changed the nature of the groups created or the participant's attitudes towards grouping.

Before beginning the five tasks described here the participants were told that later they would be set tasks where they would need to share and re-find information using these groupings, and that they should bear this in mind when they created their groups. They were also told that they should group based on what connections were significant and useful to them, that the aim was not to model all connections they perceive to exist, but to represent connections that are useful to them. The tasks were set as follows:

- *Free Grouping* - The participant was told to group the contacts however they thought they would find most useful and made most sense to them.
- *Location Grouping* - Grouping based on any form of perceived spatial significance between the contacts. This was described as being valid whether the grouping described something as narrow as a particular section of desks in an office, or as large as a continent.

- *Affiliation (Organisational) Grouping* - Grouping based on any perceived affiliation with some form of organisation whether it is formal or informal. Hereafter referred to as organisational grouping, this was described as including things such as organisations people work for, projects they are working on, clubs/societies they are involved in, or even a pub quiz team. The term affiliation was used instead of 'organisational' to avoid the suggestion that this task only referred to formal business organisations.
- *Relationship Grouping* - Grouping based on how participants perceived their relationships to people. This was described as including relationships such as friends, enemies, and acquaintances, but could also include specific roles such as managers, colleague, or supervisor.

In the ego grouping condition, using the egocentric interface, participants repeated the Free Grouping condition (or completed it for the first time depending on task ordering), grouping the contacts however they saw fit. This repetition of the tasks allowed a comparison to be made between the groupings made in the list interface and the ego interface.

While completing these tasks screen capture was used to record the participant's interactions with the interface, along with an audio recording of the participant using think-aloud protocol (Ericsson and Simon, 1993) to expose their thought processes regarding their grouping decisions and interactions with the interface. Due to the personal nature of the data being captured by the recordings participants were told they could note down the time of any instances where information they considered to be sensitive was captured, and that they would be provided with the opportunity to remove it from the collected data at the end of their participation in the study. Additionally, at the end of their participation the message headers from the participant's Facebook and email accounts were deleted. The only data retained in the course of the study was the screen capture, audio recordings, and contact groupings.

7.2.2 Group Construction Results

This results section will begin by describing the broad strategies the participants were observed using in the process of grouping and will then focus in on more specific results and observations of the contact grouping process.

When grouping contacts, participants all largely followed similar processes to each other in terms of when they created groups, how they processed the contacts into groups, and the points at which decision making tended to occur. Figure 7.1 illustrates diagrammatically the processes participants were observed undertaking to group their contacts. It provides a description that accounts for the majority of the participants' behaviours and decision making processes that were observed and revealed through the participants' hesitations, interactions, and thoughts as expressed through the think aloud protocol. The diagram helps to illustrate two prominent features of the behaviours observed during the grouping exercises. It delineates the two broad strategies participants used to create the groups, and it highlights the points in the process where participants were making decisions (highlighted in red). The two points highlighted in red were where participants would typically hesitate during the grouping process, either when deliberating over the inclusion of a contact in a group or planning the next group to create, and represent the points at which much of the decision making seemed to occur.

7.2.2.1 Adding Contacts

The two strategies participants used to add contacts to groups were either to examine each contact sequentially – highlighted in red in Figure 7.2 – or to search for specific contacts – highlighted in red in Figure 7.3. When sequentially examining the contacts the participant would consider whether each contact fitted into any of the existing groups created or a single specific group being considered. If the contact fit into a group they would add them to a group and if not they would move on to the next contact. This provided a systematic means to ensure all contacts were included in the correct groups. When using the strategy of searching for contacts participants would not search among the contacts in an ordered manner, they would recall the contacts they were looking for and go straight to their location, either by scanning for the image of their face, navigat-

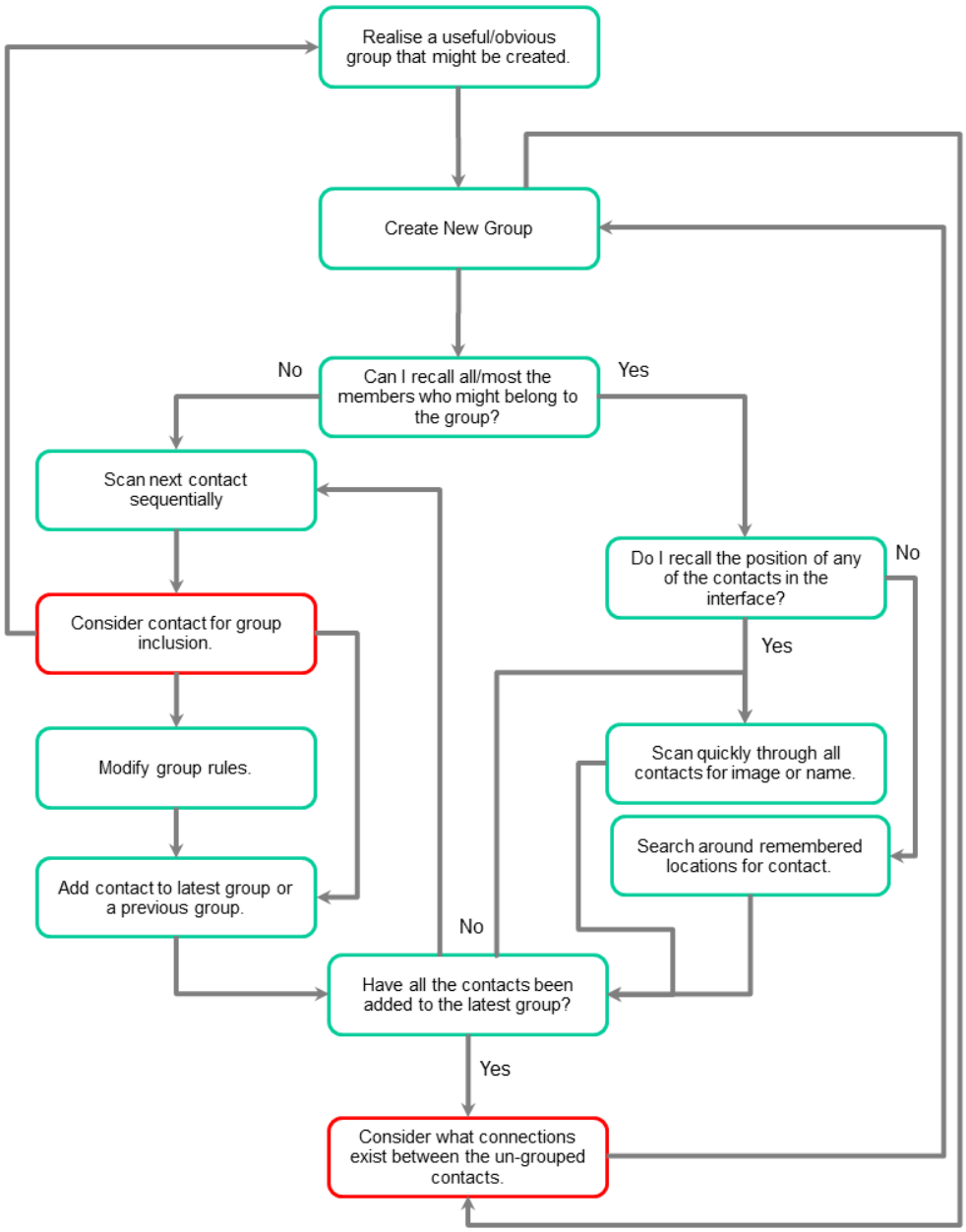


Figure 7.1: Contact Grouping Diagram

ing alphabetically, or looking for their name in the area they recalled the contact to be. In contrast to the sequential scanning approach, the participant did not spend time considering about which group might contact might fit, because they had already determined which contacts belonged to the group that they had created most recently. Often this behaviour corresponded to the creation of smaller groups and groups that represented very clear social structures to them, removing the need for careful consideration, as they would know both that they wanted to create the group and that they could recall all of the people who might belong to it.

This description of participant behaviour is to some extent a very abstract depiction of what occurred during the group construction process. The two strategies of sequential examination and searching were evident across all participants and were used the majority of the time that participants were actively grouping, but at times participants would switch between them and the boundaries between the strategies were not absolute. Participants would change and adapt their strategies as the structure of the groupings they were creating took form. At times a third strategy - or lack of it - was present, in that the grouping behaviour appeared haphazard for short periods of time, with contacts being added in no particular order to no particular groups, almost allowing the structure to emerge without too much consideration. There appeared to be no obvious reason this behaviour would be instigated. It perhaps related in some instances to participants reviewing the un-grouped contacts and adding any missed contacts to already created groups.

7.2.2.2 Group Creation

Similar to the two strategies of adding contacts there were also two broad strategies encountered with regards to approaching group creation a mixture of which might be used by participants. Participants would either create all of the groups at the very start of the exercise - coloured red in Figure 7.4 - or they would sequentially create a group and add all of the contacts to it - coloured red in Figure 7.5. When creating all of the groups before grouping any of the contacts, participants would spend a considerable amount of time thinking about the contacts

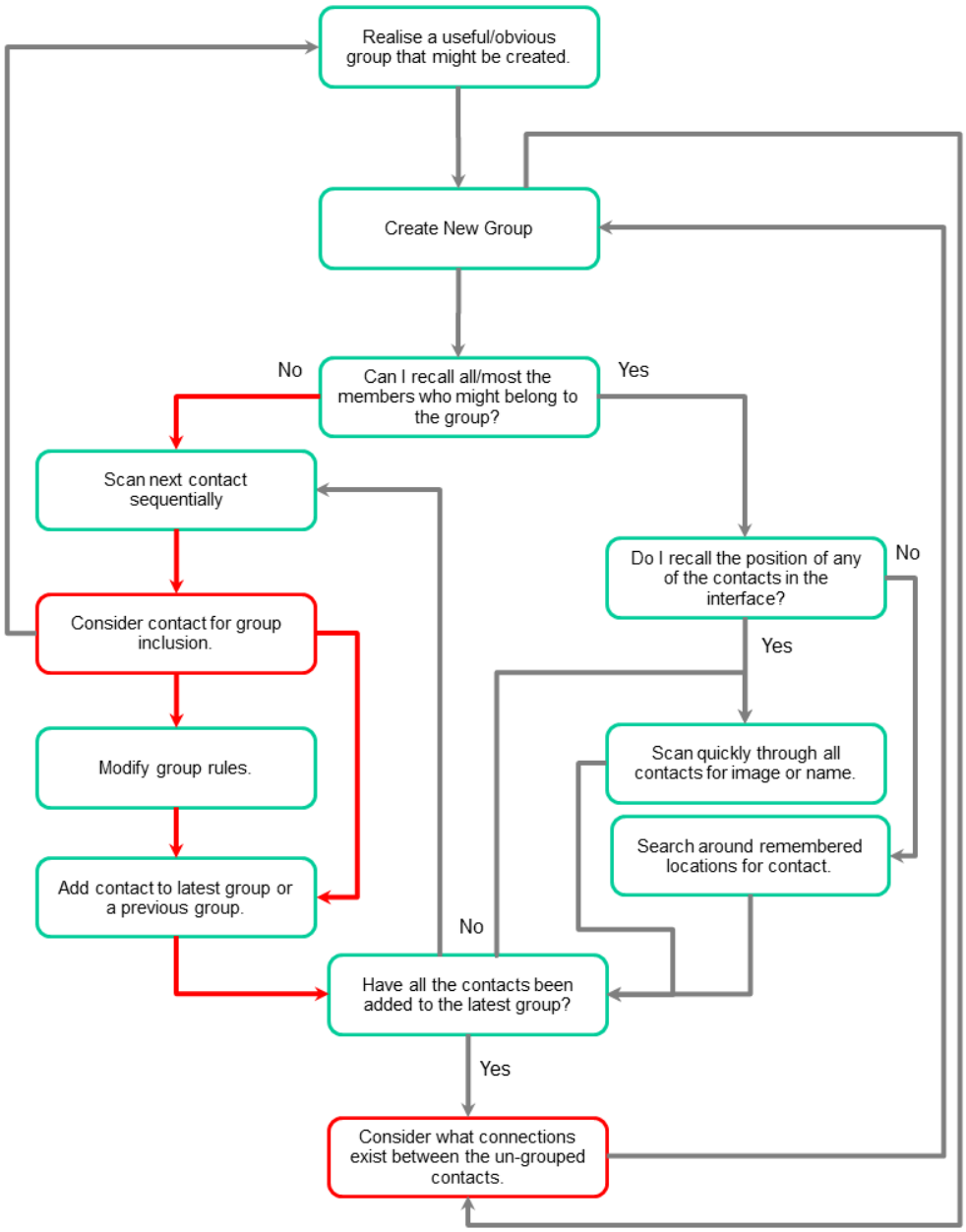


Figure 7.2: Contact Adding Strategies – Sequential Strategy

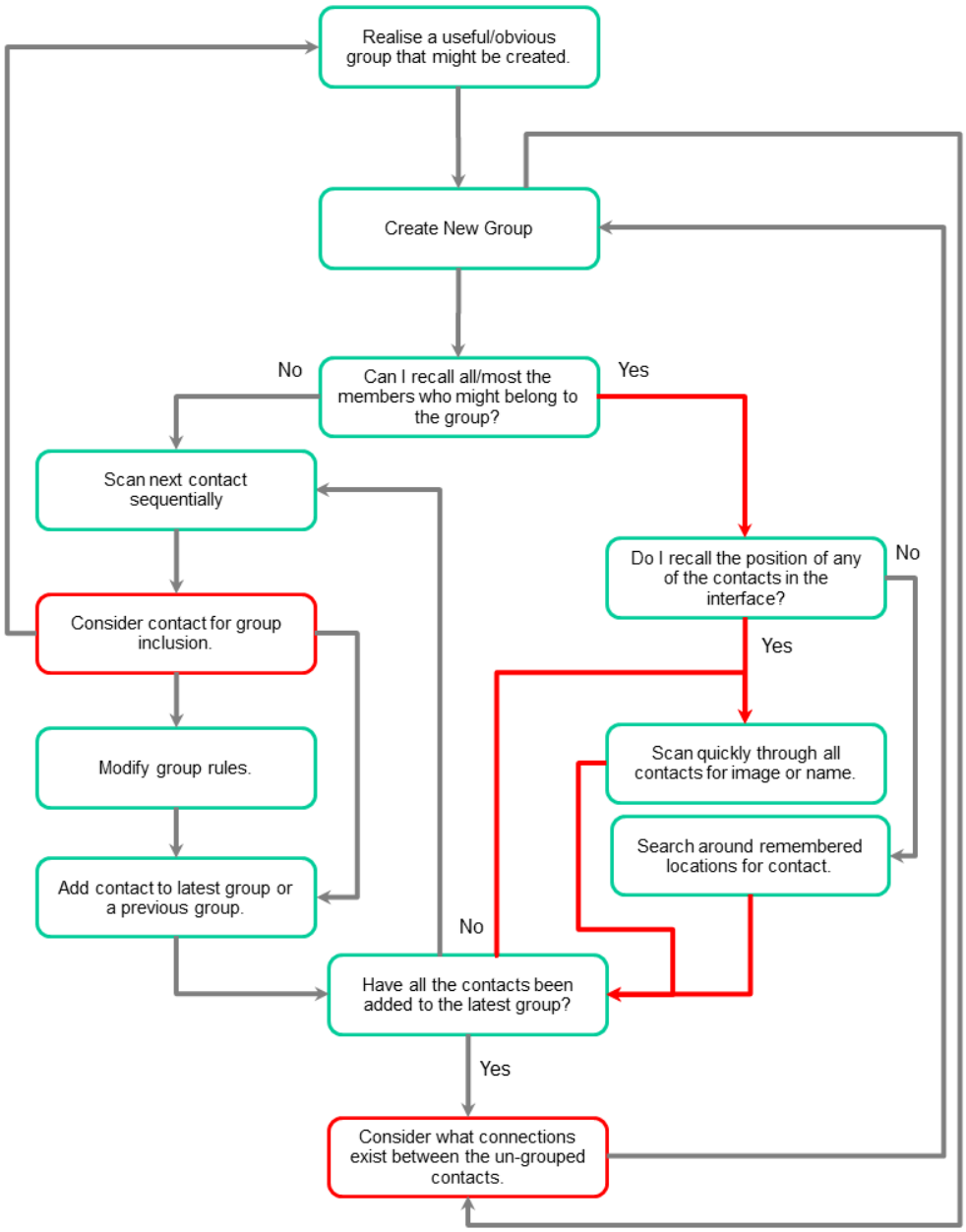


Figure 7.3: Contact Adding Strategies – Search Strategy

and the groups they perceived were there and which they anticipated would be needed to organise their contacts. They did not always necessarily succeed in creating all the groups they deemed necessary beforehand, but it was common for most of the groups to be created in one go. In contrast, when participants grouped sequentially much less consideration seemed to occur as to the overall structure of the groups that would be created, and group creation was more emergent via a process of considering obvious groups for the remaining un-grouped contacts. The approach of grouping beforehand occurred more often when participants used the egocentric interface, this could perhaps be due to the greater feedback about the structure of the groups and the ability to model relationships between groups visually, however it is difficult to determine whether there is an effect in this instance.

Some of the differences in approach might be attributed to how enthusiastic participants were about the process of grouping. Some participants seemed to particularly enjoy the exercise of modelling their social relations, they conveyed enthusiasm about remembering and identifying all of the facets of their relations and they therefore were happy to create many groups to represent that accurately. Others were highly reluctant to create new groups unless absolutely necessary, and would be much more economical with group creation, retaining fewer clearly less defined groups, which had the benefit of being simpler to keep track of and retain in memory whilst grouping the contacts, '*For the sake of not having too many groups I'm going to have continental Europe. A very British phrase.*' (Participant K).

Having created one or more initial groups, as participants began adding contacts to the group(s) they would inevitably reach an impasse where a contact would not obviously fit within the current structure - present or planned - and the participant would be forced to reassess their groupings. This would manifest as hesitation in the participant's flow of grouping while they thought about how to resolve the situation. The most simple solution to this was to create a new group, however when creating new groupings participants generally did not want to create a group solely tailored to the contact that prompted its creation. As was noted previously, participants were generally reluctant to create groups with

only one or two people in them, seeing them both as unnecessary effort to create and potentially less useful. Therefore they would often prefer to make fewer groupings with multiple people in them, even if it meant the groups needed to be broader in scope. Doing this would require the participant to consider how the contact in question might fit into a group with the other contacts present, and not solely what their strongest individual association with them might be, '*No, I'll put her in the abroad group. Because he's Turkish, and I was going to have a group Turkey and put it in there, because that's home for me. But because it's just one person, I'll put it in the abroad group.*' (Participant I). Participants would create groups to contain only one or a few contacts occasionally, but these tended to be either of particular importance to them, or a way of resolving how to group the final few un-grouped contacts, which they were less obviously able to put into groups.

As a result of this tension between group creation and economy of groups, participants would sometimes decide that on balance they would rather reassess the meaning of an existing group and change the rules for inclusion in a group, essentially broadening its scope. This would seem to be to avoid situations where a contact would likely end up in either a group of one or two contacts, or in a meaningless catch-all group. A typical example of this was a participant who while adding contacts to a group called '*Nottingham Area*', repeatedly allowed the meaning of Nottingham Area to extend to include a broader and broader radius as he came across increasingly distant contacts that didn't fit into any of his other main groups, '*Well you know actually he only lives in Birmingham, that's Nottingham Area isn't it? He only lives an hour away.*' (Participant A). This behaviour of re-defining which contacts are acceptable in a group also shows that the participants assign each group a set of rules when they create it, governing how it will be used. The hesitations in grouping would occur when those rules were stretched. At this point the participant either re-defines and expands the rules of a group or defines a new set of rules by creating a new group. The decision may be affected by how arduous creating another new grouping would be, and how inappropriate it would be to include a contact in a group. This process of rule change, was never such that contacts would be moved out of a group or

into another, the re-definition of its meaning was always an expansion that would include existing members and new ones.

Time	Duration	Behaviour	Event
00:00:00	15.066516	Create Group	State start
00:00:00	130.398696	Not Grouping	State start
00:00:15	0	Create Group	State stop
00:00:15	19.533138	Create Group	State start
00:00:34	0	Create Group	State stop
00:00:34	13.033203	Create Group	State start
00:00:47	0	Create Group	State stop
00:00:47	11.199888	Create Group	State start
00:00:58	0	Create Group	State stop
00:00:58	13.366533	Create Group	State start
00:01:12	0	Create Group	State stop
00:01:12	47.599524	Create Group	State start
00:01:59	0	Create Group	State stop
00:01:59	9.966567	Create Group	State start
00:02:09	0	Create Group	State stop
00:02:09	23.199768	Adding People	State start
00:02:10	0	Not Grouping	State stop
00:02:10	22.799772	Sequential Scanning	State start
00:02:32	0	Adding People	State stop
00:02:32	8.399916	Create Group	State start
00:02:33	0	Sequential Scanning	State stop
00:02:33	8.066586	Not Grouping	State start
00:02:41	0	Create Group	State stop

Table 7.1: Participant L Ego-centric condition coding output extract

Table 7.1 and 7.2 present extracts of the observation data analysis, which identified both the two contact adding behaviours described in the previous section and the two group creation strategies outlined in this section. Table 7.1 illustrates an

example of the prior group creation strategy, the pattern of consecutive instances of group creation can clearly be seen in the data. Table 7.2 provides an example of the repeated pattern of creation of a group, followed by adding of contacts to that group. In both examples, instances of coding of sequential scanning of contacts and searching for specific contacts can also be seen. In Appendix D complete outputs of the coding of these example extracts are provided alongside screenshots of the video analysis.

Time	Duration	Behaviour	Event
00:00:49	71.532618	Sequential Scanning	State start
00:01:59	0	Adding People	State stop
00:01:59	6.466602	Create Group	State start
00:02:00	0	Sequential Scanning	State stop
00:02:00	4.99995	Not Grouping	State start
00:02:05	0	Create Group	State stop
00:02:05	128.098719	Adding People	State start
00:02:05	0	Not Grouping	State stop
00:02:05	128.098719	Sequential Scanning	State start
00:04:13	0	Adding People	State stop
00:04:13	19.9998	Create Group	State start

Table 7.2: Participant P List condition coding output extract

7.2.2.3 Group Content

When deciding on what the content of the groups would be a variety of factors seemed to be significant in the decision making process. There often seemed to be no obvious consistent theme according to which individual participants determined how to group their contacts, with the presence and order of particular groups and types of groups varying over each condition. Looking at groups across all participants, however, revealed some common themes in initial grouping decisions and group ordering. Creating groups based upon the perceived personal importance of the contacts to the participant was one criteria that was commonly

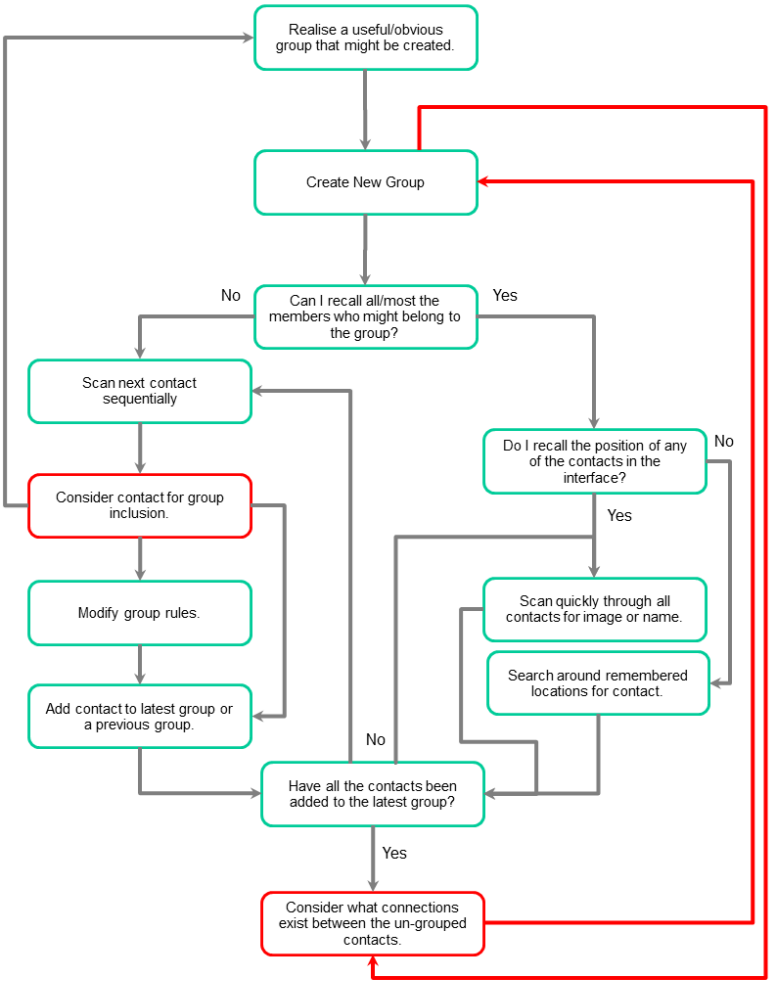


Figure 7.4: Prior Group Creation

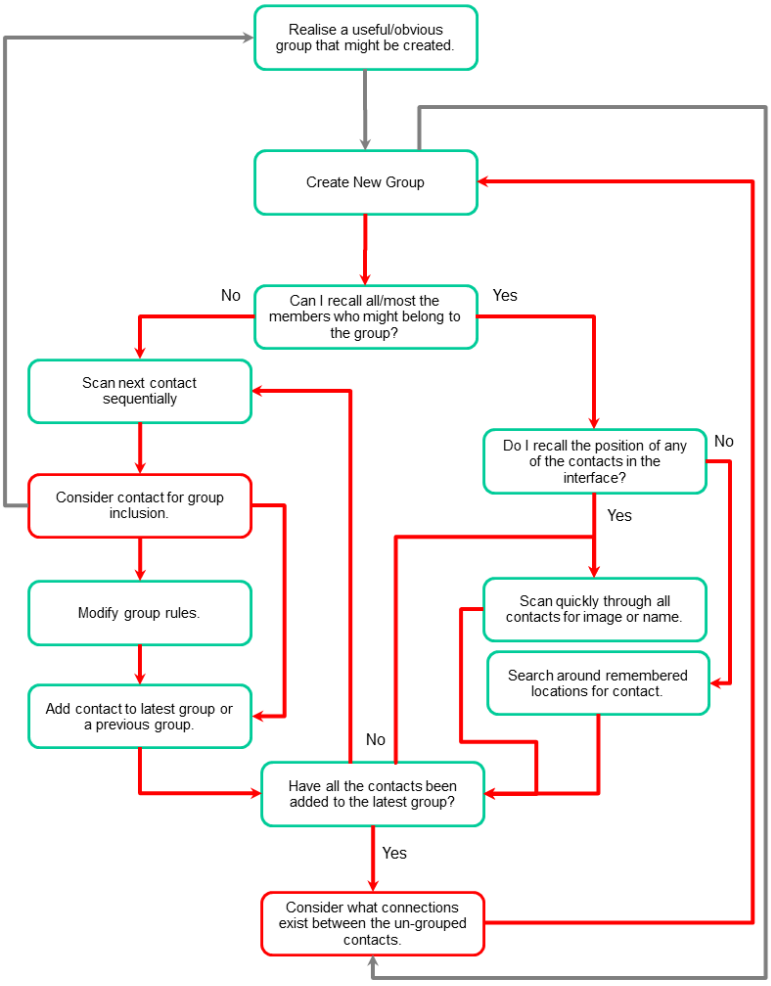


Figure 7.5: Group Creation Strategies

used. In this instance importance might be level of intimacy or how significant communication with those contacts was.

> *Participant: 'Again I'm going to start with Family, for no other reason than I feel obligated to.'*
> *Researcher: 'Why do you say you feel obligated to?'*
> *Participant: 'Just because family should come first, right?'*
> (Participant C1)

Another strategy to identify initial groups was by the number of people covered by a group, or put differently, what was the most obvious group. In this case participants would scan through their contacts and pick out a group that it was immediately obvious would include many contacts, '*So first I'm going to create a work related one, because one way or another a lot of my contacts are work related.*' (Participant I). Similar to this strategy was to begin a group with whose members the participant had the greatest amount of contact, '*Why is Office Easiest? Because that's the people I'm in touch with on a regular basis, see them every day, almost every day. It's not necessarily the people I speak to the most... well, yeah it is.*' (Participant G). These first groups were often not groups that emerged by the participant closely examining their contact data, but rather groups that they already suspected they would be creating, groups that they had a clear mental model of. Using these strategies both provided a starting point for the participant and also reduced the amount of work they had to do to make sense of their contacts. It was what was 'easiest'.

Occasionally one of these strategies would be used across a whole condition, starting with a particularly significant group and working outward, '*So the first thing I'm going to do is add myself. And the next group I think I'm going to add is an administrator group because I rank those actually quite highly on the previous section, because they're a gateway to people who help you do your job. Oh and then I'll have a supervisor group, these would be my PhD supervisors. These people outside of personal were the top ranked because communicating with them is often really crucial.*' (Participant E). More generally, however, after the first group or two no discernible pattern would emerge regarding a criteria for creating new groups, and after beginning groups would be created more by revelation of

spotting a group they should make or to sensibly group all the contacts, rather than by a predetermined criteria.

Analysis applying the grouping framework from Chapter 5 to the resultant groups created by the participants in both of the Free Grouping tasks revealed common types of groups created across multiple participants, as was the case in Chapter 5. Table 7.3 represents the encodings of these groupings using the Chapter 5 categorisation. Examining the grouping results for the List interface and the Ego-centric interface on the whole revealed only small variations in the percentage differences in group types between the two conditions.

Category	Ego	List	Total
Behaviour	**2**	1	3
Social Behaviour	2	1	3
Characteristics	**2**	**5**	**7**
Ethnicity or Identity	1	2	3
How Others Make The Participant Feel	1	3	4
Personality	0	0	0
Physical Characteristics	0	0	0
Same Age	0	0	0
Shared Name	0	0	0
Event	**1**	**2**	**3**
Ambiguous Event	0	0	0
One-Off Events	0	0	0
Recurring Event	1	2	3
Expertise	**6**	**3**	**9**
Factual Skills	4	0	4
Perception of Skill	1	0	1
Similar Fields of Work or Study	1	3	4
Social Skills	0	0	0
Family	9	12	21

Category	Ego	List	Total
Children	0	0	0
Relationships	**2**	**2**	**4**
Relatives	6	10	16
Frequency of Contact	**1**	**2**	**3**
Interests	14	10	24
Likes-Loves–Interests	0	0	0
Shared Interest or Hobby	14	10	24
Location	**21**	**10**	**31**
From Same Place	1	1	2
Live	15	7	22
Live In	12	4	16
Lived In	2	1	3
Living Together	1	2	3
Physical Proximity	5	2	7
Meeting	**13**	**8**	**21**
Know Through Another Person	5	4	9
When I Met Them	8	4	12
Organisations	**60**	**43**	**103**
Broad Organisation	11	8	19
Past Work or Study	1	0	1
Project Together	6	4	10
Role	27	15	42
Sub-Organisational Groups	10	7	17
Work or Worked With Together	5	11	16
Relationship	**32**	**31**	**63**
Friends in a...	10	12	22
Interest	0	0	0

Category	Ego	List	Total
Location	8	6	14
Organisation	1	6	7
Time	1	0	1
Intimacy of Relationship	22	17	39
Mutual Friends	0	1	1
Positive & Negative Relationship	0	0	0
Total	**345**	**273**	**618**

Table 7.3: Prototype Grouping Categorisation

As in the previous study on grouping, Organisational, Relationship, and Location came out as the most common forms of groups created. And as was also the case in the second study the nature of the groups created was often more complex than the categorisation implies. Of the groupings created, 16% were multi-dimensional in nature i.e. the groups were named or described such that they sat between categories, invoking multiple dimensions to describe the group, e.g. 'Nottingham Colleagues'. While these were categorised according to what their primary intent spoke to — in this case Organisational — they could justifiably have been categorised in more than one group, which was Location in this case.

Within these multi-dimensional groups it is again the same top three factors that are most commonly combined – with a similar low distribution through the remaining categories. The location, form or intimacy of relationship, and organisational role or affiliation, appear to act as the most important filters or modifiers with which to describe more specific sub-sets of people.

The grouping framework accounted for the vast majority of the data. The most notable difference in terms of the forms of groups appearing was the presence of a greater emphasis in some groups on emotions and behaviours. This included groups that described how participants viewed the state of some relationships, such as '*I don't know, it's difficult*', which described contacts a participant had stressful relationships with.

'Frequency of Meeting' did not occur in this data set, however, frequency of contact as a grouping basis did occur repeatedly. In light of this the framework was adjusted to include this theme in the data by altering 'Frequency of Meeting' to 'Frequency of Contact', with the intention of the new category being able to represent both face to face and digital correspondence.

Common to nearly all of the participants was a strategy of using negation of an existing grouping, to help define a new grouping. The groupings that were negated were typically already defined as part of the study, but were also in some cases groupings they had yet to create, but which they already had in mind, either because they had planned ahead or were aware the group would likely emerge from their contacts (see group creation strategies e.g Fig 7.4). These groupings based on negation were often catch all groups used to denote splits in relationships, such as those which occur within work contexts between those with whom a personal as well as professional relationship exists and those with whom the relationship is purely professional, '*Everyone who is in here is in this building, but is not an office mate or in the CDT. Everyone who is ((location))... Not a very good sorting group, sorry.*' (Participant H). In this case the participant is not sure of the utility of this grouping precisely because it is a catch all group. It rather served the purpose of collecting together contacts with something in common, but who fell outside of a more significant grouping, to the participant, that had the utility of defining more useful or important relationships. This was a fairly typical behaviour in that participants seemed to use this strategy when they wanted to talk about people outside of what they saw as their main focus of communication, work, or interaction.

While very different in general structure, the two grouping interfaces did not produce any significant differences in the categories of groups users generated. The force-directed grouping interface did however result in a significant increase of 45.26% more groups being created overall ($t(16) = 2.19$, $p < .05$). The design of this interface, particularly its facility for sub-groups and direct feedback about the structure of the groups, appears to encourage users to represent their mental model of contacts in greater detail. The increase did not manifest across all users, some of whom created more groups with the List grouping, but it does indicate

	Ego	List
A	10	5
B	7	7
C	5	5
D	4	7
E	10	10
F	8	3
G	15	8
H	14	12
I	11	12
J	6	2
K	19	11
L	10	4
M	23	9
N	8	6
O	10	6
P	15	8
T	16	16
S	8	6

Table 7.4: Ego and List Condition Group Quantities

that the interface design can be used broadly to influence the form and specificity of the groupings, if not the descriptions of the groups being created.

Portions of the increase can be attributed to a greater level of modelling of the social structures the participants perceived. Where in the List interface participants might have created two groups - Nottingham Friends, Nottingham Colleagues - within the Ego interface it was common to see relationships such as these split into three groups; a higher level Nottingham group, with Colleagues and Friends represented as individual sub-groups of Nottingham. In some cases differences between the conditions manifested because participants were prepared to subdivide groups in this manner in the Egocentric interface, where they had not been prepared to do so at all in the List interface. It also seemed to encourage some participants to create smaller, more specific groups, including some types of groups that did not emerge in the results of the List grouping condition. Creating sub-groups such as 'Partners' to denote people who were the partners of

people in a social group, rather than including them in it was one example of this. That there was a direct linking line from the social group to the 'Partner' sub-group seemed to make this diagramming of relationships more natural for some participants.

The desire to make explicit and map certain relationships occurred across the groupings of multiple participants. To resolve this participants seemed naturally to want to link people to other people, without the intermediary of creating a group, to make clear that a certain important relationship existed between two people. '*I can't link people to each other can I? ...in that case I'll call this group the "[Surname]s" which is, that's my boyfriend's surname. There are a couple of people in here, like his sister and his sister-in-law, so I'll put them in the same group*' (Participant I). This behaviour only appeared to occur in circumstances where the contact(s) in question held no greater importance to the grouper than their relationship to the contact associated them with. These might be described as secondary contacts to a more important contact, known through their association to the more important contact.

As a result of this behaviour it was also common for the strength of contact association to lead participants to group contacts 'inaccurately', as a means to resolve their desire to retain these associations and relationships, and in some instances to also minimise the amount of groups they needed to create, '*I guess she can go in pole... We don't do pole together, but I met her at my Pole instructor's wedding, so the thing we have in common is pole. Otherwise I don't really know what group I would create for her, as she has nothing in common with the other people here. So instead of having a separate group for her...*' (Participant I).

This behaviour of grouping according to strongest association manifested itself across all conditions. It was very common for participants to value strong associations over correctness of information in relation to a contact. Past and occasionally also anticipated future associations would often be invoked in group construction to describe contacts. This most typically occurred in relation to groups named to describe locations. For example, contacts whom the participant knew in one city, but who had subsequently moved away, were often grouped as belonging to the original city due to the strength of the original association,

'*This is a friend of my ex who lived in London, but now lives in Derby. I don't associate him with Nottingham so I'm going to put him in London.*' (Participant C). On the occasions where the future circumstances of contacts were prioritised in grouping them it was typically due to a judgement that it was important to reflect their new location, role, or organisation, as it had significance for how they would use a group or think of the contact.

Strength of association also seemed to emerge as a behaviour while participants sequentially examined their contacts. As each contact was encountered, participants would often explain as part of the think aloud protocol what each of the contacts represented to them, in terms of the role they played in the participant's life, '*She's an administrator, PhD Student, friend from home...*' These seemed to be expressions of the most significant associations the participants held for each contact. Grouping by strength of association, however, was largely absent as a behaviour when it came to groups that were central to the participant's social or work life. Those immediate social groups would override or become the most significant association, even if there was a strong past one, when a contact belonged to a social group in one of the participant's current locations, organisations, or local friendship groups which were considered more important or useful.

The names of the groups functioned as signifiers for the group, rather than describing strictly the rules according to which the groups had been created. The names were something to call and re-call groups by, and did not necessarily denote the meaning or content of the group, even though they often appeared to be prescriptive in their descriptions. This was most apparent where groupings referring to specific friendship groups were given a moniker, such as '*Canning Circus Crew*', because participants would occasionally find it difficult to form succinct descriptions that would delineate the group from a more generic 'friends' grouping, '*I feel bad saying friends, some of them aren't my friends. But in my head they are all a group together*' (Participant B).

7.2.2.4 Group Structure

Groups such as these also reveal two distinct underlying structures to the groups participants created. At the two extremes there appeared to be those groups that

had a clear and stable social structure and groups that had a criteria that was inherently amorphous. The stable groupings were those where the participant had a clear mental model of a particular social structure that has a constant membership, such as 'family', a specific friendship group e.g. '*Hannah and Co*', or a specific group of colleagues that changes little in membership. Essentially these were groups of people who know each other and interact together on a regular basis over a long period of time.

By contrast the amorphous groups represented one of multiple convenient or useful ways of filing or describing people, with whom a multitude of subtle relationships potentially existed, but not necessarily a significant unifying social structure. The members of these groups did not necessarily come into contact with each other, though these were groups that might contain within them social networks. Where they did contain networks they were ones which participants were happy to include members outside of that network or multiple other networks together as well. The focus of these groups was sorting contacts according to common attributes rather than social structures, such as roles e.g. 'Administrators', or a common catch-all relationship group e.g. 'Friends' - broadly the forms of groups identified in the framework of group types (Chapter 5). These less defined amorphous groupings were more common than the stable social groups, but both were present across nearly all participants.

Uniting the underlying structures of these groups was how participants rated the importance of contacts. Examining each group in relation to these ratings reveals a tendency for contacts with similar importance ratings to be grouped together. Across all participants the groups they created tended to contain contacts with only a relatively small range of ratings. This behaviour was not universally true, there were many groups that did not particularly adhere to this rule, but often even when the standard deviation was particularly low the majority of contacts would tend towards the same region of the scale - high, middle, low. The standard deviation over all of these groups was on average 1.51. Where there was a greater standard deviation a factor in this would be groups that were less well defined or focused, e.g. groups that contained all of a participant's friends. Equally groups with very low standard deviations named 'Acquaintances' might understandably

have many contacts with the same rating, as membership of that group would imply their importance is viewed at a similar level. Another feature of this result was that it was common to see many contacts all with the same rating +/- 1, but then also one or two contacts standing out with much higher or lower ratings.

Participant	Group	Standard Deviation of Group
A	Common Friends	3.43
A	Facebook Friends	2.88
A	Work	2.85
A	Non-wrok Friends	2.64
A	Staff	3.16
B	Russian Department	0.00
B	Research Groups	0.00
B	Lincoln Hall Tutoring	0.00
B	Outreach	0.00
B	PG Students	0.00
B	PG Administration	0.00
B	Societies	0.00
C	Wouldn't contact	1.37
C	University	2.14
C	Friends	1.86
C	Closest Friends	1.05
C	Family	1.50
D	Southampton	1.07
D	Family	1.83
D	TRG	1.00
D	UniNottingham	0.92
D	Professional	1.66
D	HFRG	1.59
D	Other friends	0.58

Participant	Group	Standard Deviation of Group
E	colleagues general	0.50
E	supervisor	0.00
E	colleagues shared projects	1.90
E	DTC	2.19
E	colleagues shared interests	0.58
E	housing	0.00
E	External partners	0.00
E	participants	0.00
E	intern	0.00
E	Adminstrators	1.66
E	Friends	1.83
G	External work contacts	0.00
G	PhDs	0.00
G	Friends/family	1.34
G	Staff colleagues	0.92
G	vball	2.70
G	Office	1.14
G	housing	3.11
G	ma group friends	2.57
G	departmental/official	3.13
G	german contacts	1.67
G	close friends	1.80
G	Friends	2.11
H	Political Psychology	1.17
H	partner in 'crime'	0.00
H	sunday nerds	0.58

Participant	Group	Standard Deviation of Group
H	pun mates	2.29
H	Canning Circus Crew	3.21
H	CDT	2.41
H	msc psychology	1.50
H	supervisors	1.41
H	i don't know, it's difficult	0.00
H	uni, professional	1.39
H	uni, nonprofessional	1.98
H	research contact or participant	0.58
I	Conference	0.00
I	Retford	4.04
I	Work	2.84
I	US	1.67
I	Phd people	3.50
I	Rumbles	1.73
I	Pole	3.58
I	Exes	0.50
I	Uni Staff	0.00
I	Heyford	3.70
I	Bezzies	0.00
I	Friends	3.41
J	Mex Friends	1.96
J	Notts Close Friends	1.22
J	Academic Notts	1.82
J	Pach Close Friends	2.95
J	To be in touch a bit	2.00

Participant	Group	Standard Deviation of Group
J	Family & Others.	0.00
K	Best Freinds	2.83
K	MyFamily	2.89
K	Freinds	2.34
K	Office	2.46
K	Family	1.51
K	HEMA	1.87
K	Masters	4.16
K	PhD	3.08
K	Aquaintances	1.85
K	Partner	0.00
K	notsignificant	0.00
L	other	0.00
L	I have to speak to	2.40
L	Through Emily	1.68
L	Enjoy speaking to	2.35
M	Undergrad	0.84
M	Family	1.41
M	Paris	1.73
M	Amiens	0.00
M	Work - not uni	0.58
M	Folkestone	1.00
M	Paul's friends	1.52
M	other	0.00
M	Admin	1.15
M	Office friends	1.48
M	French department	1.42
N	keep in touch	2.25
N	not in touch with	3.08
O	Family	4.00

Participant	Group	Standard Deviation of Group
O	Irrelevant	0.00
O	Acquaintances	1.14
O	Professional	1.56
O	Close friends	1.52
O	Friends	2.54
P	home friends	3.35
P	harlaxton friends	0.00
P	family	2.28
P	fellow medics	2.40
P	non medic uni friends	2.44
P	housemates	1.10
P	close contacts	0.00
P	uni friends	2.11
S	Home friends	3.25
S	Volleyball	2.43
S	Academic	2.69
S	Other	0.00
S	Sport	1.38
S	Uni friends	1.78
T	ecr sig	2.07
T	hfrg phd	1.76
T	rnib	2.12
T	family	0.00
T	iehf membership	1.67
T	vr-hyperspace	1.39
T	msr gd	0.55
T	board games	0.00
T	school	0.00
T	hannah etc	0.00

Participant	Group	Standard Deviation of Group
T	londonjew	1.48
T	london games	1.00
T	other engineering	0.00
T	other work	4.24
T	iehf	0.96
T	hfrg	2.34
T	notts jews	1.22
Average Standard Deviation		1.51

Table 7.5: Free Grouping Condition Standard Deviation of Group Contact Importance Ratings

7.2.2.5 Importance Rating

The process of rating the importance of contacts appeared to cause a degree of discomfort for most of the participants. Participants were asked to judge the importance of each contact for daily communication, and not to judge the importance of their relationship with that person. However, despite emphasising this, many participants did not feel this separated it as a judgement on the worth of the person to them when they rated them. This led participants having a sense of discomfort during the rating process. This discomfort was also present, although to a much lesser extent, during the contact grouping tasks where, for example, inclusion or exclusion of people from groups such as 'friends' or 'close friends' required making a judgement about the nature of those relationships to ensure there was some degree of internal logic to the membership of those groups.

Often these anxieties seemed to emerge when rating people they considered to be friends, due to feelings of guilt stemming from implicitly creating a hierarchy of friends. Although their friends would never see these judgements the act of simply thinking about and making concrete these judgements made the participants uncomfortable. *'It feels a bit harsh, to organise, to be like "You're okay, but I*

can't really be arsed." (Participant P) In this case the participant appears to feel guilt because the grouping, to her, expresses an implicit level of apathy she felt about maintaining that relationship.

Several participants were keen to take a different approach than most when describing relationships. They suggested that they did not like to think of their contacts in terms such as close friend or acquaintance, which had been commonly used throughout this study by the majority of other participants to indicate intimacy. Rather they said that they thought more in terms of people whose company they enjoyed and those they disliked. These approaches were not an issue for the categorising their groupings in terms of their description, as in the few conditions where this behaviour was present the coding scheme was flexible enough to represent the intent of the group, even though a different approach had been taken. With regards to the results presented here this perhaps caused fewer groups, in some cases, and slightly odd groups relative to the other participants. There was no reason to believe these approaches are not representative of how a minority of people might group, and should not have had an adverse affect on the data. A short coming of the coding scheme here is perhaps that while it captures the forms the groups might take, which may show where the emphasis or the range of groups lies, it does not record the logic of the approaches each participant took. These aspects were rather described through observation and interviewing.

7.2.2.6 Group Satisfaction

When asked about the groups they had created, having reviewed them and used them to perform information tasks, the participants generally stated they were satisfied with them. And when asked whether they would retrospectively change any of the groupings, and how they would change them, they stated that they were happy with them as they were. The complaints of participants who stated that they were dissatisfied in some way were significant, however, in that they highlighted what was important in terms of creating a useful grouping.

Those who were dissatisfied in some way all related this to feeling that the groups they had created were unhelpful or inadequate for completing the information tasks that were set. For one participant this was due to being unsure of how they

had made a distinction between two groups when they came to use them, '*With the friends and the close friends, I wasn't sure where I had put friends when I had to look for them.*' (Participant G). Several participants were broadly satisfied, but had created one or two specific groups that turned out to be unhelpful when completing the tasks. This related to participants being unsure where they had grouped contacts, because they had misplaced them or were unsure which of several groups they would be in. For others the issues stemmed from the inclusion of one or more contacts in a group that prevented it from being as useful as they wanted it to be.

> '*There's the "Nottingham Cheese Group". There's one person there I wouldn't eat out with, but with that group of friends I'd eat out with, or eat with at each other's houses quite a lot. There's one particular person who'd come up in that if I clicked on that group, who I actually wouldn't include in that group. And that would include everyone I'd actually ask. So I'd probably break that down into people who I would eat with.*' (Participant T).

The participant thought that although the person whose inclusion made the group less useful did fit into her criteria for the group, they did not fit the social criteria that she wanted to use the group for. Similarly, three other participants felt the groups they had created were more generally unsuitable for the tasks that they were set. This was related to feeling that the groups they had created were too large or general in nature to be suited to any of the specific tasks, 'Not so very satisfied. Some questions it was very pinpointed and I thought yeah, I just need to click this group. And sometimes it was related to more stressful things, and I for example just had the PhD group, and this is too big.' (Participant K). These participants suggested that the solution to this would be to create more detailed groupings.

> '*Some of them were too big, I feel like I should have had more smaller groups, because then I would have just been able to click on a group [to share information]... A London group might have people I wouldn't necessarily see or communicate with together. So I'd break them down into social groups within it.*' (Participant T).

When asked why they didn't break down the groups into smaller or more cohesive units initially, concern was expressed about maintaining the manageability of the amount of groups and the work creating them and maintaining the same consistency of detail throughout entailed, '*I already had a lot of groups and I was trying to– it was the time it was taking, and just the managing of it. It's easier to manage fewer groups, but when it comes to using them it's not so useful.*' (Participant T).

All of the participants were also asked whether they perceived any difference between the groups they created with each of the two grouping interfaces, and all except three responded that they were broadly similar in nature and that they did not perceive any particular difference between the two sets of groupings. This participant's response was typical of those who could not see any difference, '*I don't think that they were that dissimilar. I think my groupings were quite similar. I thought it had the same intention, it was just easier to [use one grouping interface] than the other.*' (Participant C) All of the three expressing that they perceived a difference in the nature of the groups described how they thought that the Egocentric interface resulted in them creating either more groups or more detailed groups, by which they meant that they broke their groups down into smaller components. 'The [egocentric grouping interface], I found that very very useful I was happy with that one... Because it was in smaller groups, I could divide the bigger groups into smaller groups, and the layout was quite clear to me.'(Participant G)

7.2.2.7 Grouping Interface Design

In terms of which interface participants preferred to perform the grouping with, opinion was evenly split between the two interfaces, with strong preferences for one or the other. The split was defined along two lines, those who preferred the List because of its static and ordered nature, and those who preferred the Egocentric interface because it gave them a better visual overview of the structure of the groups they were creating.

Those with a preference for the List interface believed it be easier to sort contacts into groupings due to its alphabetic and ordered nature, '*I would definitely say the gridded option, because it was easier to systematically go along, scan through the columns, go along the columns, up and down, side to side, so that I know that I've looked at all of them.*' (Participant B). It is possibly worth noting that of the nine List preferring participants three of them mentioned that despite finding that interface more functionally suited to the task, that they also found the experience of using the Egocentric grouping interface more enjoyable or fun.

Those who preferred the Egocentric interface did so because it provided better visual feedback to them about the structure of the groupings and how they had placed contacts in the context of those groupings they had created. '*It was just easier to visualise, I'm a visual person, it's easier to see what's going on there.*' (Participant T). This phrase, '*I'm a visual person*', was used by three of the participants to describe why the preferred the Egocentric interface, implying that they believed they were inherently more receptive to this form of interface where there was more of an emphasis on visualising connections between contacts and groups. It was common for much of the reasoning behind this preference to be focused on the presence of the links connecting contacts and groups, for many this seemed to provide a greater overall awareness of the groups and contacts they had created and how they related, '*if you see it with these links you are more likely to add people to two groups and be more aware of how people relate to individuals.*'(Participant H).

7.3 Person-centric Interface Evaluation

7.3.1 Interface Evaluation Methodology

7.3.1.1 Study Design

In the second part of this study, the 'person-centric' interfaces (see Chapter 6) for performing information re-finding and sharing tasks are evaluated. These interfaces use the contact groupings created in the first part of the study as the basis for constructing representations of the participant's contacts in the interfaces to

support performing information sharing and re-finding tasks. In order to evaluate the interfaces the participants were first asked a series of questions about their own personal data, from which information re-finding and sharing tasks could be generated specific to the participant. The aim of this was to generate a more natural series of tasks to evaluate the interface with that would allow participants to use their knowledge of their own data in completing the tasks (Whittaker et al., 2004; Elsweiler and Ruthven, 2007), as opposed to setting the same tasks for all the participants using unfamiliar data, that is not part of the participants' own personal data.

The tasks that were generated from this questioning were evenly split between information re-finding and information sharing tasks, in which participants were asked to use the prototype 'person-centric' interface to either re-find a piece of information or to share a piece of information with one or more of their contacts. Participants performed these tasks in a balanced order across the four different forms of representation of the participants' contact groupings – of Egocentric, Group Link, List, and Gravity (see Chapter 6) – each representing design a idea generated by participants in the second study (Chapter 5). The aim of asking participants to do this was to understand how the different design ideas might hinder or support participants in performing these two information tasks. In the case of tasks for information re-finding this utilised the data (emails, messages, files, links) collected about the participant's contacts and made it accessible via their contacts in the interface. The study design meant that the information that participants were asked to re-find was always present in the information that had been downloaded about the contacts. While in the case of information sharing tasks, these were performed by clicking on contacts – or whole contact groups – in order to add them to a list that represents people the information that prompted the task would be shared with.

As in the previous component of this study, a combination of screen capture and think-aloud protocol was used to record the participant's thoughts and interactions with the interface. It was suggested to participants that they should consider what information they might be accessing during the study and whether they would be comfortable with it being recorded by the screen capture. If the

tasks set required them to access sensitive material they were told they could ask for an alternative task to be set. Additionally, as in the first part of the study the same ethics measures applied and participants were able to note down the time that any sensitive material might have been recorded, so that they may request that section to be later deleted.

Finally, a semi-structured interview was conducted in which the participant was questioned about their experience with the interfaces from both parts of the study. This sought to explore the participant's views and experiences using the interfaces in order to understand how the various design ideas supported or failed to support the different information tasks being performed.

In order to analyse the screen capture and think-aloud protocol the collected data was coded in several passes (Woods and Dempster, 2011) to identify different interaction behaviours, sustained pauses in performing the tasks, and different forms of comments participants made as part of the think-aloud protocol. Similarly, transcripts of the interview results were thematically coded (Braun and Clarke, 2006) to identify commonalities in the experiences and behaviours of participants when using the interfaces.

7.3.1.2 Procedure

Participants were asked to write down one or more answers to a series of questions (below) about their personal information. The questions (below) prompted the participants to identify two types of information; information that they might want to share with others, and information they had previously been sent.

The questions participants answered were as follows:

1. For sharing tasks:
    - A piece of good news you have had.
    - One or more articles/web pages/videos etc. that you have recently seen, that may be of interest to others.
    - A document you have recently completed for a particular project.

2. For re-finding tasks:
    - A project you have been working on recently.
    - A few links that have been recently sent to you, which you might want to look at again.
    - A few files that you might need to find again that were sent to you.
    - A meeting you recently arranged via email (not event invitation)

Having answered the questions, participants were shown the 'person-centric' interface described in Chapter 6. The functionality of the system was methodically explained, highlighting how each element worked and answering any questions the participant had. Following this explanation, participants were given time to familiarise themselves with the interface. While the participant examined the interface, tasks were generated for the participant based on the answers supplied about their personal information. Once the participant was happy that they were comfortable using the interface, they were told that they were going to be set a number of tasks in which they would either need to share or re-find information, based on their previously supplied answers, using the person-centric interface. They were told to complete the tasks in their own time and not rush, only stopping when they were completely satisfied with the result. Participants were also asked to describe at the start of every task what information they were looking for, or what they wanted to share with people and what their goal was.

Participants were set tasks using four different interface conditions, corresponding to the four contact grouping presentations described in Chapter 6 of Egocentric, List, Group Link, and Gravity. The tasks that were set for the participants were either information sharing tasks or information re-finding tasks; two of each type of task were set for each of the conditions. The tasks were generated from the following list, with the researcher providing a task based on the answers that the participant gave about their personal information (where applicable).

Tasks:

- Invite people to a party at your home.

- Share the interesting article/video/webpage you have recently seen with those you think may be interested.
- Share the document you recently created with those who need to see it.
- Share the good with those who they want to know about it.
- Find the most recent email concerning a project you are working on.
- Find one of the links that have recently been sent to you. The researcher identified which link the participant had to re-find.
- Find one of the files that were sent to you. The researcher identified which link the participant had to re-find.
- Find the date / time of the meeting you recently arranged.

In each set of sharing tasks set in each condition, one task required sharing information with a single or several specific people, and one required sharing information or an invitation with a larger number of people. In the case of the re-finding tasks, one of the set tasks would involve locating a recent message, and the other would involve locating either a file, a link, or a meeting. This split in the tasks was to ensure that a balanced variety of tasks was performed with each of the contact grouping presentations.

A semi-structured interview was then conducted with the researcher querying the participant's experience with each of the different elements of the interface, including the group construction, satisfaction with the groups, use of group filtering and re-structuring buttons, accessing information through people, and preferences regarding the different presentations. The questions addressed different components of the interface design, and the participants were shown the element(s) in question in the interface so that it was clear was what was being discussed. As in the first half of this study, the observations and interviewing were coded thematically (Braun and Clarke, 2006) to identify participant behaviours and experiences in the data. The questions used in these semi-structured interviews are in Appendix E.

7.3.2 Interface Evaluation Results

7.3.2.1 Person-centric Interface Preferences

During the interviews, participants were asked to order the interfaces from their favourite to their least favourite interface to use when completing the tasks that were set them. The mean frequency of the interfaces' placements in that ordering, where the preferred interface was scored 1 and the least liked was scored 4, are show in Table 7.6.

The results of this scoring show that the List interface was the most preferred interface on average, with the Egocentric interface being slightly less preferred for performing the tasks, and the Gravity and Group Link interface being considered the least useful. Given the closeness of the results, the distribution of the placements provides a greater degree of insight than the mean. The List interface was the favourite interface to use in half of the cases, and barring one result where it was rated third, it was rated as the least preferred interface to use in all other cases. This split in preferences was similar to the one observed between those who preferred the List grouping interface and the Egocentric grouping interface.

Due to this split in preferences it is difficult to say that this indicates a clear preference for the List interface over the other interfaces. Their comments about each of the interfaces provided more insight into the participant's preferences for using a 'person-centric' interface and the reasons for the stark differences in opinion.

Participant	List	Gravity	Group Link	Ego
A	1	2	4	3
B	4	3	1	2
C	1	2	3	4
D	4	1	2	3
E	1	2	4	3
G	4	3	2	1
H	1	4	3	2
I	3	4	1	2
J	1	2	4	3
K	1	4	3	2
M	4	3	2	1
O	4	1	3	2
P	1	3	2	4
S	1	3	4	2
T	4	3	1	2
Mean	2.3	2.6	2.6	2.4

Table 7.6: Participant Ratings of Interface Preference

7.3.2.2 Gravity Interface

While the Gravity interface (see Figure 6.7) was only rated with a single first preference, participants would comment that this interface was, '*probably the most visually appealing*' (Participant E). In contrast to the messiness of the link-based interfaces, the absence of links in this design provided a much clearer presentation of the groups, with a greater emphasis on colour. However, the lack of links or grid removed any form of visual order between contacts and groups and made it significantly more difficult for many of the participants to locate contacts. As a result, it was therefore considered less useful.

7.3.2.3 List

The reasons that the List interface (Figure 6.4) received the highest number of first preferences were relatively straightforward, and similar to the reasons given when the Grid grouping interface was liked by participants. Participants felt it provided a clear, ordered, and stable representation that let them quickly perform the tasks they were engaged in. Participants would speak of this design largely in

terms of its utility: '*My favourite in terms of practicality, I think it's a lot easier to see where everyone is, they're not floating around and you can see exactly where they are*' (Participant C).

Conversely, among those participants who did not like the List presentation, a common complaint was that they had to scroll to see all the contacts and groups. Participants typically could not see all of their contacts at once due to the number of contacts and groupings made.

> '*I think normally I'd say I prefer the listing one, but at the same time I couldn't remember which of the groups I'd actually made. Sometimes I'd forget to scroll down and I'd forget, "Oh we've got PG administration right at the bottom here." It required me to remember more from the off what the groups were.*' (Participant B).

Further questioning of those participants who had mentioned expecting to like the List interface focused on whether their dislike of the List interface stemmed from this specific design detail (the inability to see all of their contacts at once). This feature of the design of the interface was partly a result of functional necessity, and partly intended to mimic the fact that in many interfaces, e.g. Facebook (in September 2015) it is typical to have a large number of contacts and as a result to have to scroll to see them all. Participants were asked if certain aspects of the design of the presentation were changed, such that they could see all contacts and or groups, whether they thought they would feel differently towards it. In each case they maintained that they would still not prefer the List interface, stating that the reasons they preferred the link based interfaces still made those interfaces more compelling to them. This suggests that their preferences may be more fundamentally influenced by the design in this case than some of examples of the issues with the interfaces they listed.

7.3.2.4 Group Link & Egocentric

Due to the similarity of these two interface designs, in much of the discussion with the participants they considered these two interfaces as relatively interchangeable. The only difference presentation-wise between the groups was that the Egocentric

interface represented social structure through sub-groupings, while the Group Link interface represented social structure through shared membership of contacts between groups (see Figure 6.6 & Figure 6.5).

The presence of the links in these interfaces was very important to some participants. The visualisation of the links between contacts and groups seemed to make the relationships clearer and more concrete for certain participants, '*...because they symbolise the connection... It adds clarity, to me, belonging, sense of tiedness*'. (Participant G). Although in contrast, for others the links were variously patronising, unnecessary, or cluttered the interface too much. The split among participants between having a preference for the link-based interfaces or the List interface was commonly framed in terms of their like or dislike of the presence or absence of the links in the interface designs. For example, one participant described how the presence of links meant that they had to perform more mental work before finding a particular contact.

> '*It was a mental process... for example, someone who is a work friend and is also in French, I don't necessarily need to see that they are both French and work, I know that for myself, and then I use that to find them in one or the other cluster. So in terms of performing tasks, in order to have this Group Link be of any benefit you would have to train your mind to kind of think what two clusters links this person, which just seems counter-intuitive.*' (Participant A).

This participant felt that since they already retained the intricate knowledge of their social relations, there was no need for the interface to attempt to highlight the structure of those relations. Rather, the participant wanted to be able to find and make decisions about how to perform tasks by applying their knowledge of the social relationships to the interface.

In other instances, those same multiple connections worked as landmarks enabling participants to navigate better around the space, and to identify groups. Additionally, visually representing the group structure provided greater awareness of what contacts they had and how they were linked, that some participants valued when completing the set information tasks. This was particularly true for those

tasks related to sharing information where the visual overview of all groups and contacts acted as prompt to remind the participants of whom they may wish to include in any group: '*Structure was important to decide who to share with, decide quickly who to share with.*' (Participant K).

The following two quotes illustrate why some participants valued the shared membership approach of Group Link and the group structure of the Egocentric interface respectively:

Group Link - '*I suppose it helped reinforce the multiple roles that certain people have to play. It's kind of useful because you've got "Lincoln Hall" training and postgraduate students there, and in the middle is [contact's name] because he's in both of those, so it helps me locate him easier. It also helps if I'm trying to organise an event, it helps me to think who do I need to— If I'm inviting postgraduate students, do I also need to be thinking of people from other groups to organise, to invite when I'm organising an event.*' (Participant B).

Egocentric - '*I could see it better, how it relates. What I liked most about this presentation was that a line was going from me to, for example, the university, to the UK. Whatever I had, there was a line. I could see that if a person belonged to a group, and another one, how they bridged together. I really liked to see that, because sometimes one person is my friend, but also a colleague. Both at the same time, and you could actually see that. So the presentation together. The overview. Because I like to have everything in one view, and then go into the detail.*' (Participant K).

While the representation of social structure was seen as a boon, the implementation, or perhaps to some extent inherent nature of this form of presentation, was that participants also found it 'messy'. Depending on the nature of the groupings they had constructed, how much overlap there was between groups, how large the numbers of contacts in the groups were, and how many groups there were, the links could obfuscate the structure and make it more difficult to locate contacts. Additionally, the automatic placement of the groups and contacts was displeasing to some participants, who found the potential for contacts to change position each time the same set of groupings was viewed to be frustrating. The participant

quoted above, who valued seeing the structure of the Egocentric view, was also frustrated by the changing positions each time it was produced, '*I didn't like was that you could not move them to where you wanted them to be. It's nice that they're floating around, but for me I need one fixed structure I feel comfortable with.*' (Participant K).

7.3.2.5 Search Strategies

As was described previously, many participants had a desire to have the whole structure of their contact groupings visible at once when performing the tasks set, and this was considered to be a benefit in the Link-based and Gravity interfaces. As such, when using these two interfaces nearly all of the participants would engage in a strategy of identifying a group they were interested in examining from a fully zoomed out perspective, where all groups would be visible, and then zooming into the region of interest. This allowed the participant to identify the relevant individual contacts, which were smaller in size than the group nodes themselves, and as a result difficult to read from a fully zoomed out position.

'*It was generally down to if there was an obvious group, if they were part of a large group or they were linked between groups and I could spot them, then I would just zone in on that. Otherwise I would filter based on stuff, whatever was relevant to that person, that was likely to get as much stuff as possible. I used the pictures if they had pictures. I was really looking at the links and the larger groups. I kind of quite liked keeping it zoomed out so I could get a big picture.*' (Participant T).

Another aspect previously touched upon which this participant highlights, was the use of group size and structure (links) as being an important component of the navigational strategies used for locating contacts from this zoomed out perspective. The participant also goes on to speak about the filtering/restructuring feature, which some found aided them in their search strategies, and the use of the pictures of the contacts. The participant seems to suggest that for them the links and the groups size were primary components of their search strategy, the elements they considered first, and that restructuring and image scanning were to some extent secondary features for consideration in their strategy.

At some point during the set tasks, every participant used the restructuring buttons and all but two participants used the filter buttons at least once. Participants used the buttons to varying degrees, some finding them useful for performing the tasks, some considering them a last resort if a contact proved difficult to find, and others occasionally using them out of curiosity or their novelty. The restructuring buttons were used to a much greater degree than the filter buttons, which were sometimes ignored completely, and otherwise only used once or a few times as the need arose.

7.3.2.6 Restructuring Buttons

The restructuring buttons were broadly used to three degrees. In the first case this included participants who did not use restructuring in any significant way, but looked but looked at the different structures once or twice. Secondly, participants who once or a few times used restructuring when a task was very clearly easier to complete with a specific set of groupings. Lastly, the third and smallest group was of those participants who restructured across many tasks – sometimes more than once.

The participants who used restructuring the most would do so during the tasks, particularly the sharing tasks, either switching between several of the structures examining the groupings and the contacts or picking one specific structure they knew would be most suitable for the task at hand.

'*So depending on the task... if I have to choose the guests, I want to invite people to try Mexican, to have a taste. So I would obviously use the Location button, because it sends me straight away to the people I would be visualising. Whereas you say the exercise sending a letter to say goodbye to Nottingham, I would use relationship, because I want to let people know I'm going back to Mexico... So relationship in that sense is important. It's easy for me to use the configuration of relationship or location depending on the task, to discriminate the type of data I will be needing.*' (Participant N).

Selecting an appropriate structure seemed to provide two benefits to those participants: a means of visual filtering and access to groups more tailored to the task

at hand. By selecting a specific structure they wanted to use, participants could visually identify the contacts they were interested in quickly, and filter out the irrelevant contacts. This was often particularly true of the location-based grouping structure, where it was common for participants to only be interested in a specific geographic region when using it. The participants could therefore quickly focus in on the contacts they were interested in through one or more groups as they had confidence they were not missing anyone in other groups. And as the task was such that they were thinking in terms of location, the grouping was tailored to their purpose.

The strategy of switching between structures could be related to participants' use of features of the structure to navigate through the interface, where in one particular structure the participant would remember it contained some sort of obvious landmark, that made the contact(s) they were interested in quick to locate.

In contrast, those participants who did not use the restructuring as such, either saw no need to do so or cited the benefits to them of being very familiar with just one set of groupings, as working repeatedly with the same grouping resulted in them being able to remember the contacts in the groupings and the grouping structure more.

'*...it was useful to see who was in Nottingham, and that time the switching from one vision to the other vision was a bit overwhelming for me. So I just stick to one, so I knew exactly where it was. I like to be familiar with one grouping, so I can find something. I can see that it is sometimes useful to switch to location if you want some location based stuff going on, just Nottingham people invite, whatever. But I did not feel very confident doing that somehow.*' (Participant K).

This participant highlighted another issue in that the complete restructuring of groupings in the interface could be uncomfortable for participants who were attempting to learn a new interface and perform tasks with it. As well as the ease of knowing where the groups were, this also provided a greater sense of stability and control. This desire for stability and familiarity for some participants

was particularly demonstrated by one participant who having restructured their groupings during one of the first tasks set, then never changed back from it.

'*Again that wasn't a conscious decision. I think I clicked location or affiliation a couple of times at the beginning and that was working well for me, and therefore there was this lack of a need to broaden my horizons to the app further, having already found some that worked well that I could always begin to commit to memory, in terms of where I thought people were going to be on them*'. (Participant A).

This was interesting in that it demonstrated, for this participant at least, that using the free groupings that the participant created on the basis that made most sense to them for grouping, was less important as a factor than having a set of groupings the layout of which was familiar and committed to memory.

7.3.2.7 Filtering Buttons

The contact filters were rarely used by participants and for most they were considered primarily as a 'last resort' option. They would sometimes use it if locating one or more contacts was proving difficult or was not quickly apparent. This was also true of restructuring for some participants. Restructuring or filtering provided a means to instigate change that might reveal, narrow down, or aid navigation to where the contact might be, '*I couldn't find Allan where I expected him to be, so I changed it. I think what triggered me to use those buttons was where I couldn't find something and my expectations weren't realised. If I can find something straight away I don't need to use that*'. (Participant D). As this participant points out there was often no need to employ these elaborate re-orderings and filtrations, simply visually navigating the groups and contacts sufficed.

Social awareness and the visual filtering that participants did of their own accord were other reasons that participants were reluctant to filter. Even though participants knew they might filter down the results via a relevant category of group, there was a sense that sometimes people not directly relevant can be important to be aware of, '*something that may only be mainly important to PhD students may also happen to be interesting to another party*' (Participant K).

When all of their groups and contacts were present on screen, participants were happier that all of the potentially interested parties could be contacted when sharing data, as it was easier for the participant to maintain a broad social awareness. As previously described, participants seemed to be happy to visually filter out groups and contacts that were of no interest to them through knowledge of what the groups represented, and focus in on contacts and groups that were of concern to them.

7.3.2.8 Importance Filter

The only aspect of the interface discussed where participants were both unequivocal and approaching unanimity was in the ambivalence and apathy expressed towards the importance filter. The primary reasons given for this dislike were that using numbers held no meaning for them, that they forgot that the bar was there, and that they didn't remember how they had rated people or importance was not necessarily static in terms of the tasks they were completing. '*I was thinking about what I'd rated people and it all felt rather arbitrary. I mean they might be important, but in different contexts, so I just didn't touch it.*' (Participant C).

The numerical rating scale was considered by nearly all participants to be arbitrary in this way. Using numbers resulted in ratings not holding any meaning that the participant could apply to the tasks they were performing. It meant that participants were unable to remember how they should expect people to be rated and as well as forget that the filter was even there, as the numbers did not prompt participants with a strategy to locate contacts. Additionally, there was an element of poor interface design in participants' lack of use of this feature. It was consistently mentioned that a possible reason for its lack of presence in the participants' minds was its position at the very bottom of the screen.

As a concept, the importance of contacts did seem to be present. However, participants spoke more about importance in terms of their groupings than they did individuals. Individuals were primarily discussed in terms of people who were of particular importance - favourites/VIP contacts whom they wanted to highlight or group. Groups were seen as holding their own inherent importance,

which therefore already brought order to the importance of the contacts. While discussing restructuring groups one participant elaborated on the importance slider's relevance to the structure.

> *I think possibly the naming of the categories. They're four relatively distinct categories that distinguish between geographical location and relationship, which again almost replaced the bottom bar, because there the level to which someone appears important to me appears as a cluster, and I can select what you might call a ten out of ten important cluster, rather than using the bars.* (Participant A).

The participants' comments on this perception of inherent group importance reinforces the group standard deviation data (see Table 7.5) of importance ratings within the groups. This data showed a tendency for similarly important contacts to be grouped together, with many groups having fairly low standard deviations, demonstrating that participants, whether consciously or not, group contacts according to a similar perceived importance. The comments of the participant here would suggest that value judgements are made as a strategy by participants at a group level as well as an individual level, allowing many contacts to be considered within a task in one go.

7.3.2.9 Sharing Tasks

Overall, participants thought that the person-centric interface most aided them in performing information sharing tasks, and was less useful when to the information re-finding tasks. None of the participants explicitly expressed a particular dislike for performing the sharing tasks through the interface, with most participants suggesting it was either adequate or a useful approach. This more positive experience was due to the awareness the person-centric interface provided them of their contacts when sharing, giving them confidence all relevant people were included, and the ability to quickly share with multiple contacts.

> *Sharing was good. If I was sharing something I'd have to think about who I was sharing it to, and write a list or... You'd think, "Oh, probably forgot someone." And you'd write "If I've missed someone out..." Whereas with*

> *this, if I was confident that my groups were comprehensive, and then I looked at the group and selected from that, I'd be pretty confident not to have to say, "Have I missed someone out?* (Participant D).

When performing the sharing tasks a common strategy was that once a relevant group had been identified, whose members would all want to receive the information, the participant would select the group as a whole to share the information with, rather than examining and adding each individual contact. Completing the task in this way was more efficient for participants: 'I certainly liked to invite whole clusters where possible, for example, a party would require the whole of the PhD office to be there. Or leaving Nottingham to click on clusters that were exclusively of Nottingham people. That was definitely one strategy'. (Participant A). However, the same participant went on to point out that this strategy was dependent upon and varied with their confidence in their groupings.

> *I did notice a couple of times the extent to which I very manually liked to go through each cluster, by looking at each individual person, up and down, which is interesting because it demonstrates the extent to which I didn't entirely trust the software to do it for me... where I clicked on a given cluster, for example, PhD colleague, and then when it got to Nottingham, knowing full well that that person will be in both, I still went through as I was doing it to just check they were still there.* (Participant A).

How participants chose to perform the sharing tasks when they involved multiple people was closely linked with the level of confidence participants had in their groups. When they believed that there was no need to examine the contacts in the group, because they were confident there would be nobody in the group who should not receive the information, then these group-level interactions were a benefit. When participants held some level of doubt, even if it was small, then the need to check each of the individual members of a group took over, and the efficiency of performing actions at a group level was lost. This was another reason why those people who were dissatisfied with the groupings were unhappy with how they had constructed their groupings. When performing these sharing tasks their confidence was much lower, and there was a greater need to check all the individual contacts of the groups.

7.3.2.10 Re-finding Tasks

Impressions of performing the re-finding tasks were split. Participants gave opinions on re-finding in two parts. In one they discussed the act of finding individual contacts in the interface, and in the second they talked about the presentation of the related information.

Contact Location

Participants tended to be quite critical of the process of locating individual contacts, as they believed that the process of looking at the groups and then the contacts was less efficient than it would be with an email client, where the name of the contact could be searched to retrieve a range of related information - as was also common in the diary study results in Chapter 4. '*If I'm honest, I prefer an email client, because then I can search, so I can search your name and if I was looking for a specific email like you, I could search "Paul"... and that would filter it down more effectively than this where I'm looking at related information, trying to work out where things are.*' (Participant T). While the main criticisms related to locating a person being more cumbersome, a variety of complaints also addressed shortcomings relating to features not implemented that were present in their email clients. In particular the lack of filtering along other dimensions other than person was considered a significant drawback.

Several participants did offer a contrary opinion of the efficiency of locating contacts. They believed that they could much more efficiently locate contacts by going straight to the grouping they knew the contact would be in:

'*I guess it kind of reduces the content of one's inbox down to just the person of interest in a much quicker way than typing their name in a way an inbox search engine would. Normally you'd have to type the beginning of their email. The conference organiser, for example, before you could get that up. Whereas here I could just click on the person, I knew his name, and I knew what cluster or clusters he belonged to, so I think it was very time efficient.*' (Participant A).

As this participant implies, this efficiency is dependent upon having knowledge or familiarity with where the contact would be located. Again returning to the idea

that interface is dependent upon having confidence the groupings match their beliefs of where things should be.

Association of Information Through a Contact

Having identified the desired contact, participants then clicked on it to reveal all of that contact's related information. Participants were asked whether they thought associating information with people in this manner was a useful approach. Two participants considered this not to be a useful approach, another two considered it to be 'about the same' in terms of how they searched for people to access their information. The remaining majority considered it beneficial in some way to access information. As was intended, the design and use of this aspect of the interface by the participants mimicked the way that people thought about information, as identified in the diary study in Chapter 4. Rather than information being accessed in terms of messages, it was instead accessed in terms of contacts, people.

'*I thought it was good. Often I'll find myself searching for an email, and it was sent from someone. That would be a really quick way of doing it. I'll just be, "Oh, I know Gary sent me that email." And it would be a really quick way of showing everything he'd been involved with.*' (Participant D).

Similar to results in the diary study, in this instance the participant discusses the process of re-finding the desired information is in terms of their memory of a person. In particular, participants tended to be most enthusiastic about the way in which the interface design exposed both the attachments and the links associated with the contacts, and presented them separate to the messages they were sent in. Reasons for thinking this was a particularly beneficial feature related to instances where they communicate a lot with a particular person, and messages and attachments therefore become increasingly difficult to locate, '*if you can't remember which email you were sent a link in, but you know it's somewhere in the conversation, then that can be quite annoying.*' (Participant B).

Another aspect of this approach, that one participant pointed out, was that by bringing together all of the communications with a person in one place, it made

the participant aware of how they were interacting with that person, how they spoke, and what they spoke about:

'*I guess it would make things a bit easier. It makes you aware of what you are talking about with one particular person. So if you say you think you are really casual with someone or you only talk about work.'*
'How did you feel about that?'
'The truth comes out!
(Participant H)

The participant's comments underline how the greater social awareness that this aspect of the interface can provide could potentially be uncomfortable in some circumstances for people.

7.4 Discussion

7.4.1 Group Structure

MacLean et al. (2011) point out that when listing contacts people would forget people they knew, as well as groups of people, suggesting that people are not particularly good at remembering the structure of our own social networks. Similarly, the way in which participants would become stuck when grouping contacts and make decisions to resolve the impasse they had come to, whether it was through creating new groups or by revising the meaning of groups, is indicative of participants having quite a strong sense of how contacts may be grouped. However, this process was not necessarily entirely an expression of their mental model of their contacts, as the outcome of the process was often modified by factors such as what was a more convenient action to take, although it did have to adhere to the participant's own internal logic of how their contacts related to groups. Participants would often speak in terms of emotions about how contacts fitted into groups, giving the impression of the participants having only a sense of what was acceptable to them, rather than an attempt at creating a representation of a pre-existing model.

However, the construction of groups itself did seem to represent the build-up of a

model. Each group created would have a set of internal rules associated with it, governing who could be included within it. As participants grouped their contacts they felt a need to balance a number of factors, including the number of groups in terms of the amount of effort required to make them and then remember them all, the amount of contacts in each group in terms of not having too few or too many contacts so as to render the group meaningless, the strength of association with a group, the importance of the contacts, and the nature of contact with members of a group. Moments of hesitation in grouping were often points where these factors came into conflict with the existing group structure, necessitating a decision as to how best group a contact without violating their sense of what was acceptable.

The process of contact grouping was highly influenced by how participants balanced the competing tensions of a desire for expediency in the task of grouping and the desire for groups to reflect correctly how they perceived their contacts. This manifested in how much effort participants were prepared to put into modelling the intricacies of their relationships to the contacts versus filing them and how tempted they were to revise their rules for the groups - how much detail would they sacrifice - to include more contacts versus creating new groups. This had the impact that the more enthusiastic participants were, the greater the number of groups they created and to some extent the greater the amount that the groups represented their real social relations. This is possibly similar to the effect that Malone (1983) found in office file management, with people either adopting the approach of being 'filers' – those who spend time organising information – or 'pilers' – those who pile information, allowing clutter and dedicating little time to organising information. This behaviour was also evident in digital contexts, Whittaker and Sidner (1996) identified a variation of this dichotomy within email organisation, and Abrams et al. (1998) within web bookmark organisation. In the context of this study, 'pilers' were those participants who tended to lump lots of people into broad groups regardless of social structures or nuance, and 'filers' those who were motivated to model fully their social interactions. However, if it is indeed a similar effect, the distinction is less clear as all participants had to 'file' to some extent due to the nature of the task set being one in which they

were asked to order their contacts.

Importantly, this study found that the design of the egocentric interface affected these behaviours, increasing the amount of groups participants were willing to create and encouraging a greater degree of modelling of their social relations. The result implies that it may be possible to influence the structure of how people construct groups, though not necessarily the types of groups created, through interface design. Encouraging these more detailed groups could be beneficial in that large ill-defined groups were considered less useful by participants. Further research may find other methods through which other aspects of group construction may be influenced. Understanding how to influence the groups people create through design could conceivably be beneficial where systems are most effective with groups meeting certain criteria. For example, in this case encouraging more specific groups was sometimes seen as more useful for performing tasks, as they were better defined and therefore it would be clear from just the group name who would be contained within it. Kelley et al. (2011) also examined more than one system of manually grouping contacts. While they also found that this affected how contacts were grouped, unlike this study they found that people tended not to use sub-groups when this was available to them as an option, though perhaps crucially, in contrast with this study, it was not explained that they could do so.

McCarty et al. (2007) describe people as labelling social groups in an ad hoc manner, with the best available label they have at the time. This was often true of the strategies participants used for creating and identifying groups as they went, dependant on factors such as what group could most obviously be created next based on the remaining contacts. Contacts were commonly grouped according to their moment of strongest or most important association with the participant, even if this association was no longer true, never was true, or had not yet occurred. For example, the context in which a person was first met, e.g. a particular city, was commonly used to describe a group of contacts even if they no longer lived there, as it described the shared context of the group. The choice between grouping according to a contact's present circumstances and a stronger association would sometimes give pause to participants when grouping,

as they decided how to best resolve these contradictions. The use of strongest association as a determinant of how to group contacts could potentially be used as to determine how contacts should be grouped when there is contradictory information available.

Where people first met was important in both this study and the previous contact grouping study in Chapter 5 for how groups of contacts are recalled. In the second study this behaviour was identified as a distinct category of group description, and in the present study this category of group was again observed. This behaviour can be linked to the strength of association determining how contacts are grouped. Indeed, where people first met emerged as being particularly important in determining group identities, as was who they were met through, albeit less frequently.

The interfaces used here focused on representation of social groups and not representation of how people remember contacts in relation to each other or the relationships between contacts. Mapping social relationships at a more individual level was sometimes important to participants, particularly when they wanted to indicate people at the periphery of a group, people known through other people, e.g. partners of friends. This was also highlighted in the diagrams participants drew in Chapter 5 of how they would represent their contacts as group. In some of the diagrams produced in Chapter 5, participants made clear how they viewed the trajectory of contacts. In some cases, this was moving between groups, e.g. colleague to friend, or moving within groups, e.g. were they core members or at the periphery of the group, were they moving towards the core of the group or drifting away from it. Selectively mapping some of these relevant contact-to-contact relationships might provide a way to identify the extent of group membership and provide a means to alleviate the conflict participants had when grouping between contacts who were only peripheral members of a group, neither belonging in it nor outside of it.

The three findings of contacts being grouped by their strongest association, that they were created using internal rules, and that they tended to be grouped according to the perceived importance of the contacts has possible implications for the approach of using automatic detection to identify contact groups based

upon social network structure (e.g. Jones and O'Neill, 2010). Their research would suggest that while identifying groups relating to the structure of the social network would potentially correctly identify a proportion of groups participants perceived, a large proportion of potentially useful groups, which do not strictly correspond to social networks, would go unidentified. Jones and O'Neill (2010) demonstrated that there was greater correlation between manually created groups and groups derived from automatically generated network graphs where smaller subgroups participants perceived were combined into a larger network. Many of the factors which define groups of people are multi-faceted and subjective. For interface design, these more subjective and ad hoc decisions may be more useful, desirable, and dependant on the context of the group's use, than representing a completely accurate social network.

The results did suggest that these amorphous groups – those not bound to a strict social structure – would often correspond to specific potentially identifiable and usable attributes, such as roles and locations. But in either case it would be difficult to identify fully the make-up of these groups by different means of data, due to participants being happy to group contacts 'inaccurately' where associations with some contacts no longer represented current circumstances. The perception of contact importance seemed to be suggestive of another useful factor to consider in automatic generation of groups, but this would require input on those perceptions, which as demonstrated here made participants uncomfortable.

That at times, and to varying degrees, participants were dissatisfied with their groups, that nearly all of the participants engaged in revision of the rules for inclusion in one or more groups, and that they often lacked confidence in the groupings they had created brings into question the idea of the 'ideal' group proposed here, and also how you measure accuracy of automatic detection when the user's own picture of their contacts is uncertain and ill-defined in certain aspects.

In the second study of this thesis (Chapter 5) it was reported that the average number of groups in the diagrams participants created to represent their contacts had a mean of 6.61 and a range of 3-14 groups. As a comparison, in this study the average number of groups in the Ego and List conditions respectively the mean

number of groups was 11 and 7.61 with a range of 4-23 and 3-16. The results for the List condition are very similar to those from those of the diagrams in the second study of this thesis. The Ego results, however, are a bit higher but still broadly similar. In other similar studies of contact grouping, discussed in greater detail in Chapters 2 and 5, Kelley et al.'s (2011) privacy contact sorting study found a range of groups from 2-30 and a mean of 6.5 groups, Jones and O'Neill's (2010) privacy contact sorting study found a similar range of 4-20 groups, but a higher mean number of 16.7 groups, and Leskovec and Mcauley (2012) found recorded the highest mean of 19 groups. The results from these studies are not wildly divergent from the results found in the two studies described in this thesis, but reasons for the several higher results are worth considering. One factor that may have influenced why numbers are to the lower end in this thesis' studies might be attributed to the number of contacts involved. In this study participant's contacts were restricted to only those they reciprocated contact with over the last 6 months, resulting in fewer contacts to group compared to other studies which used all of a participants contacts. Equally, as is highlighted in the results here, the design of the Ego interface condition significantly increased the number of groups created by the participants. The mechanism by which participants perform these tasks, therefore, is another factor that lead to variance in the number of groups created.

7.4.2 Contact Awareness

Awareness was an important factor contributing to the split between those preferring the more traditional List presentation and the two link-based presentations. Those who disliked the List interface thought that the awareness provided by other interfaces was important for two primary reasons; they valued seeing how groupings and contacts related to the wider structure, and a desire to be visually supported in remembering the existence of all of their groups and contacts. As Card (1996) describes, the benefits of visualisation, in this case by representing structure through links and groups, accrue when visualisations are used as workspaces, as they reduce the cost of performing a task by shifting the burden from slower cognitive to faster perceptual processes, by enabling more informa-

tion to be conveyed through the structure of the visualisation. Risden et al. (2000) performed a similar comparison between a graph visualisation and a tree list of web data, where they argued that the ability to see the global relationships through the structure of the graph visualisation gave participants an advantage in performing certain retrieval tasks. Although, as in this study they similarly found no preference between the two forms of interface among participants.

Participants found the greater awareness provided by the Group Link and Egocentric useful as it served as a prompt for noticing and remembering people with regards to the sharing and inviting information tasks. This was particularly beneficial in ensuring a greater chance of including relevant people who might otherwise be forgotten or ensuring no one is accidentally forgotten, when completing these kinds of tasks. Again corresponding with MacLean et al.'s (2011) observation that people are poor at recalling their own social structures. However, one participant offered a - perhaps cynical - counter view to the benefits of this when discussing contact awareness, '*If I forget someone, maybe it's not that important,*' (Participant K) implying that if they did not remember the contact then maybe it was not important to include that person anyway.

Another behaviour that is seen in information management when grouping the contacts was the use of groups as reminders of tasks they needed to perform. Barreau and Nardi (1995) describe how files would be organised by people such that, when these groups were seen again, they would serve as reminders of something they needed to do. Similarly, several participants created groups with names such as 'Keep In Touch', or 'To Be In Touch'. These groups seemed to reflect anxieties about the state of certain relationships and possibly a desire to be encouraged or reminded to stay in contact with certain people. This behaviour, however, was not very common occurring only in the data of a few participants, but stands out among the data and is notable along with the other groups created that deal with emotional states relating to contacts rather than descriptions of them. One of the benefits Whittaker et al. (2004) identified in their evaluation of ContactMap was that it promoted 'social reminding', whereby simply the visual presence of contacts aided remembering to perform communication and other social obligations to those people.

7.4.3 Group Confidence

Manual grouping was used in the first part of this study to generate 'ideal' contact groupings, as opposed to generating them automatically. Automatically identifying groups can be challenging in that the accuracy of the groups can vary dramatically (Jones and O'Neill, 2010) and is unlikely to predict perfectly what users would perceive to be accurate. By having participants create the groups according to how they perceived them to be, it was hoped that this would provide optimal groupings that participants would be able to use due to understanding of the group's membership being a minimal factor in the task completion.

The consistent rules participants tried to maintain when constructing groups appeared to encourage optimal or high confidence groupings at first. They not only provided a framework to aid participants in grouping contacts, but it also meant that members of that group could all be treated in the same way when using the interface at a later date. A group of close friends that has been constructed consistently to meet certain rules of intimacy can then be used at a later date to include or exclude all members in communications without the need to carefully consider each of the members if the creator of that group is confident in how they have constructed a group.

This, however, was commonly not the case, and group confidence emerged as an important issue for participants when using their groupings despite tailoring them to their personal perspective. While some participants did have absolute confidence in how they had created their groupings and could use the groups to perform tasks at a group level, others harboured some degree of doubt. This doubt undermined the benefits of being able to perform tasks with entire groups without examining the contacts. The doubt, even when slight, meant that participants would scan each individual contact in a group to make certain of its content, even when the participant felt they knew who would be in the group. This seemed to be the same problem that one participant had with the idea of automatic detection of groups '*If it was accurate. If it was 98% I might use it. That might be worse because I would trust it to be accurate*' (Participant T). The small amount of doubt about how they had grouped the contacts often undermined their use of the interface for tasks as a whole.

7.4.4 Group Importance

The importance rating participants gave to each of the contacts showed that contacts who were considered to carry similar levels of importance were likely to be included in groups with each other. This was demonstrated through a low average standard deviation of the ratings in each group (see 7.5), implying that groupings on the whole tended to be rated at the same or approximately the same level of importance in terms of the participants' contact with them.

From this it might also be inferred that these groups carry a weighting of importance commensurate with the importance of the contacts grouped within them. This might be expected to occur as people included in a group together typically share something in common, and might therefore share a common perception as contacts by the participant. This finding was also reflected in the comments of the participants who were disparaging about the importance filter feature of the interface. Some participants commented that this feature was not useful because they felt groupings they had created already inherently embodied a level of importance that was implicit to the grouping they had created. Importance for them was therefore reflected in the decisions of which groupings were relevant.

It was also evident that these judgements of importance when made explicit created discomfort. While the relative importance of contacts did seem to be a factor in how they were grouped or represented, as identified in the free drawn diagrams in Chapter 5, interactions that encouraged participants to explicitly address this difference in importance (such as rating, and sometimes grouping based on relationship), made participants uncomfortable. Participants felt guilty passing judgement on people they knew and often liked, and found differentiating between different types of importance (practical versus personal importance) difficult. Adamic et al. (2011) similarly found in examination of friendship and trust on the Couch Surfing website that people reacted unfavourably to using its reductive rating scale of friendship in order to leave feedback about users. As in this study the subjective and multifaceted nature of friendship resulted in its use meaning different things to different people, and that it was difficult for people to quantify friendships along a single unidimensional scalar. The implication of this for design is that utilising elements of contact performance in must be done

either by providing sufficient complexity to describe a relationship, or through more subtle and indirect or selectively means, e.g. choosing favourite contacts. Fisher et al. (2007) in their SNARF interface present a more subtle and nuanced approach to this by highlighting and grouping contacts based upon the users behaviour with them, making the groupings incidental and taking the judgements out of the participant's hands. And indeed their approach of mimicking how people triage incoming emails was considered useful by the participants.

The strategies through which participants determined what their first groups were going to be are also potentially significant to as they reveal that not only do the groups have a description, a criteria that defines them, but they are also thought of in terms of their importance (in this instance, in terms of their relationship rather than as contacts) to the participant and the amount of contact they have with them. These attributes seemed to be determined by how foremost they were in the participants' minds.

7.4.5 Information Tasks

That participants tended towards sticking to a familiar pattern once they had established a pattern for locating groups and contacts, which was similar to behaviours and comments in the diary study and wider research (e.g. Jones et al., 2004), where people are disinclined to change set behaviours, even if they know they are or might be inefficient, because for them it feels most comfortable and is the path of least resistance, '*even if that will take longer and I can actually save some time, which is not a great thing to admit to, but it's just an automatic reflex to do rather than strategize.*' (Participant B).

7.4.6 Person-Centric Interface Design

Preferences for the two broad forms of interface – list-based and link-based – in both the grouping and information task parts of the study were largely determined either by a desire to have systematic and ordered processing and/or searching of the groups, or by a desire to have a clear visual overview of the groups created and their structure. Although participants expressed strong pref-

erences for one interface or the other, this preference might suggest that these are two important design goals for manual grouping interfaces, rather than different preferences for how to group; neither requirement is necessarily mutually exclusive and both could be achieved within a single interface. One participant envisioned one such alteration where grids of contacts are present, within a linked format that visualises group structure.

Participants often seemed to use their memory of certain aspects of the structure, size of the groups, and contact images to aid them in navigating to a contact or a group. However, to a degree participants were relying on learning the location of groups and contacts to speed up their navigation. For some, this familiarity was the most important navigation strategy.

The process that was identified in the diary study, in Chapter 4, of recalling a person along with the information being re-found was also reflected in this study. The prototype appeared to mimic successfully the processes identified in the first study, with participants explaining their re-finding process in the same terms as in Chapter 4 of recalling certain attributes to recall and use to navigate, and referring to how the interface reflected their process for re-finding information. As a result a majority of participants found this specific aspect of the interface to be positive or beneficial in some form.

In particular, the separation of links and attachments from the messages was seen as useful. While several participants missed the context of the original messages, this again seemed to reflect the same association of information with people. It was the information that was important to participants, and that they associated with the contact, and not the messages that contained the information. Participants talked about wanting the following items to be separate and accessible in this manner: phone numbers, addresses, links, and attachments.

While it was only described by one participant, that the interface provided greater awareness of the participant's communication behaviour was an interesting point that may merit further investigation. This effect is similar to what the interfaces that raised awareness of social behaviour described in the literature review in Chapter 2 were attempting to achieve by revealing the wider context of our com-

munication behaviours. It would be useful to determine whether greater awareness of communication behaviour, if this is indeed a consequence of the design, is a positive or negative attribute. It is conceivable that in some circumstances this could make people feel uncomfortable, or more positively it might encourage people to change their behaviour having assessed how they are behaving with others.

7.4.7 Limitations & Methodological Issues

One issue with the methodology was that it arguably introduced too many unfamiliar design concepts to the participants at once. While they were shown all of the individual features of the interface and given time to try them all out, it was common for participants to say that they forgot certain features were available to them for large parts of the tasks, or in the case of the importance bar, often all of the task. Features such as group filtering, structure re-ordering, and importance filtering were largely considered optional and not often perceived as necessary to complete a task and so might not be considered for use even when they might have aided the participant. Although equally, some participants, as they tried out features and decided they had utility, this caused them to increase their use of certain features.

Differences in approaches to grouping and preferences for different interfaces may be down to differences in how they visualised the purpose of the interface or the contexts in which they would be used. The need to group all of the contacts and desire to minimize the amount of groups created meant that some of the participants either made groups that were too broad or revised their internal rules for inclusion in certain groups. This had the effect that some of the groups became much less useful to them when performing information tasks. Although the methods used here provided valuable insights into contact grouping, given that the intention of using manual grouping in this study was partly to achieve more ideal or accurate groupings of contacts for participants to use, the results suggest that the idea of an accurate grouping is a misleading ideal that this method cannot produce.

When contrasting the results from the repertory grid, in Chapter 5, and the groups created in this study, through the contact grouping process, it is clear that the results of the contact grouping hid to some extent the language that was prevalent in the second study that was important for developing many of the sub-categorisations of the grouping framework. For instance, it was common in the repertory grid results for descriptions of Location groups to occur in multiple forms with an emphasis on different forms of association, such where people were met, where they live or lived and where they are from. This was not always evident in the larger groupings of the prototype, where Locations were largely described by one word geographic descriptors e.g. Nottingham. When participants were grouping individual contacts and thinking aloud to explain their decision making process, however, the language seen in the results of the repertory grid was again evident: *'I'll put this person in Retford because I met him in Retford, even though they're living in Sheffield,'* (Participant I).

Taking into consideration these comments, the similarities in the results of the group coding between the second and third study bears out the use of repertory grid as a methodology to elicit how people tend to group their contacts. While most of the group types that emerged from the repertory grid were also found to be present in the contact grouping exercise, some were less prevalent or not found in this instance. As these had a low occurrence in the second study it is hard to determine whether they were not present in the results due to chance or because they were not relevant to the way people practically group contacts. In examining the results of the repertory grid it is clear that it elicits much more varied and specific forms of groups. This may be that it allows a much finer grained expression of relationships or that only certain forms of these specific sub-groups will commonly emerge within practical applications which may have a tendency towards broader groupings.

There were at times occasions where participants had become obviously bored due to the length and repetitiveness of the task at hand. This had the effect that for some of the results, participants were creating many fewer groups that were broad in nature and rushing through contacts, compared to their previous behaviour in earlier conditions. While the balanced ordering of the conditions was

intended to mitigate the potential impact of this on the results, it was a significant factor affecting the results in certain conditions for several of the participants. This also has implications not only for the methodology of this study, but also the design of grouping systems generally. While it was not thought that this work would demonstrate manual grouping to be a more plausible solution to group creation, it does confirm that manual grouping is too repetitive and time intensive solution to maintain interest for many people. Nor is this promising for the issue of maintaining groups as contacts emerge, disappear, and move between groups, as well as the question of there being a degree of discomfort for some people in the implicit judgements about how they group certain contacts.

7.5 Conclusion

This chapter sought to evaluate the 'person-centric' design ideas identified by studies in Chapters 4 and 5 concerning performing information re-finding tasks, and representation and relationships with personal contacts respectively. In doing so it aimed to address the question of what the decision making processes involved in contact group construction are, how contacts and contact groups should be represented in order to support information sharing and re-finding tasks, and what benefits this approach might provide for performing information sharing and re-finding tasks.

7.5.1 Group Construction

With regards to group construction, the results from this study make it clear that there are a number of common interrelated behaviours influencing the creation, structure, and description of contact groups. As in Chapter 5 the resultant groups in this study were again defined in terms of attributes describing a specific subset of the contacts, which were dominated by a focus on the overarching categories of Location, Relationship, and Organisation. The multi-dimensional nature of these groups was again also evident, with a significant proportion of groups being made up of combinations of these, and other, attributes in order to accurately describe the groups.

The process of group creation and the ongoing act of adding contacts to groups that led to their definition, however, was to some extent a form of reification that created challenges. The perceptions participants had of contacts contained a variety of abstract ideas that appeared to be difficult to reduce and make concrete, at least within the confines of the system, or to some extent involved tacit knowledge that was difficult to make explicit. The ability of participants to model their relationships according how they wanted to define groups was constrained by factors such as a desire to not to create so many groups that maintaining and recalling the groups became difficult, a desire to represent strong relationships between contacts, and the need to group all contacts. Group definition was also in tension with the series of goals participants had in creating the groups, that they should accurately reflect the internal rules they set for the groups, appropriate for the tasks they'd be used for, that they could have confidence in what people would be included in a group without checking, and that the group remained meaningful to them.

In order to create groups within the goals and constraints they imposed on their creation, the participants used a range of strategies to create a particular view of the complexity of their social relations. Factors that influenced the outcome of the groups produced by participants included:

- The strategies people grouped by – creating all the groups at once versus as they went – and the strategies they used to identify groups - contact amount, importance, location.
- Whether the groups created by the participant reflected stable social structures, or were more amorphous in nature being based around a particular shared attribute associated with its contacts, such as living a particular city, for example.
- The point of strongest association of a contact with a group.
- Judgements of contact amount and relative importance of the contacts in terms of their communications.

By examining how participants define groups, the strategies they use to identify

groups, as well as the goals and constraints people place on groups this research has identified factors that are important for the decision-making process involved in contact grouping. Identification of these factors provides greater insight into what constitutes a meaningful contact grouping to a person, and what factors will lead them to determining what groups should be created. This can in turn be reflected in requirements for supporting users to either manually creating groups of contacts, or providing criteria for automated detection of contact groupings, that can reflect to a greater extent the goals and constraints a person would place on group creation;.

7.5.2 Group Representation and Supporting Information Tasks

While the research conducted in this study revealed useful findings with regards to people's preferences about specific aspects of how groups should be represented and the goals that the interface should seek to address, it did not manage to identify a specific interface design out of those evaluated that embodied most of these facets and that was clearly a preferred approach for performing information sharing and re-finding tasks. Preferences for the different interface designs explored in this research were relatively evenly split among participants. It did, however, identify that how groups are represented can have an impact on people's behaviour in performing grouping tasks. When the design of the interface supported sub-groups and visual feedback of an inter-connected group structure in the egocentric grouping interface – compared with the list-based interface – this influenced the degree to which participants modelled their social relations resulting in more groups.

Representing the structure of groups and contacts was also found to be valuable for creating greater awareness of the user's contacts. There was a split among participants between those who preferred a more traditional list presentation of their contacts, and the two link-based presentations. In the case of the link-based interfaces, the awareness provided was important because participants valued seeing how groupings and contacts related to the wider structure, and because seeing the whole structure reminded them of the existence of all of their groups and con-

tacts. This supported participants in both navigation around the interface, and helped them to include contacts in information sharing tasks who they might otherwise not have remembered or considered. In contrast the list style interface provided less overview of the contacts and groups, but was much easier to scan and process contacts as part of those tasks. This highlighted the need for the 'person-centric' interface design to find a bridge between these two goals of both representing the structure and providing overview to support the users' memory through visualisation, and an easy to scan order that supports the user in efficiently performing actions with the interface.

Conducting the information tasks with the prototype's interfaces demonstrated a mixture of positive and negative reactions to the design ideas implemented to support the tasks being performed with all of the different contact representations. This included aspects such as filtering and ordering contacts according to specific criteria, filtering according to importance, and extraction and ordering of different types of information associated with contacts, such as links and files. In each case the underlying concept identified for inclusion in Chapters 4 & 5 were considered relevant to these person-focused interfaces, even if they were not useful in this instance or the specific implementation was not as participants would desire.

While it was previously identified in Chapter 5 that contacts and groups are perceived as having different relative levels of importance that is implicit to the user, the implementation of that in this interface was very negatively received as a means of ordering or filtering contacts. The system in this case required explicit numerical judgements to be made about contacts that made participants uncomfortable and which did not map well in its implementation to how participants viewed contact importance when performing the tasks. While this was a very negative result, it did highlight the need for the use of this in interfaces to be more oblique. Value judgements on contacts are important in performing information tasks (Fisher and Dourish, 2004), but their use should not force the user into making explicit judgements, rather other means such as analysis of the users past behaviour might allow importance to be predicted and used to support a task.

In contrast, the implementation of the premise of associating individual informa-

tion items with people, as identified in Chapter 4, was successful in demonstrating that participants were happy to access information by thinking in terms of people to navigate to it. Additionally making files, and URLs easily accessible, by making them available when selecting contacts, was also positively received and showed that participants were largely happy for this to be done without the surrounding context of the messages they were sent in (as long as the context was still retrievable).

This evaluation has examined a variety of designs for representing contacts within an interface, in order to support information tasks. From this a range of design considerations for how 'person-centric' interfaces can support re-finding and sharing tasks have been identified, such as the goals and constraints that influence what make meaningful groups for performing information tasks, and the value of representing structure between contacts for performing information re-finding and sharing tasks. While this evaluation has demonstrated that people were happy to utilise the associations they recalled between information and people as a means of accessing information, and therefore that exploiting this to a greater extent has the potential to help support people in performing information tasks, it has less clearly demonstrated the value of making people the central focus of interfaces for managing information. In future research it will be important to design a 'person-centric' interface that takes into consideration the findings from this evaluation and that is evaluated over a longer period of time (similar to Sawyer et al. (2012)) as part of the user's everyday information behaviours. This will enable the question of how 'person-centric' interfaces can be designed to integrate with people's existing information practices to be addressed.

Chapter 8

Discussion

At the start of this thesis, the diversification of information forms and the ever increasing amount of information people encounter (Carpenter, 2009; Lenhart et al., 2010) was contrasted with the relative lack of changes in the tools used to process information with over the years (Kljun et al., 2013) and the need to manage all of this information (Jones and Anderson, 2011). As a result of this it was suggested that there may be increasing challenges in performing tasks with our personal information. With this in mind, the thesis began by focusing on a specific information task – that of re-finding information – with the intention of identifying where difficulties might lie in performing these information tasks, and how people could be supported further in performing them. The research identified that while for the most part information re-finding tasks are considered to be easy to complete, difficulties still occur in 10% of tasks and completion of tasks is not necessarily straightforward even if perceived as being easy. Examination of the re-finding process identified the importance of contextual memories for enabling information to be located. As part of this process, keywords, time, and people linked to the information were identified as key elements of re-finding strategies once the participant had reached the 'source' of the information. Being able to recall from memory and then utilise these elements was important for enabling information re-finding tasks to be performed successfully and efficiently.

The literature review revealed that previous research had focussed much more

on using time and keywords to perform personal information tasks and had been researched to a much greater extent than utilising the links made between people and information had. While there have been valuable contributions to this such as the *ContactMap* (Whittaker et al., 2004) and *SNARF* Fisher et al. (2007) systems, and Sawyer et al.'s (2012) system for recording physical social interactions for re-finding, the overall body of work was shown to be quite small. It was clear from the diary entries and literature that, unlike the other aspects people remembered when re-finding their information, links between people and information were often unable to be used as part of the process of locating the information by the user – in part due to the difficulty of capturing social interactions (Sawyer et al., 2012). As a result there has been a lack of literature on the design and utility of person-centric interfaces, and a limited understanding of how a 'person-centric' approach might be used to support information tasks and what form they should take to support those tasks. More effectively utilising the links between people and information was therefore identified as a means of reducing the number of difficult or frustrating information tasks, by providing a further or improved means for interacting with and accessing personal information.

This was explored in the literature review, which identified information re-finding and information sharing tasks as two areas where this association might be beneficial. The research conducted in this thesis aimed to identify how 'person-centric' interfaces could be used to support information re-finding and sharing tasks by exploring how people wish to organise and interact with information via representations of people they associated with it.

This chapter discusses how the research conducted in this thesis is situated in relation to the literature, and synthesises the findings of the studies to identify factors that influence contact group construction, and to propose high level design guidelines for 'person-centric' interfaces. This research had three key aims: first, to identify the role of factors that enable the process of information re-finding, second to identify how to represent people in interfaces to support information sharing and re-finding tasks, and third to demonstrate and evaluate 'person-centric' interface approaches in providing novel support for information tasks, and identify design implications for their use. The final part of this discussion

revisits these aims, and considers the extent to which they have been met by this research.

8.1 Group construction

In order to explore how contacts (people) might be best utilised and represented within interfaces to support information retrieval and sharing tasks this thesis has considered how people think about groups of contacts as means of ordering them. The results of this have been the identification of factors which influence the construction of groups and implications for how automatically grouping contacts might be approached.

8.1.1 Implications for Automatic Grouping

Automatic detection of groups suffers from problems typical of many areas of automation, in particular that it can produce incomprehensible (though structurally valid) results, which lack meaning or sense to the end user. In the context of identifying social structures, automatic detection has been criticised in the past (Whittaker et al., 2004; McDonald, 2003; Jones and O'Neill, 2010) due to the potential for it to misinterpret, miss completely, or not have available important data when identifying groups. This in turn can mean that it often lacks validity when compared to the user's perception of what meaningful groups are.

The research of Jones and O'Neill (2010) demonstrates some of the complexity of this problem. While they could algorithmically identify at best 79.5% of social groups accurately from Facebook contacts based on ties between contacts – a tie being a connection between contacts, in this case in the form of having accepted a 'friend request' – they could not do this with any consistency, sometimes identifying as low as 16% accurately. Accuracy was higher when the person's social network contained distinct clusters and the participant chose to group contacts to reflect those ties between contacts. Accuracy was lower when the participant had less distinct clusters and chose to represent this through multiple groups. As in this thesis participants would often subdivide groups into categories such as "work other" and "work friends". Combining these groups into one single

group was found to increase accuracy as it reflected the ties between people the algorithms detected more closely.

While this illustrates how some groups of people can be identified by the mutual 'friendship' ties between them, it also highlights the myriad of other factors that might influence how people perceive groups to exist. Examining only ties ignores factors such as the commonalities between contacts which the grouper may perceive to be important connections that do not necessarily manifest in communication or mutual connections data. Or, viewed from another perspective, McCarty et al. (2007, p. 158) suggest that ignoring these other ties may be a reason that personal network visualisations fail to match categories generated by people, as those categories '*tend to follow a certain dimension, such as types of behaviour the respondent engages in... whereas social structures do not necessarily conform to one domain.*'

Similarly, when people create groups manually they do not create them perfectly, as the literature often assumes them to. Examining how people grouped their contacts in Chapter 7 revealed a tendency for participants to group along a number of criteria that would conflict with factual data that automation might use. The strongest association a participant had with a contact would often lead them to intentionally group the contact somewhere other than might be logical, for example, the place where a person was first met might lead someone to group their location incorrectly as that place, alongside other related people to that point in time, if their home had changed in the meantime. Contradictions such as these have implications for automatic group detection in terms of how automation should identify and handle the need to group inaccurately when presented with seemingly accurate data, and how it might analyse past histories of interaction with people to take into consideration aspects such as what are the strongest 'social' ties with people at different points in time.

Another issue that should be considered is that of the variance of manual grouping processes identified in this thesis. On many occasions participants would bend the internal rules they made for groups or continually expand them beyond their original definitions in order to reduce the number of groups being created and to ensure all contacts were grouped somewhere. This has implications for how

automated group detection is validated. Studies such as Jones and O'Neill (2010) have used manual grouping to validate the accuracy of automatically generated groupings. But if these manually created groupings themselves are inconsistent, or only describe one possible way a person has of understanding their social groupings, then a different criteria for judging the accuracy of automatically generated groupings may be required.

The way people think of and group others, is to some extent too individualistic, multifaceted, and context-dependent for consistently high levels of accuracy in automated groupings as judged according to correspondence with the groups the user would make themselves. It may therefore be better to examine how a process of forming and identifying meaningful and useful groups might be expedited through automatic detection, rather than by trying to identify them for people. The research of Amershi et al. (2012) is an example of this. Rather than identifying a person's social groups, it identifies and suggests a variety of possibilities for selection based on a variety of possible common known traits, such as location.

Work in automatic detection, such as that of Leskovec and Mcauley's (2012) on detection of groups based on network structure as well as profile features, and McCallum et al.'s (2005) detection of roles based on message content, suggests that the more complex multiplex groupings discussed here are potentially detectable as network features. Indeed Leskovec and Mcauley (2012) demonstrate that multiplex social dimensions very similar to those found in this thesis (Chapter 5 & 7), such as 'people with PhDs living in S.F. or Stanford' can be detected. Being able to identify these groups, however, only addresses one part of the problem, as it becomes necessary to determine which of these groups the user would perceive as being important.

8.1.2 Factors Influencing Group Construction

This thesis has identified a variety of factors which may influence the construction of groups by examining what people considered to be meaningful representation of their contacts through grouping. These factors are directly applicable to man-

ual grouping processes, but also to varying degrees have implications for how automatic detection and creation of groups might work.

In the third study (Chapter 7) – building on the results of the second study – participants were asked to create groups of contacts manually. As was expected, the manual grouping proved to be an overly intensive system for use in real world situations, as has been identified in other studies (Whittaker et al., 2004; Mazzia et al., 2012). This approach, however, had two advantages - it provided an opportunity to understand the reasoning behind the construction of groups and how people would ideally construct them, and it provided a point of comparison with manual grouping that often exists as a task within social network websites for managing contacts.

Figure 8.1 summarises the factors that were found to influence, constrain, and drive the construction of groups and illustrates how they fit together. As the diagram highlights, these findings could be divided into three areas: factors driving the creation of groups, factors determining the internal structure of the groups, and the goals group creation should aim to achieve. Within each section the diagram describes each of the factors influencing these three areas which are summarised in detail below. As the diagram indicates with arrows, each of these areas follow on from each other, each influencing the next. The reasons for group creation influence how a group might be defined, and how the group is defined affects the extent to which the goals of a useful group are met. The cube in the centre of the diagram represents the concept of the group, whereby the axes of the cube are symbolic of the idea identified in construct elicitation exercise (Chapter 5) that group definitions are multiplex in nature – that they are combination of multiple forms of construct that define it, for example, 'Nottingham Colleagues'.

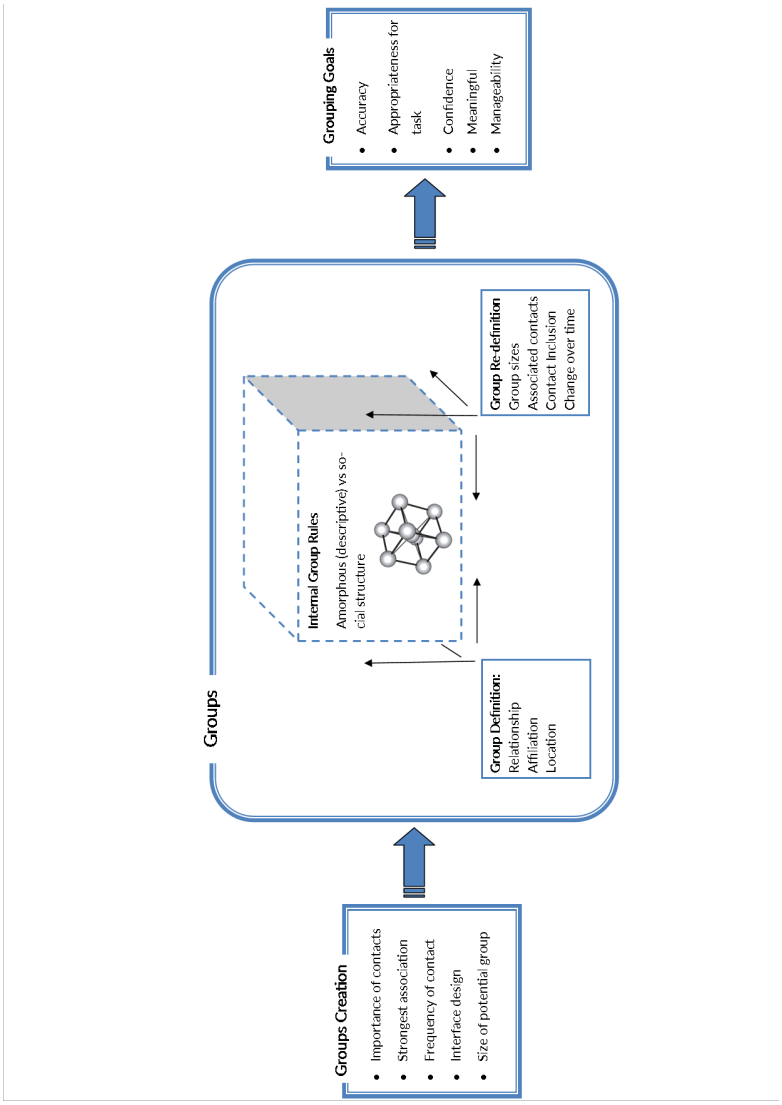


Figure 8.1: Contact Grouping Diagram

Group Creation

The following factors (shown in Figure 8.1) were identified as determining the definition of groups created or the inclusion of contacts in a group:

- **Importance of Contacts** – These are groups that people seek to generate due to their social importance and includes groups such as those that represent family and close friends.
- **Strongest Association** – The strongest association in common between a series of contacts will often determine how a group is defined. Where a set of contacts were first met, for example, often serves as a strong association, or as in the previous point strong or important relationships also have this effect.
- **Size of Potential Group** – The appropriateness of the size of potential groups also determines how groups are defined. There is a desire for groups not to be neither too numerous nor so small as to be meaningless, which influences which groups will be created. Groups that are obvious in terms of covering a large number of contacts, such as a generic 'Colleagues' group, for example, are often common candidates for this reason.
- **Frequency of Contact** – The frequency with which there was contact with the contacts was a strategy for determining which groups were created. This might be grouping frequent contacts together, but also as a factor determining which groups were created based on the frequency of contact of its potential members.
- **Interface Design** – The design of the interface can influence specificity and number of groups likely to be generated.

Besides the last point in this list, *Interface Design*, all of these factors were identified through analysis of the think aloud protocol data. These factors represent the themes that emerged when participants vocalised grouping strategies they were using to identify groups to create. For example, one participant described how they grouped a contact as follows 'This is a friend of my ex who lived in London, but now lives in Derby. I don't associate him with Nottingham so I'm

going to put him in London.' This represented an instance of the strategy of grouping by strongest association. For the final point of interface design, this was identified through the comparison of the number of groups created using two separate grouping interfaces, one that used a list of contacts and groups, and one that used a graph structure that enabled sub-grouping. The design of the latter grouping interface encouraged more groups to be created and a greater degree of modelling of social structures.

Groups

As described by the contact group categorisation described in Chapters 5 and 7 groups that are created are defined along one or more dimensions ('Nottingham Friends' would be two dimensions), the most common of which can broadly be described as relating to affiliation, relationship, and location. Figure 8.1 illustrated the dimensions that define groups as dimensions of a cube representing a group. The figure also describes factors which can cause the definition of the group to be re-defined, which act to some degree as constraints on group definitions. These are described below:

- **Group Sizes** – The desire to minimise the number of groups in use can result in groups being re-defined to be inclusive of more contacts.
- **Associated Contacts** – Contacts that are strongly associated with members of a group, but not part of it according to it's definition may be included within the group on the basis of their relationship to the group or a member of it.
- **Contact Inclusion** – If all contacts need to be grouped, at the end of a grouping process this can mean a few contacts remain, which don't fit into any of the existing groups. To ensure completeness this can also result in group re-definition rather than to create new small groupings.
- **Change Over Time** – If groups are permanent in nature they may need to be re-defined to handle cases where new contacts are met and also where relationships to contacts change over time.

The first three of these points were identified as part of the think aloud protocol

analysis. These were themes that emerged when participants hesitated in their grouping process and needed to consider how best to group contacts, leading them to re-define what constituted their rules for inclusion in a group. An example of this was a participant who repeatedly re-defined the area Nottingham included in order to minimise the number of groups he created – '*Well you know actually he only lives in Birmingham, that's Nottingham Area isn't it? He only lives an hour away.*' The final point of group membership changing over time was a identified as a significant issues in the literature (Viégas et al., 2004; Fisher and Dourish, 2004) and was indicated as significant consideration in some of the free hand representations of groups in Chapter 4.

Figure 8.1 shows that within each group there is an internal structure of contacts. Groups were found to have either structures that approximated real social structures, such as those between friends or family, or ones that were descriptive in nature in that membership meant the contacts were described by the group, but did not require all of the contacts to be connected by the same social structures.

Grouping Goals

Another pressure on group construction is the ultimate goal of the group. Grouping construction should aim to generate groups that adhere to the following series of goals, which seek to make groups useful in the tasks that they relate to. The factors which influence this are as follows:

- **Accuracy** – Groups should be created to be as accurate as possible and not include erroneous contacts. What makes a group accurate will be dependent on what the user of the group understands it to represent, what their personal model of the group's intent is and what contacts that dictates to them it should contain. Importantly this is distinct from what the group is named as may imply is correct.
- **Appropriateness for task** – Groups should be created with the task they will be used for in mind. Groups which suit the task they are being used for are likely to be used more efficiently as the user will have to spend less time searching for contacts if all relevant contacts are easily accessible.

- **Meaningful** – Groups should not be so small or so large as to cause a group to lose meaning. Groups should ideally represent a logical set of related contacts, this can be undermined by some of the factors that might cause group re-definition/constraints.
- **Manageability** – Similarly to the previous point having too many or too few groups can make a grouping system unmanageable in terms of causing there to be too many groups for the user maintain awareness, or too few groups causing the groups to lose meaning.
- **Confidence** – In order to allow users to perform actions with whole groups rather than checking each individual contact it is important they have confidence in the group. All of the previous points contribute towards this sense of confidence. Groups that are accurate and appropriate in size and to a task will likely give the user confidence as to what a group represents in terms of contacts. Lack of confidence about the contacts included groups may reduce the effectiveness of groups as a strategy for organising contacts.

These goals to some extent serve as counterpoints to some of the constraints that cause group re-definition. Factors causing re-definition of groups, for example, often involve introducing elements of inaccuracy which undermine goals, such as accuracy, meaning, and confidence. The goals described here apply equally to manual group construction and automatic detection of groups, which must seek to address these goals and consider more than just the ties between people.

These factors were identified from themes emerging from the think aloud protocol and interviewing when participants performed information tasks with the 'person-centric' interface using their previously created groups. Using the groups highlighted instances where they were unsatisfactory in some way. These goals represent what groups should aim for to avoid reasons that participants considered them poor for performing tasks. They are interlinked to some extent, for example one participant spoke of a task where in a group 'there's one person there I wouldn't eat out with, but with that group of friends I'd eat out with, or eat with at each other's houses quite a lot. There's one particular person who'd come up in that if I clicked on that group, who I actually wouldn't include in that

group. And that would include everyone.' This touched on interrelated issues of the group being accurate, meaningful, and having the confidence to use it.

The value of the factors identified in Figure 8.1 was that they highlight how group creation was often a process of compromise, a trade off between time, group size, group structure, and group meaning. When manually created groupings are discussed in the literature, they are often accepted simply as an accurate reflection of a person's perception of their social structures (e.g. Jones and O'Neill, 2010). The results of this study demonstrate how the groups that are created are not completely static, they are the result of compromises made that may change from time to time and are dependent on which and how many contacts are included in the grouping exercise. Additionally, that people have an overall lack of awareness of their own social structures contributes to further uncertainty about how groups will be created at any given time, since groups were often created in an ad hoc manner as they went along, often without much prior thought to the structure.

Also important was that the nature of the groups created could be influenced by the design of the grouping interface. This suggests that consideration should be given to the process of grouping when designing a grouping interface with regards to what the ideal outcome from the grouping activity is – as well as the other factors described here that influence group construction.

8.2 Implications for Information Tasks

In the second part of the third study (Chapter 7) participants were set a variety of information re-finding and information sharing tasks, with each of the person-centric interfaces developed to incorporate the findings from Chapters 4 & 5.

8.2.1 Information Sharing Tasks

The differing interface designs as well as the grouping interface designs highlighted the value of social awareness. Participants' responses to the interfaces were split. This largely divided along those who preferred the List interface due to its stable

and ordered nature and those who preferred to be able to take in the entire structure of their groupings at the same time in the link-based interfaces. The List interface was perceived as allowing the information tasks to be performed in a straightforward manner, but the inability to see all of the groups at once was considered problematic for others. There was also a strong preference expressed towards being able to see all of the groups and contacts, and their structure. McCarty et al. (2007) describes how people are rather poor at recalling all of their own social structures, the link interfaces would seem to provide support by improving peoples' ability to recall their contacts. In the same way that the participants in Jones et al.'s (2004) study did not want to lose their folders because they helped them to see the relationships between information, reminded them of things they needed to do, and helped people to make sense of their information, so did the contact groupings and visual feedback about the structure of their social network. This was particularly relevant for information sharing tasks which could benefit from this greater level of social awareness when considering who to share information with. It served both to allow participants to rule out sections of their social network quickly, but also reminded them of the presence of certain contacts or groups of contacts they might otherwise forget to include. This is similar to the benefit of social reminding – being reminded to honour communication commitments and to keep in touch with contacts – that Whittaker et al. (2004) sought to support with their person-centric interface. Being able to see all of their contacts at once, people are reminded of their social obligations to those contacts.

Another important implication for the design of person-centric interfaces was that of group confidence. In order to benefit from using the groups they had created, the participants needed to have confidence that they knew who all of the members of those groups would be. When grouping the contacts, grouping each contact sequentially and methodically, and giving each group internal rules meant that participants were able to have more confidence in the correctness of their groupings. Performing information tasks, particularly information sharing tasks, with whole groups, only works if the participant has confidence they know who they will be sharing information with. If participants doubted the validity

of the group in any way they were compelled to check each individual member of a group, negating any benefit of operating at the level of whole groups. This has particular implications for managing privacy on social networks and the uptake of using groups to manage privacy. If people are to use groups to determine who has access to information about them, it becomes increasingly important that the user has confidence in those groups.

8.2.2 Information Re-finding

Participants in the prototype evaluation did not like the unified point of access for both their Facebook and email accounts. This was because they considered that they would not want to have these contacts together, for many of their contacts they would only communicate with them through one medium or the other. This desire could relate to people who have work and non-work identities that are perceived as being bounded from each other to some extent (Kossek et al., 2012). In this case email would typically be part of a work identity and Facebook would relate to a personal non-work identity.

As a result there was therefore no benefit to presenting these two sources together, as they wanted to specifically go to Facebook or specifically go to their email accounts as the sources of the information. It did not make sense to participants to access the information elsewhere. This was similar to the process observed in the diary study in Chapter 4, where navigating to the source of the information was a critical first step in the re-finding process. While in this case there is the added factor of using the different mediums for different forms of communication with different sets of people, it still highlights again the importance of the 'source' that information is encountered in the process of managing personal information. This has possible implications for systems which attempt to integrate our personal information to provide a unified point of access for personal information management or provide universal search, such as *Gnowsis* (Sauermann et al., 2005) and *YouPivot* (Hailpern et al., 2011) respectively. If people are uncomfortable with this level of integration in certain circumstances, design of information management systems such as these will need to address this desire for segregation of information sources and the desire to navigate using memory

in a series of familiar steps.

While navigating to individual people was not a positive experience for many of the participants, as Whittaker et al. (2004) also found, participants typically preferred some form of more direct person search. Accessing information through people, however, was found to be positively received by most of the participants. This was particularly true with regards to separating links and attachments from the messages that contained them and making them accessible as items in their own right. When participants remembered information that was connected to a person and contained in a link or attachment, they preferred to find just the link or attachment, and did not feel the need to navigate to the message first. This suggests that people were happy to navigate according to the link in their memory between people and these pieces of information, and did not necessarily need to follow the exact same path they first accessed it through, so long as the information was ordered in the same manner in which they thought about it and was obviously more accessible to them than the alternative of how they might normally access it.

8.3 Person-Centric Design Guidelines

The results of the studies described in this thesis have a range of possible implications for the design of interfaces which may be to some extent 'person-centric' in nature. This section summarises the implications for design that can be found in this thesis. The implications described here are high level, but serve to highlight how the results of this thesis might be taken into consideration in design. Further work would likely be able to expand upon these aspects to make more concrete guidelines for design.

- **Structure**
  - *Links* – Supporting richer pictures of contacts, by providing the means to communicate the structures inherent within contacts can be valuable in bringing order to a large numbers of contacts. In this thesis this has been achieved through the use of groupings, and links between

contacts and groups, but this might be expressed by any other means of making explicit structures within the contacts that the user might perceive. By making explicit links to create structure between contacts this can provide a means of ordering contacts as part of information tasks and can aid navigation to and awareness of contacts.

– *Complexity* – While it can be beneficial to provide additional information to the user through visualisation of structure between groups and contacts, this should not overload the user with information. The visualised elements should meet the needs of the user by providing meaningful structure and order to contacts appropriate to the task it supports. Users will recall many details relating to social structure and standing when prompted by seeing the contacts together. The design should not necessarily be to model all details of a social structure, but to aid in ordering the contacts to perform tasks.

– *Awareness* – For the purpose of social reminding and information sharing, or any task that involves interaction with multiple contacts, it can be important for the user to maintain awareness of all relevant contacts and groups, as it can sometimes be difficult for users to recall or keep in mind all of their contacts. Providing increased awareness through design can aid users by reminding them of contacts who are relevant to a task, who might otherwise be forgotten. Design of interfaces should aim to provide a means of maintaining awareness of contacts, most simply, for example, by making all contacts or groups viewable at the same time.

– *Interaction* – If the interface relies on users grouping contacts, how users are permitted to interact with the interface can influence the groups they create. Aspects such as whether a structure between groups is visible, how groups and contacts are ordered and made visible, and the ability to create different types of groups may potentially affect the extent to which users identify or model groups of contacts. The interaction with contacts for grouping should therefore be appropriate to the types of tasks that will be performed with a group.

  - *Static Elements* – If there are many elements to interact with (in this context contacts) people learn the position of them on the screen in a similar manner that people recall paths through folder structures. When using these elements should remain static and retain the same locations each time the interface is used to support the user in accessing those elements.

- **Ordering or Filtering Elements**
  - *Ordering Contacts* – Location, organisation, relationship or other behavioural facets, such as frequency of contact, are the most appropriate dimensions with which to provide order or filtering to a body of contacts besides contact groupings. Each of these dimensions can be expressed in different forms; organisation might be expressed in terms of a the name of a particular organisation or the roles people play within it. When using this elements to provide order or structure consideration should therefore be given not only to what is the most appropriate form of order, but also how it is expressed – data permitting.
  - *Memory* – While the focus in this thesis has been to make people the primary unit of interaction, it may be that in many interfaces it takes a more equal, or supporting, role alongside other means of ordering and locating information such as time, title/subject, source, file type. Design should therefore consider what elements are most suited to driving the interaction and how the use of people works together with these other elements to support a task.
  - *Contact Importance* – Contacts and groups have different relative levels of importance, which are implicitly understood and determined by the user. This has been a useful means of highlighting contacts in past research (Fisher et al., 2007) and was highlighted as something participants cared about in this research. However, the use of this should not involve the user making many explicit judgements about contacts to define their importance as this can make the user uncomfortable.

Rather, this can be done through, for example, behaviour e.g. through frequency of contact.

- **Association Of Information With Contacts**
  - *Information* – Files, URLs, and information about a contact, such as phone numbers, when sought as part of an information task can be made more easily accessible by making them available in relation to contacts without the context that they were encountered within (e.g. within an email). The context should, however, be recoverable in relation to the file, link, or information.
  - *Group Level Interaction* – Allowing users to perform actions that utilise groups to include all the contacts within it can enable more efficient interactions. However, this requires that the user has enough confidence in their groupings to do this without checking individually who is within in each group.

This research identified 'person-centric' interfaces as an area where there was limited research into how they should be designed to support information tasks. The guidelines above seek to be begin to address how to design for information re-finding and sharing tasks. With further research it will be possible to both extend these principles and provide more specific practical guidance about how they are best applied in different contexts.

8.4 Using a focus on 'person' to support information re-finding and sharing tasks

The primary aim of the research described in this thesis has been to investigate how the links between people can be used to support information re-finding and sharing tasks. To this end, as is outlined in Chapter 1, Section 1.2, the research began with the objective of identifying the role of factors that enable the process of information re-finding. A number of factors, including person, source, and time were found to be significant, as is described in Section 8.4.1. A second objective, to identify principles to support the representation of people and their networks

in person-centric interfaces for information sharing and re-finding, emerged after this research confirmed the importance of 'person' as an underutilised attribute for performing information tasks. Section 8.4.2 indicates how this objective was met. A final objective, to demonstrate and evaluate 'person-centric' interface approaches in providing novel support for information tasks, identifying design implications for their use, followed logically from these findings. This objective is discussed in Section 8.4.3.

8.4.1 Objective 1: To understand the process of information re-finding and identify the role of factors that enable or affect its completion.

Examination of everyday information re-finding behaviours, in the diary study (Study 1 – Chapter 4), aimed to identify where greater support could be provided when performing information re-finding tasks. The study identified five primary aspects that were remembered in relation to information and used by subjects when they navigated back to re-find that information. First and foremost amongst these aspects was the *Source* of the information, the environment in which the information was encountered. Source in this sense is the structure that most closely contains the information that they can remember, such as a specific folder structure (e.g. a folder structure containing all the research for a thesis). Participants would use navigating to the source of the information as the starting point for the re-finding process, and recalling this would therefore play a critical role in determining the difficulty of a re-finding task. After the source of the information was located, participants would use a variety of other memories in order to locate the desired information, most often the other four attributes discussed here besides *Source*.

In addition to the *Source*, the other four other most commonly recalled attributes in relation to the information sought were *Keyword / Subject* (what the information was about), *Time* (when the information was received or last accessed or altered), *Person* (any person involved in the creation or later communication of the information sought), and File Type. Depending on the task and the inter-

face, memory of these attributes was important to varying degrees in locating the information, with - as previously described - *Source* being the most important to recall. While interfaces commonly either supported or focused on the use of most of these attributes, the lack of support for *'Person'* stood out (see Literature Review – Chapter 2). Despite being as prominent as keywords and time in the process of recalling information, interfaces were often neglecting its utility or were unable to utilise it to the same extent. Due to this, the role that *'Person'* can play as a key attribute for supporting performing personal information tasks was made the focus of this research.

The role of *'Person'* and the links made between them and information were not just connected to recalling who sent or created it. Important links between people and information were also made as part of information tasks regarding a variety of other aspects, such as recalling the need to share information with someone. This also sometimes included recalling a link between information and people, but also knowing that this link was not recorded as meta-data that could be used as part of an information task. These observations emerged from the diary study (Chapter 4), which identified the information needs that emerged as part of re-finding tasks. This revealed that participants were most commonly looking to re-find information about time sensitive events, guidance about a subject, information about products or service offerings, information about other people, and information concerning themselves. Just over a third of these information needs were related to checking information. This was by far the most common task. These tasks occurred where the goal was predominantly for the person to reassure themselves that they had correctly remembered a piece of information. Participants commonly lacked confidence in their memories and therefore needed to check information.

Analysis of the diary entries highlighted how many of the needs and tasks undertaken when re-finding information were about or related to people. A link between people and information was identified as existing in the following instances:

- Information sent to someone.
- Information received from someone.

- Information about someone else.
- Information that someone else may be interested in.
- Information about yourself.
- Information authored or created by someone.
- Information created for someone.
- People at or going to events and meetings, and their availability.
- People as a source of information.

The results of the study showed that difficult re-finding tasks were typically instigated by instances where the information being sought had not been accessed for a long time, and in cases where the user was not involved in the process of locating the sought information originally e.g. someone else found it for them initially. While the perceived difficulty of performing the majority of information re-finding tasks was rated as being low by participants, difficult information re-finding tasks still accounted for around 10% of tasks, and could take relatively long periods of time to resolve or remain unsatisfied. Further to the diary entries, interviews with participants also highlighted the need to take into account the role of frustration in performing tasks, which although considered easy in their execution could not be performed as efficiently or simply as might be wished. Elements such as mitigating frustration – and other negative affective responses – or improving the efficiency of the process therefore need to be addressed, as well as difficulty. It is in this role therefore that improving the ability of knowledge of a 'person' to be utilised in re-finding can provide greater support for personal information tasks. Depending on the application being used, the user may be limited in one way, or another, in terms of what aspects of the information they recall can be used to locate the information they require. Looking at the role of 'person' in this process of re-finding opens up a further means of locating information, increasing the chances of finding it promptly and without frustration.

While there is more that can be understood about the processes taking place as part of information re-finding tasks, it is clear that there exist strong connections that people make between information and other people and that they readily

recall them in relation to information. It is also clear that those connections can be leveraged to a greater extent to support information tasks. This research has outlined how the process of re-finding is carried out and underlined the important untapped potential of memories of 'person' in the process. It has also considered the purpose and needs being addressed by performing re-finding tasks, and as a consequence the multiple senses that 'person' can be considered as part of them. As a result of these findings, the research conducted throughout the rest of this thesis has focused on understanding how representation and interaction focused on people can help to support personal information tasks.

8.4.2 Objective 2: Having identified recall of links between people and information as an important, but underutilised part of re-finding search strategies in Objective 1, identify how to represent people in interfaces to support information sharing and re-finding tasks.

In addressing the first objective, connections between people and information were identified as being an important aspect of the information re-finding process, which could be utilised to a greater extent by interfaces to improve support for performing personal information tasks. Towards increasing the support provided by this approach, the next step taken was to establish a basis for how best to design interfaces that make use of this attribute and its multifaceted connections to information to support personal information tasks. Very little of the research in this area provided an understanding of how to represent people or contacts as a means of performing information tasks. Gaining an understanding of the structures and relationships people perceive between their personal contacts was therefore considered to be critical for enabling the design of interfaces that allow information tasks such as re-finding and sharing information to be performed in a 'person-centric' manner. To this end, the second study interrogated how people choose to represent their contacts and the relationships that they perceived between their contacts. This was detailed in Chapter 5, which focussed

on addressing these two elements of representation and relationships.

Representations of personal contacts were studied through a freehand diagramming exercise where participants drew how they would represent all of their contacts, however made most sense to them. Contact relationships were examined by having participants compare contacts through a modified repertory grid task that elicited what commonalities were present in how they perceived different aspects of their personal contacts. From analysis of these two exercises, salient features were identified in the data that could aid the design of person-centric interfaces by providing a basis for organising and interacting with contacts according to important facets that reflected their own understanding of their personal contacts.

The results from the diagramming exercise were categorised into different forms of contact group representations, each of which emphasised different factors important for grouping contacts.

- Spider diagram – Emphasised group structure, particularly the presence of links between groups and sub-groupings.
- Grid diagram – Emphasised the presence of multiple interlocking social dimensions and criteria for determining groupings.
- Venn diagram – Emphasised the overlap between groups and contacts, and their relative sizes.
- Axis diagram – Emphasised the relative importance of different contacts and groups of contacts.

From the repertory grid contact comparisons, it was possible to create a formal set of categorisations describing the different forms of relationships participants perceived between contacts (see Figure 5.1). This was then applied to the results of the comparisons to examine the make up of the types of relationships participants identified. This categorisation was validated through the results of the contact groupings created with the person-centric prototype system in the third study (Chapter 7). This demonstrated their applicability for describing the contact groups people create or use in practice to perform information tasks,

demonstrating that eliciting the range of relationships perceived between contacts could be applied to understand and describe contact grouping. Application of the categorisation to both the repertory grid constructs (Chapter 5), and groups created by participants using the prototype person-centric interface (Chapter 7), identified the following four attributes as being the most common types of relationship shared between contacts at the broadest level:

- Organisation – Work and non-work related organisational structures.
- Location – Places contacts are located in from specific buildings to broad geographical areas.
- Relationship – Forms of social relationships between groups of people, friends, family, colleagues etc.
- Events – One-off and recurring events that contacts participate in (Note: this did not rank as highly in the Chapter 7 results)

At the lowest level, the comparisons highlighted a wide variety of nuances in thinking about people, including past, present, and future states of relationships, defining groups of people through negation of other groups of people, and using words to describe the intimacy of relationships to qualify groupings. The categorisation provided a means of distilling complex relationship descriptions into categories that describe the intended form of relationship between the contacts. This also highlighted that participants were sometimes using multiple elements of the categorisation in single group descriptions. These contact groups were multiplex in nature and could be best described using a variety of single or multiple social dimensions to construct a description. A group consisting of 'Sheffield Friends', for example, would be described as being a combination of both a Relationship and a Location in order to accurately describe it. This further illustrated how the different forms of groups identified could be used in combinations to specify or identify groups according to a range of criteria and at different levels of specificity.

Expanding on the findings of study two, two interfaces were developed (see Chapter 6) that allowed participants to create groupings of their own personal contacts

in slightly different ways, one emphasising structures of links between contacts, and one emphasising an ordered list format. In study three (see Chapter 7) these interfaces were used as a means to explore the decision making processes involved in contact grouping and the factors influencing the construction of contact groupings. Understanding what factors influenced the way people construct groups, can in turn be used inform how interface design can support the use of groups and how they should be represented.

Group construction was influenced by factors driving group creation, group redefinition, and the goals that groups should aim to fulfil to be useful. As outlined in this Chapter in Section 8.1.2. The factors described in 8.1.2 can be used together to inform the design of 'person-centric' systems, as is discussed in Section 8.3. This might either be in terms of the design of person-centric interfaces that utilise groups, or the automated detection of groups. For example, in the case of interface design, in order to achieve the goal of keeping groups manageable, an obvious design implication of these factors would be to have limits on the number of groups that should be displayed on-screen. Or in order to ensure groups are meaningful, using the more common dimensions used to define groups might be an appropriate criteria to order groups or contacts by. In the case of automatically detecting groups, dependent on the goal of the system, consideration might need to be given to limits on the size of individually detected groups, whether groups are going to be based on social structures or specific facets in common, and how elements such as importance or strongest association can be identified and factored into the design to identify meaningful groups.

In seeking to address this objective this research focused on grouping contacts together as a means of accessing and sharing information. To this end it has identified a range of factors that are important for the representation of people, first by asking people about the relationships they perceive between contacts and how they would choose to represent them, and second by examining the decision making processes involved in creating groups manually. These factors make a contribution to an understanding of how design should represent contacts and and seek to guide group construction. It does this both in terms of the structures that are valuable for making sense of and utilising groups of contacts, and in terms

of the relationships that define how they should be ordered and generated.

8.4.3 Objective 3: Demonstrate and evaluate 'person-centric' interface approaches in providing novel support for information tasks, and identify design implications for their use based on the findings of Objective 2.

The final objective of this research was to apply the earlier findings to the design of a person-centric interface and evaluate it. This brought together the previous research identifying the value of connections between information and people, and the role this can play in improving the experience of performing personal information tasks, and the subsequent investigation into representing people in interfaces for personal information tasks that elicited a series of different forms of structures and relationships for representing personal contacts. However, a key objective remained - to understand how these findings could be used to guide design of a 'person-centric' interface.

The design of the person-centric prototype interface (described in Chapter 6) was influenced by analysis of the different relationships participants identified between contacts, and the different ways in which they represented their contacts in study two (see Chapter 5). This primarily involved different forms of presenting and ordering the contacts to support re-finding and sharing tasks, for example, allowing contacts to be ordered on the basis of location or relationship, or visualising overlaps between groups. Specifically, the interfaces developed aimed to aid people in performing information re-finding and information sharing tasks with their own personal data. These two areas identified as having the potential to benefit from a person-centric approach. Evaluation of these interfaces (Chapter 7) sought to determine the utility of the representations based on those factors.

One of the primary considerations to emerge from the results of the evaluation (Chapter 7) and the second study (Chapter 5), was the role of representing the structure of the contacts. The results demonstrated that how contacts are ordered is not the only consideration in facilitating information tasks. This study showed that how we represent the relationships between contacts, between contacts and

groups, and between groups is also key to supporting tasks. As is described in Section 8.2 this research identified a number of reasons for this, including the ways in which interface design can increase social awareness or encourage the creation of larger or more specific groups.

Analysis of the results of study three also identified a number of elements that were important in visualising the structure of the contacts. Connecting links were found to be an important means of conveying structures between both contacts and groups, being useful for both visualising membership of contacts to groups, group (parent) and sub-group (child) relationships, and relationships between contacts. The value of this structure, however, was found to need to be tempered by consideration of the task's goal. Providing too much information about social relationships was found to be either unnecessary or confusing. This need to retain clarity was similar to the findings discussed in Objective 2, where the number of groups created was identified as being a factor in group construction, in that participants were often keen to limit the number of groups created, in part due to a desire to keep all the available groups in mind and reduce complexity (see Section 8.1.2). This also applied to using the groups for information tasks and as with links should be considered as part of managing complexity to an appropriate level for the task. A clear conclusion therefore emerged: the right balance of complexity needs to be struck when visualising structure in order to clearly convey relevant information about social connections, not necessarily all of it. A full and detailed map of social connections was not necessary, and could be unhelpful, when using a person-centric interface for most information tasks.

As described in the previous objective, structure was also found to be significant for group creation. In the evaluation of the prototype, participants were asked to create groups to perform information tasks using two interfaces, one of which was a gridded list of contacts and groups and one which visualised the structure between groups and contacts via links – enabling the creation of sub-groups. Using the latter interface led to a tendency among participants to model relationships between their contacts in greater detail. The implication being that providing feedback about structure – or determining the structures that are allowed to be created – can have an impact on the resulting groups generated, which in turn

may have an impact on how groups are used to perform a task. Consideration must therefore be given to how interface design can encourage users to create groups in the most beneficial form for the tasks they will be performing.

An important issue that was identified for person-centric representations was that of *confidence*. When participants used contact groupings as a means to share information, their confidence in whether they could remember the members of the group was a critical factor in determining how those groupings were used. When there was confidence that groups contained only the contacts expected then they would perform group level actions with a group; when this confidence was lacking each group member had to be checked greatly lengthening the time to complete sharing tasks. As Figure 8.1 highlights earlier in this chapter, a number of factors during group construction may lead to groups whose scope is revised and become less ideal reducing confidence. Equally, when locating a contact in a re-finding task, being unsure of how people had been grouped could lengthen the time taken to search for a contact. This suggests that design should support or encourage people to create groups they will have high confidence in or help maintain their confidence and reduce uncertainty once groups are created.

Another issue that was identified as requiring consideration in design was the use of importance as a means of ordering contacts. Design which encourages or requires users to make explicit relative judgements about the *importance* of contacts was identified as being highly undesirable as it made participants uncomfortable. An interesting observation, however, is that grouped contacts tended to have similar levels of importance, indicating that participants were happy to make these judgements implicitly. Importance was also identified as being an important attribute in Chapters 5 & 7, for providing order to contacts or groups. That there is simultaneously a desire to not make judgements concrete, but also acknowledgement of the value of using the inherent differing levels of contact importance presents a distinct challenge. Design should therefore be sympathetic to this and aim to minimise explicit judgements if importance is used to order contacts. One means of addressing this has been through behaviour, for example, highlighting frequent contacts (Fisher et al., 2007), as this allows the user to abrogate responsibility for making the judgement.

The interface prototype created in this research (Chapter 6) to evaluate the factors identified in Objective 2 was largely successful in demonstrating the concepts identified. From this prototype, the evaluation of those concepts enabled the implications of using them to be identified as well as implications for the design of person-centric interfaces. While there has been some previous research into 'person-centric' interfaces, very little of this research has sought to provide a basis for how they should be designed to support information tasks. The research described here provides a foundation that informs the design of such interfaces. In the previous section of this chapter factors that have implications for the design of 'person-centric' interfaces were detailed. This highlighted aspects such as the need to generate confidence in groups and the value of providing awareness of contacts as important design goals. There however, remains much scope in future work to conduct more focused research that demonstrates concrete examples of how these design implications should be implemented.

Chapter 9

Conclusion

9.1 Introduction

This thesis began by considering what the impact of the increasing amount and diversification of information sources and forms might be on our ability to perform personal information tasks. In order to understand more about how personal information tasks are currently being performed information re-finding tasks were initially focused on. By examining how people are performing everyday information re-finding tasks, using a diary study (Palen and Salzman, 2002), the aim was to identify where greater support for those tasks might be provided. The diary study (Chapter 4) identified time, and keywords or subjects, and people as being important memories that enabled information re-finding. From these, the connections that were being commonly recalled between people and information were identified as being an area that had yet to be exploited fully as a means of aiding personal information tasks. People often recalled a person in relation to a piece of information being sought, but they were also often unable to leverage that knowledge to their advantage easily when re-locating information, commonly due to a lack of support for locating information in that way. This prompted the question of how to design interfaces that are focused on people for performing information tasks – that are 'person-centric' in nature – and concomitantly how people think about and wish to organise their contacts. By first establishing what relation-

ships and structures people perceive between their contacts, this could provide a basis for designing interfaces that represent their contacts to support information tasks.

In order to design interfaces that provide a means of accessing or sharing information through interaction with personal contacts equally requires a means of organising and representing those contacts. Similar to how the diary study revealed some of the prominent dimensions that information is recalled and ordered by, the design of 'person-centric' interfaces required a similar understanding of what dimensions influence how to structure and order contacts. The next step was therefore to understand what relationships people perceive between their contacts, with which people can be organised and visually represented, and how to structure them within interface designs. Knowledge elicitation tasks adapting Repertory Grid (Fransella and Bannister, 2004) to identify relationships between contacts, and free-hand diagrams (McCarty et al., 2007) to identify different forms of 'person-centric' contact representations revealed a range of salient factors for designing 'person-centric' interfaces based on how personal contacts are perceived. These included aspects such as the prominence of organisations (both work and non-work related), location, relationship forms, the varying importance of contact groups, and providing feedback about links, overlaps, and hierarchies between groups. The findings about how people perceive their contacts and how they re-find information with them was then used as a basis for designing a 'person-centric' interface with which to evaluate the utility of the different identified factors in aiding information re-finding and sharing tasks. This demonstrated that how 'person-centric' interfaces represent contacts does have an impact on their use. It also identified factors that influence the organisation of groups of contacts to enable the performance of information tasks with 'person-centric' interfaces. Further it demonstrated that connecting people and information is a valuable means of supporting personal information tasks.

This chapter describes the contribution made by the research detailed in this thesis, it examines the limitations of the research approach taken, and considers what direction future work might take to build on this research.

9.2 Research Contributions

The following section outlines how the studies conducted contribute to the literature across multiple areas.

9.2.1 Study 1 – Re-finding Information

- *Developed two new and one modified analytical frameworks for the analysis of re-finding events in context.*

  Analytical frameworks were developed to describe the diary study data. These frameworks described information re-finding in terms of information needs, information recall, and information tasks. While in other diary studies (Dearman et al., 2008; Church and Smyth, 2009) diary entries were considered in terms of categorisations describing information needs, these studies either did not focus on re-finding, or did not consider the relevance of memory and the task the information was being sought for. These analytical frameworks extend work done looking at how information needs are addressed to describe how, what, and why information re-finding events occur. This also contributes to research on recall of meta or contextual information (Blanc-Brude and Scapin, 2007; Capra and Pérez-Quiñones, 2003; Elsweiler and Ruthven, 2007). These studies identified attributes people recall about information within specific contexts, such as email or system files. This work contributes to this body of literature by providing a natural cross-section of recalled attributes that is representative of re-finding as a whole. This highlighted the salience of the source of information, time, keywords, and people as aspects commonly recalled and utilised across information re-finding tasks

- *Reinforced the results of previous studies with regards to information re-finding events and the process of re-finding, and added to that work by examining re-finding across both work and home contexts, as well as across digital devices.*

  A range of previous diary studies have been conducted, such as Dearman

et al. (2008); Church and Smyth (2009); Sohn et al. (2008), which examined people's information re-finding behaviours capturing a cross-section of information re-finding events. These previous studies have largely focused on a particular context of information re-finding, such as mobile phone use. The diary study conducted in this thesis contributes to this body of work by providing a broader perspective on information re-finding behaviours and instead providing a cross-section of behaviours across both personal and work contexts as well as across all digital devices. Additionally, by applying the information re-finding analytical frameworks to describe the data this provided a more complete picture of re-finding tasks. For example, by considering the information needs, task, and process together this made clear that most common information re-finding tasks were about finding information about events, that much of this was driven by performing checking tasks where information was sought to re-mind or reassure the user, and that for these tasks the use of time was a prominent part of the re-finding process.

Additionally, the results of the diary study also corroborated the existing literature on the re-finding process, confirming that re-finding is a distinct process from information search, and that people navigate to information they are re-finding as part of what Teevan et al. (2004) calls 'Orienteering' and Capra et al. (2005) describe as two stage iterative process of navigation – first locating the source of the information and then working to locate the information within that source.

- *Identified information re-finding tasks as largely being perceived as easy to complete.*

  Previous research (Blanc-Brude and Scapin, 2007; Teevan et al., 2004; Boardman and Sasse, 2004) has suggested re-finding information is a commonly difficult or problematic task. However, the results of participants rating the difficulty of the tasks they performed in the information re-finding diary study suggest that re-finding tasks are not typically considered difficult. One reason as to why the results of this study differ is that it examined it a cross-section of difficulty in everyday re-finding, whereas in other studies

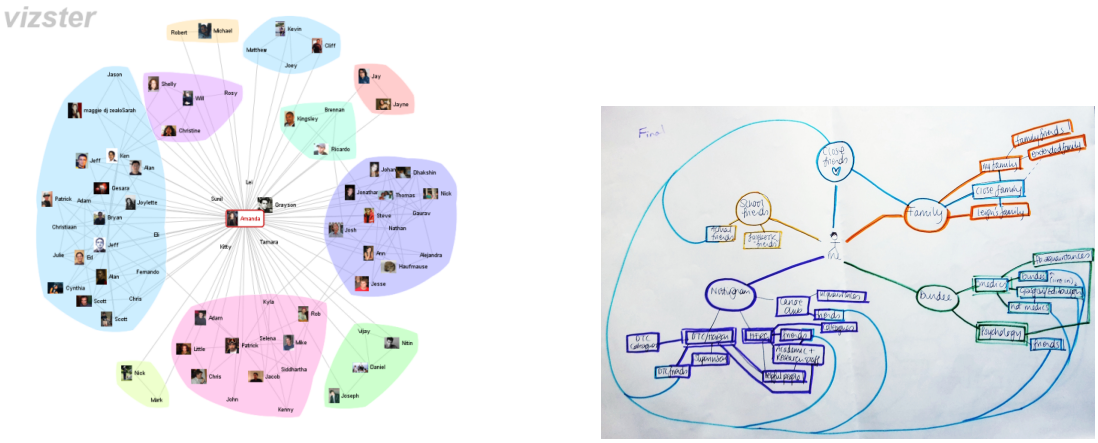
this has sometimes included less naturally occurring or arbitrary re-finding tasks. Tasks such as these can be inherently more difficult, because the information sought may be out of recent memory, whereas in this study many of the re-finding tasks were recently re-accessed – 53.9% being re-found within a week of last encountering the information (see Figure 4.2) and 76.1% of the information being re-found within one month. As this thesis and others (Elsweiler and Ruthven, 2007) have identified, an important factor in re-finding difficulty is the amount of time lapsed since it was last accessed. This research paints a more balanced picture of typical re-finding tasks, in which difficult tasks are common, but infrequent relative to the number of daily re-finding tasks occurring. The results also point towards the need to take a more nuanced approach to task difficulty in information re-finding, and address the role of factors such as frustration, which was an evident issue in some tasks perceived as being easy. By exploring difficulty of these tasks in greater detail, this may provide greater scope for improving the experience of re-finding and allow efforts to improve support for re-finding to be more effectively targetted at improving efficiency and mitigate factors such as frustration.

9.2.2 Study 2 – Contact Grouping: Categorisation and Representation

- *Identified distinct types of representations of contact groupings that people perceive in order to structure their contacts and make visible relationships between them.*

  The categorisation of free-hand drawn diagrams which represents the different ways participants viewed contacts constitutes a series of different ideas about what people value in terms of organising and structuring contacts as groups of contacts. As demonstrated in this thesis these different forms of ordering contacts that the categories identify can be translated into design.

  The different emphasis participants placed in their diagrams to some extent reflects the two different approaches identified in the literature review; those

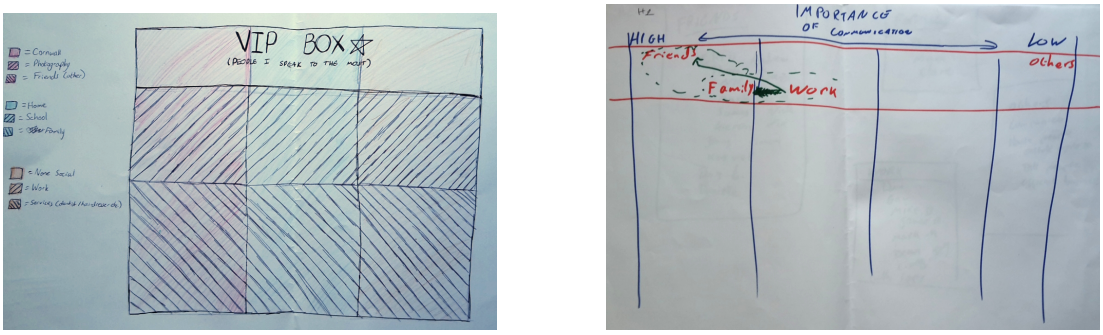


(a) (b)

Figure 9.1: (a) Vizster (Heer and Boyd, 2005) (b) Spider Diagram

interfaces focused on awareness of social structures (e.g. Offenhuber and Donath, 2007) and those focused on information tasks through identifying and understanding social interactions (e.g. Fisher et al., 2007). Diagrams such as the *Spider* and *Venn* diagrams emphasised social structures and diagrams such as the *Grid* diagrams and the *Axis* diagrams emphasised the complexity of social interactions and ordering.

In Heer and Boyd, (2005)'s *Vizster* system, for example, the design represents contacts as a graph, in which contacts that know each other are clustered together and highlighted in different colours as groups (see Figure 2.4). Visualisations such as this highlighted the social structures between contacts to the user. This is similar to the *Venn* and *Spider* diagrams, which created structure between groups through links and overlaps between groups and sub-groups to make clear the structure of how different groups of contacts fit together. Figure 9.1 shows this similar focus on structure and colour to Vizster to convey the make up of different contacts.



(a) (b)

Figure 9.2: (a) Grid (b) Axis Diagrams

In contrast, the Grid and Axis diagram types reflected aspects of the literature focused on understanding social interactions. This includes literature such as SNARF (Fisher et al., 2007) which sought to triage email through functionality that highlighted emails based on behaviour and interaction, such as filtering based on contacts you most frequently communicate with. Similarly these diagrams tended to emphasise ordering of groups according to attributes such as importance or intimacy of relationships.

- *Adapted repertory grid technique to use it in a novel application.*

  As part of the methodology of the second study (Chapter 5) repertory grid technique (Fransella and Bannister, 2004) was adapted by applying its construct elicitation to identify relationships between contacts, which were identified in a similar manner as in social network studies such as (Lubbers et al., 2007). Applying repertory grid in this manner is similar to techniques used to construct social network graphs, where comparisons are made between people by the participant to build up relationships in the graph (McCarty et al., 2007). The methodology provides another means of eliciting relationships between multiple people, but crucially focusing on identifying all of the relevant relationships in a domain, rather than on putting specific forms of relationship into context.

- *Described a categorisation which identifies the breadth of forms and ways people think about relationships between their contacts. This was done both*

at a high level identifying the concepts the groups were describing and at a lower level highlighting the different forms of expressing a particular concept.

Analysis of the repertory grid construct elicitation produced a categorisation identifying both the high level themes that describe the forms of relationship perceived between contacts and the lower level themes within them which reflect the different ways that those high level themes can be expressed. The most prominent high level themes identified from the contact relationship constructs were all present in the data of at least 80% of the participants and were as follows:

- Organisational (20%)
- Forms of Relationship (14%)
- Location (13%)
- Characteristics (10%)
- Expertise (10%)

These themes, both at a high and low level, were valuable both within this research in contributing to design, by identifying means of ordering contacts in the 'person-centric' prototype design (Chapter 6) but are also relevant to research on contact grouping, for example, within social networks (Jones and O'Neill, 2010; Leskovec and Mcauley, 2012; Amershi et al., 2012). As was discussed in Chapter 8 (see Section 8.1) understanding how users choose to group people may have relevance for research on automatic detection of groups, where there is a challenge in identifying groups that are appropriate to a task.

9.2.3 Study 3 – Person-Centric Prototype Evaluation

- *Explored the decision making processes involved in group construction. Understanding this was significant for understanding what factors people influence group construction, as summarised in figure 8.1. This may provide useful insights into group creation either for manual, semi-automated, or*

automated group identification processes.

There has been little literature exploring the creation of groups of contacts for performing information tasks. Some research has concentrated on either filtering/ordering content or managing privacy settings on social media (Jones and O'Neill, 2010; Mazzia et al., 2012). While this research has considered the forms of groups people tend to create, which is also explored further in this thesis, none of them examine the decision making processes involved in generating groups. Previous research has focused more on what groups it is possible to detect or create, rather than exploring what constitutes a meaningful group to a user. Jones and O'Neill (2010), for example examined how manually created groupings can be used to validate automatically generated groupings, but did not consider how those manual groups are arrived at, or the implication this has for automatic group generation. Understanding the processes involved in grouping provides useful insights into group creation either for manual, semi-automated, or automated group identification processes.This thesis both outlines the decision making process involved in creating manual groups, as means of identifying the user's ideals for group construction, and the factors that influence how and why they are constructed and possibly re-defined (see Figure 8.1).

- *Evidenced how interface design can impact on the on the construction of user generated groups.*

  Examination of group construction using two manual grouping interfaces – one grid based and one visualised as an egocentric graph – illustrated how interface design can impact on the content and construction of contact groups. In the evaluation of the 'person-centric' interfaces, the graph visualisation interface significantly affected group construction in terms of the number of groups created by causing participants to model their contacts to a greater extent, resulting in greater specificity of groups. Previous work on contact group construction does not consider the role of the interface in this process. While novel interfaces for performing group construction are presented in these papers – whether manual (Whittaker et al., 2004), semi-automated (Amershi et al., 2012), or automatic (Jones and O'Neill,

2010) – they do not examine the impact of interface design on group construction. As this research demonstrates the design of their interfaces can have an effect on the groups constructed beyond the novel feature they are demonstrating.

- *Identified factors that have implications for 'person-centric' interface design, related to how contacts and groups are represented.*

  This research identifies a range of factors which have implications for the design of person-centric interfaces (see Section 8.1). These factors primarily relate to two aspects, how the representation of contacts and groups are structured, and how contacts are organised/filtered. In the literature review 'person-centric' representations were identified that either aimed to represent structures of social networks (Heer and Boyd, 2005) or highlight information behaviours with people (Viégas et al., 2004). The representations explored in this thesis relate more to the former kind.

  In the literature, the purpose of these representations was primarily to represent social structures in order to enable people to understand and reflect on their behaviour on social networks (Offenhuber and Donath, 2007) (see Figure 9.1). While reflection can be a valuable activity it does not practically contribute to performing personal information tasks, which this thesis aims to support. Where previous research has emphasised the representation of people through social structures to support reflection, this research also demonstrates how it can practically be applied to support the tasks of information sharing and re-finding. Representations of this kind had been used in the literature in the past (Heer and Boyd, 2005), but they had not sought to aid specific information tasks. Many of the representations that emerged from the free-hand drawing task data in Chapter 2 were akin to interfaces in the literature in how they represented contacts (see Figure 9.1, suggesting that representations of this kind do emphasise the structures many people wish to represent about their contacts.

  In this research the representations of social structures added practical support to the tasks being performed in three ways. Firstly, the graph visual-

isations of contacts and groups facilitated interaction to perform tasks by providing means to locate and interact with contacts through social structures. Secondly, this structure also provided a greater sense of awareness of contacts who may be relevant to a task. This was two fold, in that by making visible all of the groups at once it reminded participants of all of the groups of people that may be relevant to a task, and also by making all of the contacts present reminded participants of contacts who may not have otherwise been considered at all. Thirdly, it provide a means of modelling relevant relationships and using them to perform tasks at a group level, such as sharing information with a group of people. This highlights how visualisation can be used to offload some of the cognitive burden of recalling some social information about contacts and move it to faster perceptual processes (Card, 1996), by essentially acting to remind user's of their contacts. As reported elsewhere (e.g. Whittaker et al., 2004) one of the wider benefits of a person-centric approach to interface design is that the focus on people helps to support social tasks, such as reminding people of obligations to people, or as in this case reminding to include people in sharing tasks.

By identifying factors that may influence the design of 'person-centric' representations, this thesis adds to previous work as it considers the practical implications of its use for supporting information tasks, something that the existing research has only done to a limited extent.

- *Validated the categorisation of contact relationships described in Study 2 as being suitable for describing the make-up of contact groupings (with minor amendments).*

  The categorisation of contact relationships described in Study 2 (Chapter 5) identified different forms of describing relationships between contacts at both a high thematic level (e.g. location-based relationships) and a lower level that expresses more about the nature of the relationship (e.g. living together). By applying this categorisation to the results of the contact grouping exercise (Chapter 5), with the participant's own contacts, the categorisation was validated as being of value not only for delineating dif-

ferent forms of relationship, but also in being able to describe usefully the make-up of different forms of contact groupings. This contributes to a limited amount of data on manual group construction (Leskovec and Mcauley, 2012; Jones and O'Neill, 2010; Whittaker et al., 2004). While studies concerning contact grouping created do also identify different forms of groups in their data (many of which are also described by this categorisation) this work represents an attempt at a more formalised categorisation of the extent of contact relationships that can be used to describe contact groups. This has value both for research and design. Future studies can refer to this categorisation to describe contact groups and contrast results. While research or interface design involving group identification, suggestions, or ordering and filtering can refer to the and the relationships – and how common they are – to determine which forms of relationship are most suitable or most relevant.

- *Confirmed value of links made between people and information when re-finding information as described in Study 1.*

  In the information re-finding diary study (Study 1 - Chapter 4) the association between people information for re-finding task was identified as being useful for information re-finding tasks. In particular it highlighted that people often recall information in terms of a person they associate it with, and that making information accessible in this manner should therefore support people in the information re-finding process. Replicating this process of recalling a person in association with the information in the 'person-centric' interface is quite similar to some extent the work of Whittaker et al. (2004), which associated emails with people – as in this study. By associating related information such as links and attachments with a person, rather than with a message, this study differed from their work. Associating information directly with a person in this manner was considered beneficial due to the ease of access from not having to locate the messages they were located within. That this method of accessing information was considered positive goes towards confirming that this can be a useful means of organising and accessing information, that reflects one aspect of how people re-find

information.

9.3 Implications of this Research

The research presented in this thesis provides a foundation for understanding how interfaces focused on the role of people in managing personal information can be designed, by exploring some of the underlying processes involved in their use. This is achieved by focusing in greater depth on the process of information access via people, the representation of contacts and relationships between them, and the process of contact grouping. This was done through an evaluation of novel 'person-centric' interfaces and the study of people's current information behaviours. This highlighted pertinent behaviours that may have implications for how users can be supported in their personal information tasks – particularly those of information re-finding and information sharing. Equally one of the justifications for pursuing this line of investigation is that social behaviours are intertwined with information production, use, and memory. This would also suggest that 'person-centric' interaction can be beneficial in other arenas as our ability to capture and interpret social behaviour increases. The findings and approach of this research could therefore be applied more broadly.

In examination of everyday information re-finding this research has helped to establish three analytical frameworks that describe how, what, and why people re-find information, and through the use of these frameworks explored what typical re-finding behaviour looks like. This work makes a valuable contribution to information re-finding research by providing a basis for future studies to describe information re-finding behaviours, the results of which can be compared across studies. Having an understanding of our information re-finding behaviours – as is provided in this thesis – can be used as a basis for guiding the design of systems that can improve the support provided to people in performing personal information tasks.

The most obvious applications of this research relate to similar communication mediums to those explored in thesis. Enterprise email clients in particular, such as Microsoft Outlook, are increasingly adding modes of viewing or accessing in-

formation related to contacts, as well as about people, within organisations – although the degree these features are well integrated into the work flows of users is unclear. Integrating people to access information could also be extended to the design of operating systems as a point of integration, bringing together information from the local file system together with sources from other applications, similar to Rekimoto (1999); Krishnan and Jones (2005) in their use of time and task based desktop environments. At a step slightly further removed, managing information in relation to people and understanding how best to represent people in interfaces could also be a valuable in instances where a large amount of information might need to be held about people, such as can be the case in project management.

This thesis also argued that associating information with people in an information system first requires some form of interface through which our contacts are presented and managed. In response, this research focused on the question of how to order and represent contacts to support information re-finding and sharing tasks. In answering this question a categorisation was developed through which to describe and quantify the different forms of ways that people group contacts. This identified organisational groups, location groups, relationship groups and presence at events as being the most common connections made between contacts. Examining this in more detail, however, demonstrated that groups are multiplex in nature and can be thought of and defined in multiple ways at the same time. It was not that each group was definitively one of these types of group, it was rather that for the most part each type of group identified could be applied to each contact. How to represent the relationships between people was also examined. Participants were encouraged to draw their own representations of how they perceived their groups of contacts. From the analysis of these diagrams attributes such as overlap between groups of contacts, the structure and relationships between contacts and groups, the relative importance of contacts and groups to each other, the location of contacts and groups, and the multiplex nature of contact grouping were identified as being important when representing contacts.

As has been discussed in Chapter 7 a potential application of this research is in the

design of social media platforms. Specifically an understanding of what forms of group are appropriate, useful or meaningful, how they should be constructed, and how they might be represented can contribute to the design of features for how privacy is managed, how events are managed, and information is shared. Equally, this work may prove beneficial for identifying appropriate means of organising and accessing information for the purpose of research into social media, or where there are large datasets relating to people require visualising.

The value of 'person-centric' interfaces, in contexts such as this, will to a large degree rest on the ability of systems to present contacts in socially meaningful ways, whether this is through manual grouping of contacts (e.g. ContactMap - Whittaker et al., 2004, automatic grouping of contacts, or identification of social behaviours (e.g. SNARF - Fisher et al., 2007). The results presented in this thesis suggest that the most promising avenue of research is the approach of semi-automation of group identification (e.g. Amershi et al., 2012). By identifying possibilities for groups based on the forms of groups people perceive, such as those identified in the contact grouping framework (Chapter 5), which can be used as building blocks to create *ad hoc* groups, it affords greater flexibility for the user. This has the advantage that it mitigates the inability of automatic detection to produce accurate groupings (MacLean et al., 2011) and the tendency of people to create compromise groups when manually creating groupings (Chapter 7). The impermanent nature of the groupings means that it also avoids the issue of group evolution over time (Fisher and Dourish, 2004; Bartel and Dewan, 2013) and the ability to adapt groups to context specific situations. This should in turn provide greater confidence in the groupings created, which was identified as an important component for using groups to manage information. Only when participants were absolutely confident that they knew who was in each group would they perform group level actions. Without this confidence every individual member of a group had to be checked. This need for confidence is an issue for systems that want to operate with interactions on group levels, such as privacy controls in social networks (Mazzia et al., 2012; Kelley et al., 2011; Jones and O'Neill, 2010). Without absolute confidence in groups people will be less inclined to use them or will use them poorly due to inclusion or exclusion of the wrong contacts.

9.4 Limitations

This research has sought to understand how to design interfaces that are focused on people or contacts as a means of supporting personal information tasks. One of the reasons for the focus of this research has been that links between information and people have been utilised to a lesser extent than other similar primary attributes, such as those of source, subject, and time (see Chapter 4) for supporting personal information tasks. However, it is not always the case that interfaces are required to focus extensively on one attribute in this manner, it is common for interfaces to provide a range of means of viewing or interacting with information in conjunction with each other, to use as appropriate. In the information re-finding study (Chapter 4), for example, participants used a range of information to support their navigation to the information they were looking for. Most obviously interfaces for file browsers often make use of all of these attributes, without giving priority to a particular attribute to describe and order files. How the use of links between people and information might work best in conjunction with these other attributes was beyond the scope of this research, and was not explored. However, understanding how using links between people and information in conjunction with these other attributes would equally make a valuable contribution towards supporting personal information tasks. Even in 'person-centric' interfaces it would often not be appropriate to use this concept in isolation of other existing concepts for understanding and interacting with information.

Further to this, while this thesis focussed on generating realistic tasks in its evaluation of the prototype system, by eliciting from participants items within their own personal information to re-find or share, the tasks set were still artificial in nature. The task elicitation allowed the process that participants went through in performing the tasks to be as natural as possible in a lab based study, but this can not fully account for how participants might behave were the tasks performed naturally as they occurred and over a longer period of time. While this work establishes a foundation of research into understanding how the links made between people and information can be used in interface design that supports people in performing information tasks, further research needs to be done to

understand how the design implications of this research fit with natural everyday behaviours. Next steps must be to take the conclusions and design concepts arrived at in this largely lab-based context, and further explore how they fit to 'natural' information behaviours. For example, would use of interface features such as those described in this thesis, if implemented, be adopted by users of software that integrated them in the long term, or would they be largely ignored as new features can often be (Jones et al., 2004)? And would the processes described here be undertaken differently alongside other means of performing the same task?

Another limitation of this thesis is that while the sample sizes used in these studies were appropriate to the research being performed, it was not logistically feasible to extend this scope to large sample sizes. In the contact grouping study (Chapter 5) and the contact grouping prototype (Chapter 7), it is clear to see what forms of groups are most common, but larger sample sizes over a more balanced range of demographics could reveal more representative information about the forms of groups being generated, and in particular would clarify the prevalence of some of the less common types of groups. Similarly, in the information re-finding study (Chapter 4) the categorisation of information re-finding need types was unable to describe, to the same degree of detail several of the types of tasks being performed, such as the need to find information personally created, as they were less common and not enough data was available to fully break them down into more specific and insightful task categories

Two factors which may have influenced the representativeness of the results presented in this thesis were the samples of participants used and the variability of participant enthusiasm during their participation. While a variety of demographics were represented by the participants in the studies in this thesis, there was a heavy bias in particular towards university researchers and PhD students. While this might have altered the emphasis in terms of the occurrence of certain results, this was not considered to be detrimental to the aims of this thesis as it was concerned with understanding the processes that underpin the behaviours in the use of 'person' to perform information tasks, the understanding of which should not obviously have been affected by these instances.

The other factor of participant enthusiasm was also significant in both the re-finding study (Chapter 4) and the group construction part of the prototype evaluation study (Chapter 7). In the former, participant enthusiasm meant that the diligence of participants in recording re-finding events was variable and the results did not represent a complete record of every participant's re-finding behaviours. In the latter, grouping over each of the conditions was a time consuming and repetitive process, which some participants completed with more enthusiasm than others, resulting in greater or lesser numbers of groups and levels of specificity to be created by some participants. This variability does not affect the conclusions drawn from the research, but rather highlights that there is scope for further research to be done with larger samples and observations in more natural environments.

9.5 Future Work

Much of the research described in this thesis has been necessarily exploratory to some degree due to a relative lack of prior research in this area of design, despite the salient relationships 'person' has to information. The focus of this work has therefore been to understand the processes and relationships that might underpin the use of a person-centric interface. The first two studies provided an in-depth exploration of a specific information task and the representation of contacts. The final study implemented design ideas based on findings from the prior studies, where a much of the value lay in exploring the responses to those ideas, understanding the boundaries of person-centric interface design, and why these features did or did not serve a particular purpose. It also sought to draw out the wider implications of those responses. Previous research has focused only on the role that interacting with people can play in the process of performing a task. This research considers both the process of the tasks being performed, and how the ways in which they are represented can affect that process. In future work this could be extended by focusing further on exploring design concepts, including or expanding on those identified in this research, to understand more about how the design of 'person-centric' interfaces can influence and support information tasks, and to further define and establish guidelines for their design.

Looking at the process of contact grouping by creating another prototype interface, but in this instance focused purely on grouping and integrated with a email client or social network, such as Facebook, could provide a large amount of data on grouping behaviours. By integrating participation in the study with an application already used by the participant the study could be conducted in the participants' own time in a less obtrusive manner – providing similar benefits to the mobile diary study in Chapter 4 – lowering the barriers to participation and using real world data over time in a context that is clear to the participants. Generating a larger and more focused data set in this manner could serve to further extend or validate the existence of some of the behaviours described in this thesis, such as the egocentric grouping interface design leading to an increase in group creation.

Another obvious option for further study in this area would be to again create another prototype, but rather than increase the amount of participants, increase the length of time participants are studied. As in Sawyer et al. (2012), this new prototype would attempt to integrate with or replace the tool that participants use as part of their current practices for a particular task. Data could be captured through a diary study, similar to Chapter 4, or through direct observations. This would serve to demonstrate the utility of some of the person-centric design ideas discussed here in practical rather than hypothetical situations, providing a rich data set through which to analyse the information behaviours of participants and their use of a person-centric design.

References

Abrams, D., R. Baecker, and M. Chignell

1998. Information archiving with bookmarks: personal web space construction and organization. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, Pp. 41–48. ACM Press/Addison-Wesley Publishing Co.

Adamic, L. A., D. Lauterbach, C. Teng, and M. S. Ackerman

2011. Rating friends without making enemies. In *ICWSM*. Citeseer.

Amershi, S., J. Fogarty, and D. Weld

2012. Regroup: Interactive machine learning for on-demand group creation in social networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

André, P., M. L. Wilson, A. Russell, D. A. Smith, A. Owens, et al.

2007. Continuum: designing timelines for hierarchies, relationships and scale. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, Pp. 101–110. ACM.

Antonucci, T. C.

1986. Hierarchical mapping technique. *Generations: Journal of the American Society on Aging*.

Bakshy, E., I. Rosenn, C. Marlow, and L. Adamic

2012. The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web*, Pp. 519–528. ACM.

Ballagas, R. and S. P. Walz

2008. Gaming Tourism : Lessons from Evaluating REXplorer , a Pervasive Game for Tourists. *Pervasive computing. Lecture Notes in Computer Science*, Pp. 244–261.

Barreau, D. and B. A. Nardi

1995. Finding and reminding: file organization from the desktop. *ACM SigChi Bulletin*, Pp. 39–43.

Bartel, J. W. and P. Dewan

2013. Evolving friend lists in social networks. In *Proceedings of the 7th ACM Conference on Recommender Systems*, Pp. 435–438. ACM.

Belkin, N. J. et al.

1993. Interaction with texts: Information retrieval as information seeking behavior. *Information retrieval*, 93:55–66.

Bergman, O., R. Beyth-Marom, R. Nachmias, N. Gradovitch, and S. Whittaker

2008. Improved search engines and navigation preference in personal information management. *ACM Transactions on Information Systems (TOIS)*, 26(4):20.

Beyer, H. and K. Holtzblatt

1997. *Contextual design: defining customer-centered systems*. Elsevier.

Blanc-Brude, T. and D. L. Scapin

2007. What do people recall about their documents?: implications for desktop search tools. In *Proceedings of the 12th international conference on Intelligent user interfaces*, Pp. 102–111. ACM.

Boardman, R. and M. A. Sasse

2004. Stuff goes into the computer and doesn't come out: a cross-tool study of personal information management. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, Pp. 583–590. ACM.

Boren, T. and J. Ramey

2000. Thinking aloud: Reconciling theory and practice. *IEEE transactions on professional communication*, 43(3):261–278.

Braun, V. and V. Clarke

2006. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101.

Broder, A.

2002. A taxonomy of web search. In *ACM Sigir forum*, volume 36, Pp. 3–10. ACM.

Brown, B. A., A. J. Sellen, and K. P. O'hara

2000. A diary study of information capture in working life. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, Pp. 438–445. ACM.

Capra, R., M. Pinney, and M. A. Perez-Quinones

2005. Refinding is not finding again.

Capra, R. G. and M. A. Pérez-Quiñones

2003. Re-finding found things: An exploratory study of how users re-find information. *arXiv preprint cs/0310011*.

Card, S. K.

1996. Visualizing retrieved information: A survey. *IEEE computer graphics and applications*, 16(2):63–67.

Carpenter, B. E.

2009. Observed relationships between size measures of the internet. *ACM SIGCOMM Computer Communication Review*, 39(2):5–12.

Church, K. and B. Smyth

2009. Understanding the intent behind mobile information needs. In *Proceedings of the 14th international conference on Intelligent user interfaces*, Pp. 247–256. ACM.

Consolvo, S. and M. Walker

2003. Using the experience sampling method to evaluate ubicomp applications. *IEEE Pervasive Computing*, 2(2):24–31.

Czerwinski, M., E. Horvitz, and S. Wilhite

2004. A diary study of task switching and interruptions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, Pp. 175–182. ACM.

Dawson, A., D. Brown, and V. Broughton

2006. The need for a faceted classification as the basis of all methods of information retrieval. In *Aslib proceedings*, volume 58, Pp. 49–72. Emerald Group Publishing Limited.

Dearman, D., M. Kellar, and K. N. Truong

2008. An examination of daily information needs and sharing opportunities. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, Pp. 679–688. ACM.

Donath, J.

2002. A Semantic Approach to Visualizing Online Conversations. *Communications of the ACM*, 45(4):45–49.

Donath, J., A. Dragulescu, A. Zinman, F. B. Viégas, R. Xiong, and Y. Assogba

2010. Data portraits. In *SIGGRAPH 2010*, Pp. 1–8.

Dumais, S., E. Cutrell, and J. Cadiz

2003. Stuff I've seen: a system for personal information retrieval and re-use. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval.*, Pp. 72–79.

Dunbar, R. I.

2016. Do online social media cut through the constraints that limit the size of offline social networks? *Royal Society Open Science*, 3(1):150292.

Ekmekcioglu, F. C., M. F. Lynch, and P. Willett

1996. Stemming and n-gram matching for term conflation in turkish texts. *Information Research News*, 7(1):2–6.

Elsweiler, D. and I. Ruthven

2007. Towards task-based personal information management evaluations. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, Pp. 23–30. ACM.

Ericsson, K. and H. Simon

1993. *Protocol analysis*. MIT Press Cambridge, Mass.

Fisher, D., A. Brush, B. Hogan, M. Smith, and A. Jacobs

2007. Using social metadata in email triage: Lessons from the field. In *Symposium on Human Interface and the Management of Information*, Pp. 13–22. Springer.

Fisher, D. and P. Dourish

2004. Social and temporal structures in everyday collaboration. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, Pp. 551–558. ACM.

Fisher, D. A.

2004. *Social and Temporal Structures in Everyday Collaboration*. PhD thesis, UNIVERSITY OF CALIFORNIA, IRVINE.

Fransella, F. and D. Bannister

2004. *A manual for repertory grid technique*. Academic Press, London, New York, SF.

Fruchterman, T. M. and E. M. Reingold

1991. Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164.

Gemmell, J., G. Bell, R. Lueder, S. Drucker, and C. Wong

2002. Mylifebits: fulfilling the memex vision. In *Proceedings of the tenth ACM international conference on Multimedia*, Pp. 235–238. ACM.

Gruen, D., S. L. Rohall, S. Minassian, B. Kerr, P. Moody, B. Stachel, M. Wattenberg, and E. Wilcox

2004. Lessons from the remail prototypes. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, Pp. 152–161, New York, NY, USA. ACM.

Hailpern, J., N. Jitkoff, A. Warr, K. Karahalios, R. Sesek, and N. Shkrob

2011. Youpivot: improving recall with contextual search. In *Proceedings of the*

SIGCHI Conference on Human Factors in Computing Systems, Pp. 1521–1530. ACM.

Hart, S. G., M. F. California, and L. E. Staveland

1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research Advances in psychology. *Advances in psychology*, (52):139–183.

Heer, J. and D. Boyd

2005. Vizster: Visualizing online social networks. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*, Pp. 32–39. IEEE.

Jaimes, A., K. Omura, T. Nagamine, and K. Hirata

2004. Memory Cues for Meeting Video Retrieval. *CARPE'04 Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, Pp. 74–85.

Jones, S. and E. O'Neill

2010. Feasibility of structural network clustering for group-based privacy control in social networks. *Proceedings of the Sixth Symposium on Usable Privacy and Security - SOUPS '10*, P. 1.

Jones, W.

2008. How is information personal. *Personal Information Management: PIM*, Pp. 5–6.

Jones, W. and K. M. Anderson

2011. Many views, many modes, many tools... one structure: Towards a non-disruptive integration of personal information. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*, Pp. 113–122. ACM.

Jones, W., A. J. Phuwanartnurak, R. Gill, and H. Bruce

2004. Don't Take My Folders Away! Organizing Personal Information to Get Things Done. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, Pp. 1–4.

Katifori, A., C. Vassilakis, and A. Dix

2008. Using spreading activation through ontologies to support personal infor-

mation management. In *Proceedings of the 2008 International Conference on Common Sense Knowledge and Goal-Oriented Interfaces-Volume 323*, Pp. 36–45. CEUR-WS. org.

Kellar, M., C. Watters, and M. Shepherd

2007. A Field Study Characterizing Web-Based Information- Seeking Tasks. *Journal of the American Society for Information Science and Technology*, 58(7):999–1018.

Kelley, P. G., R. Brewer, Y. Mayer, L. F. Cranor, and N. Sadeh

2011. An investigation into facebook friend grouping. In *IFIP Conference on Human-Computer Interaction*, Pp. 216–233. Springer.

Kelly, D.

2006. Evaluating personal information management behaviors and tools. *Communications of the ACM*, 49(1):84–86.

Khan, R., S. U. Khan, R. Zaheer, and S. Khan

2012. Future internet: the internet of things architecture, possible applications and key challenges. In *Frontiers of Information Technology (FIT), 2012 10th International Conference on*, Pp. 257–260. IEEE.

Kim, J. Y., M. Cramer, J. Teevan, and D. Lagun

2013. Understanding how people interact with web search results that change in real-time using implicit feedback. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*, Pp. 2321–2326. ACM.

Klein, G. A., R. Calderwood, and D. Macgregor

1989. Critical decision method for eliciting knowledge. *IEEE Transactions on systems, man, and cybernetics*, 19(3):462–472.

Kljun, M., J. Mariani, and A. Dix

2013. Transference of pim research prototype concepts to the mainstream: successes or failures. *Interacting with Computers*, 27(2):73–98.

Kossek, E. E., M. N. Ruderman, P. W. Braddy, and K. M. Hannum

2012. Work–nonwork boundary management profiles: A person-centered approach. *Journal of Vocational Behavior*, 81(1):112–128.

Krishnan, A. and S. Jones

2005. Timespace: activity-based temporal visualisation of personal information spaces. *Personal and Ubiquitous Computing*, 9(1):46–65.

Kules, B., M. L. Wilson, M. Schraefel, and B. Shneiderman

2010. From keyword search to exploration: Designing future search interfaces for the web. *Foundations and Trends® in Web Science*, 2(1):1–97.

Lenhart, A., K. Purcell, A. Smith, and K. Zickuhr

2010. Social media & mobile internet use among teens and young adults. millennials. *Pew internet & American life project*.

Leskovec, J. and J. J. Mcauley

2012. Learning to discover social circles in ego networks. In *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds., Pp. 539–547. Curran Associates, Inc.

Liu, T.-Y., Y. Yang, H. Wan, Q. Zhou, B. Gao, H.-J. Zeng, Z. Chen, and W.-Y. Ma

2005. An experimental study on large-scale web categorization. In *Special interest tracks and posters of the 14th international conference on World Wide Web*, Pp. 1106–1107. ACM.

Lubbers, M. J., J. L. Molina, and C. McCarty

2007. Personal networks and ethnic identifications. *International Sociology*, 22(6):721–741.

MacLean, D., S. Hangal, S. K. Teh, M. S. Lam, and J. Heer

2011. Groups without tears: mining social topologies from email. In *Proceedings of the 16th international conference on Intelligent user interfaces*, Pp. 83–92. ACM.

Malone, T. W.

1983. How Do People Organize Their Desks ? Implications for the Design

of Office Information Systems. *ACM Transactions on Information Systems (TOIS)*, 1(1):99–112.

Marinagi, C., C. Skourlas, and P. Belsis

2013. Employing ubiquitous computing devices and technologies in the higher education classroom of the future. *Procedia-Social and Behavioral Sciences*, 73:487–494.

Mazzia, A., K. LeFevre, and E. Adar

2012. The pviz comprehension tool for social network privacy settings. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, P. 13. ACM.

McCallum, A., A. Corrada-Emmanuel, and X. Wang

2005. Topic and role discovery in social networks. In *Ijcai*, volume 5, Pp. 786–791. Citeseer.

McCarty, C., J. L. Molina, C. Aguilar, and L. Rota

2007. A Comparison of Social Network Mapping and Personal Network Visualization. *Field Methods*, 19(2):145–162.

McDonald, D. W.

2003. Recommending collaboration with social networks: a comparative evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, Pp. 593–600. ACM.

Muller, M. J.

2003. Participatory design: the third space in hci. *Human-computer interaction: Development process*, 4235:165–185.

Neustaedter, C., A. J. Brush, M. A. Smith, and D. Fisher

2005. The Social Network and Relationship Finder: Social Sorting for Email Triage. In *Proceedings of the Fourth Conference on Email and Anti-Spam (CEAS 2005)*.

Newman, A.

2008. *Protectors of privacy: Regulating personal data in the global economy*. Cornell University Press.

Nicholas, A., J. H. F. The, and C. Link

2007. The Spread of Obesity in a Large Social Network Over 32 Years. *New England journal of medicine*, 4(357):370–379.

Offenhuber, D. and J. Donath

2007. Comment flow: visualizing communication along network path. *Poster presented at IEEE InfoVis*, 7.

Palen, L. and M. Salzman

2002. Voice-mail diary studies for naturalistic data capture under mobile conditions. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, CSCW '02, Pp. 87–95, New York, NY, USA. ACM.

Papachristos, A. V.

2011. The coming of a networked criminology. In *Measuring Crime & Criminality: Advances in Criminological Theory*, Pp. 101–140.

Preece, J., Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey

1994. *Human-Computer Interaction*. Essex, UK, UK: Addison-Wesley Longman Ltd.

Quan, D., D. Huynh, and D. R. Karger

2003. Haystack: A platform for authoring end user semantic web applications. In *International Semantic Web Conference*, Pp. 738–753. Springer.

Rekimoto, J.

1999. Time-machine computing: A time-centric approach for the information environment. In *Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology*, UIST '99, Pp. 45–54, New York, NY, USA. ACM.

Reyes, C.

2016. Eliciting data on social relationships: The use of hand-drawn network maps in tracing the perception of digitally mediated social ties. *International Review of Social Research*, 6(4):256–268.

Rieman, J.

1993. The diary study: a workplace-oriented research tool to guide laboratory

efforts. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, Pp. 321–326. ACM.

Ringel, M., E. Cutrell, S. Dumais, E. Horvitz, et al.
2003. Milestones in time: The value of landmarks in retrieving information from personal stores. In *Proc. Interact*, volume 2003, Pp. 184–191.

Risden, K., M. P. Czerwinski, T. Munzner, and D. B. Cook
2000. An initial examination of ease of use for 2d and 3d information visualizations of web content. *International Journal of Human-Computer Studies*, 53(5):695–714.

Rivadeneira, W. and B. B. Bederson
2003. A study of search result clustering interfaces: Comparing textual and zoomable user interfaces. *HCIL*, 21(5).

Sauermann, L., A. Bernardi, and A. Dengel
2005. Overview and outlook on the semantic desktop. In *Proceedings of the 2005 International Conference on Semantic Desktop Workshop: Next Generation Information Management D Collaboration Infrastructure-Volume 175*, Pp. 74–91. CEUR-WS. org.

Sawyer, B., F. Quek, W. C. Wong, M. Motani, S. L. C. Yew, and M. Pérez-Quiñones
2012. Information re-finding through physical-social contexts. In *Workshop of CSCW 2012 on PIM*.

Sluka, J. A.
1996. Peace process images, symbols and murals in Northern Ireland. *Critique of Anthropology*, 16(4):381–387.

Sohn, T., K. A. Li, W. G. Griswold, and J. D. Hollan
2008. A diary study of mobile information needs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Pp. 433–442. ACM.

Stanton, N., P. M. Salmon, and L. A. Rafferty
2013. *Human factors methods: a practical guide for engineering and design*. Ashgate Publishing, Ltd.

Sun, X., S. Sharples, and G. Kefalidou

2012. User requirements for an interactive system to stimulate serendipity. *Contemporary Ergonomics and Human Factors 2012*, Pp. 369–376.

Tamler, H.

1998. How (much) to intervene in a usability testing session. *Common Ground*, 8(3):11–15.

Teevan, J., E. Adar, R. Jones, and M. A. Potts

2007. Information re-retrieval: repeat queries in yahoo's logs. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, Pp. 151–158. ACM.

Teevan, J., C. Alvarado, M. S. Ackerman, and D. R. Karger

2004. The perfect search engine is not enough: a study of orienteering behavior in directed search. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, Pp. 415–422. ACM.

Tulving, E.

2002. Episodic memory: from mind to brain. *Annual review of psychology*, 53(1):1–25.

Tyler, S. K. and J. Teevan

2010. Large Scale Query Log Analysis of Re-Finding. *Proceedings of the third ACM international conference on Web search and data mining*, Pp. 191–200.

Ugander, J., B. Karrer, L. Backstrom, and C. Marlow

2011. The anatomy of the facebook social graph. *arXiv preprint arXiv:1111.4503*.

Viégas, F., M. Wattenberg, J. Hebert, G. Borggaard, A. Cichowlas, J. Feinberg, J. Orwant, and C. R. Wren

2013. Google + Ripples : A Native Visualization of Information Flow. In *Proceedings of the 22nd international conference on World Wide Web. International World Wide Web Conferences Steering Committee*, Pp. 1389–1398.

Viégas, F. B., S. Golder, and J. Donath

2006. Visualizing email content: portraying relationships from conversational

histories. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, Pp. 979–988. ACM.

Viégas, F. B., D. H. Nguyen, J. Potter, J. Donath, S. Hall, and G. Tech
2004. Digital Artifacts for Remembering and Storytelling : PostHistory and Social Network Fragments. *Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, 00(C):1–10.

Vogl, S., U. Zartler, E.-M. Schmidt, and I. Rieder
2018. Developing an analytical framework for multiple perspective, qualitative longitudinal interviews (mpqli). *International Journal of Social Research Methodology*, 21(2):177–190.

Whittaker, S., Q. Jones, B. Nardi, and M. Creech
2004. ContactMap: Organizing communication in a social desktop. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 11(4):445–471.

Whittaker, S. and C. Sidner
1996. Email overload: exploring personal information management of email. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, Pp. 276–283. ACM.

Wilson, T.
1997. Information behaviour: An interdisciplinary perspective. *Information Processing & Management*, 33(4):551–572.

Woods, D. K. and P. G. Dempster
2011. Tales from the bleeding edge: The qualitative analysis of complex video data using transana. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, volume 12.

Yukun, L., D. Elsweiler, and X. Meng
2010. Towards Task-Organised Desktop Collections. In *Proceedings of the 1st Desktop Search Workshop. 2010.*, Pp. 23–26.

Zubiaga, A., V. Fresno, R. Martinez, and A. P. Garcia-Plaza
2013. Harnessing folksonomies to produce a social classification of resources. *IEEE transactions on knowledge and data engineering*, 25(8):1801–1813.

Appendix A – Information Re-finding Diary Data Examples

Diary Entries and Extracts of Corresponding Interview

Entry 1

Domain: Work Difficulty: **Easy Recency:** Week **Success:** Yes

Diary Entry Need: what date am I going to newport

Participant C: That was in relation to my internship...

Diary Entry Prior Knowledge: that it was in October and the email was sent by mark

Participant C: ...I knew the person who was organising it was called Mark [Surname].

Diary Entry Find: I searched for marks name in outlook and fortunately it was the last email he sent

Participant C: So what I did was, I errm searched for his name in the email client, and I was scanning through for the word October, because I knew it was in October as well. I searched for Mark's name in Outlook, fortunately it was

the last email he'd sent. So it was information I'd already seen and we were in the midst of trying to organise it, and someone had wanted to know exactly what date I was going.

Entry 2

Domain: Personal **Difficulty:** Neutral **Recency:** Month **Success:** Yes

Diary Entry Need: wanted to show my mum an outfit that an actress as wearing in a video

Participant C: I wanted to show my mum and outfit an actress had. I presume it was something I had seen somewhere else, and it had come up in conversation with my mum, and I was showing her the same Youtube video.

Diary Entry Prior Knowledge: it was in a YouTube video I'd seen before :)

Participant C: It was something I'd used on that computer,

Diary Entry Find: I knew it was in my YouTube history and searched through the videos until I saw the actress

Participant C: . . . so I just looked at the last searched. . . so you know if you log into Youtube, it records your history, so that is what I searched through to get to the thing. It was through browsing the history.

Interviewer: Did you know the subject when you were browsing?

C2: I don't really look at that many youtube videos so it was quite obvious.

Entry 3

Domain: Work **Difficulty:** Easy **Recency:** Week **Success:** Yes

Diary Entry Need: looking for a section in a paper

Participant C: So this seems like it was looking for a section in a paper,. . .

Diary Entry Prior Knowledge: that it was in my bookmarks on chrome

Participant C: ...and that it was in my bookmarks on Chrome. I think it was actually a book, this was a visual analytics book that Mark [Surname] had sent me. Visual analytics covers five main topics like human factors, data mining blahdy blahdy blah. And he'd sent it to me to read the human factors section.

Diary Entry Find: went on to chrome, then bookmarks and looked through the list

Participant C: So I'd saved it for later if that makes sense, so I'd bookmarked it, he'd sent it me as a book. So I hadn't read the actual section that I had saved for later if that makes sense.

Entry 4

Domain: Work **Difficulty:** Easy **Recency:** Week **Success:** Yes

Diary Entry Need: to get a journal paper by Steve benford

Participant C: I was writing a paper. This paper won best award at CHI, so I was using his paper as a reference of how to write an introduction.

Diary Entry Prior Knowledge: the words Steve vendors and 2012

Participant C: And I used the words Steve Benford and I knew it was 2012. It was something I'd read before because it was best paper, and I think it was sent around before as well.

Diary Entry Find: typed the words in scholar

Participant C: So it is almost trying to get ideas for it. But I was looking for it on Google Scholar, and the words I used were Steve Benford 2012.

Interviewer: Did you remember anything else about the paper?

Participant C: I remember that it was about discomfort and that it had won best CHI paper, so if I hadn't found it from Google Scholar straight away then that would have been my next port of call. The main thing was that it was Steve Benford's paper.

Entry 5

Domain: Personal **Difficulty:** Easy **Recency:** Year **Success:** Yes

Diary Entry Need: when matts birthday is

Participant C: When Matt's birthday is.

Diary Entry Prior Knowledge: it was in the month of october

Participant C: Because I knew it was in October but I thought it would be on Facebook, but it wasn't there.

Diary Entry Find: I looked on Facebook but wasn't there, then I messaged him and double checked

Participant C: I knew it was in the month of October, I looked on Facebook, but it wasn't there, so I messaged him. The year before he'd put it on Facebook, errrm and done like an invite and everything, so I presumed it would be on Facebook. But I don't know if it was that he'd changed the settings or he'd invited everybody.

Entry 7

Domain: Work **Difficulty:** Easy **Recency:** Week **Success:** Yes

Diary Entry Need: when the PhD meetings are? *Participant C:* Human Factors meeting, when the PhD meetings are. I think there was some change in the times or the dates recently.

Diary Entry Prior Knowledge: that its in the diary

Participant C: I knew they were in... basically they were in one time then they were changed to another time or a different Monday of the month. So that's what I was checking for. I'd seen them previously, in my calendar, seen the emails, accepted them.

Diary Entry Find: looked in calendar

Participant C:

Entry 8

Domain: Personal **Difficulty:** Difficult **Recency:** Week **Success:** Yes

Diary Entry Need: music video *Participant C:* I'd watched this video quite a number of times. My friends had come up at the weekend and they had. You know the Gangnam style video? There are parodies of them, and we'd been searching for them.

Diary Entry Prior Knowledge: phrase gangnam

Participant C: They had shown me this navy one, and I think I'd searched for "navy Gangnam style", I hadn't searched it before, but I knew it was on Youtube. If that makes sense.

Diary Entry Find: navy hangman style video, comedy value

Participant C: So I Googled "navy Gangnam style video comedy", so it was something somebody had shown me before, so I hadn't searched for it before. I think someone had posted it on Facebook as well and I tried to look through and find it.

Entry 9

Domain: Work **Difficulty:** Neutral **Recency:** Month **Success:** Yes

Diary Entry Need: information about cdm techniques *Participant C:* So for the CDM papers [Supervisor's Name] had suggested that I speak to [Colleagues Name] who has done a lot of work and done lots about it.

Diary Entry Prior Knowledge: the persons name who sent it

Participant C: And she sent me two papers by email and errrm... I searched for her name, but I didn't look at them straightaway because went away for my internship, and then went back to look at them while I was there.

Diary Entry Find: email search

Participant C: So it had been a month or so by the time I looked for them It was just an email search for her name.

Entry 10

Domain: Personal **Difficulty:** Difficult **Recency:** Few Months **Success:** Yes

Diary Entry Need: Photoshop tutorial on how to ink *Participant C:* I'd looked through a few Photoshop tutorials. I'm trying to teach myself how to ink using a graphics tablet. And errrm, I'd saved a few videos for later.

Diary Entry Prior Knowledge: it had been sent to me on facebook

Participant C:

Diary Entry Find: searching through the history of a conversation on Facebook

Participant C: They had been sent to me on Facebook, so I searched through the history to find the video. The conversation between... we'd had a conversation about it and it had come up. We'd spoke about it because that person is also interested in the same thing, we'd talked about inking, and they'd errm sent me the link.

Appendix B – Extracts of Code from Prototype Development

Extract of Code from Person-Centric Interface

The code listed below represents a small extract of the JavaScript code that generated the visualisation and defined the interaction with the person-centric interface prototype. This code works alongside the HTML code (not shown) to define how the visualisation behaves. It begins by receiving a JSON file containing the contacts' data and the results of the group construction exercises. Much of this code is concerned with setting up the variables that define how the visualisations were structured, it's interaction, and inputting the content from the JSON file.

```
1
2  d3.json("static/PersonCentric.json", function(error, data) {
3
4  var overlay;
5  var triggerBttn;
6  setupOverlay()
7
8  $('#message-content').hide();
9  var messageList = data.messages;
10
11 var minDate = new Date();
12 var maxDate = new Date();
13
14 var locationFilters = [],
```

```
15 affiliationFilters = [],
16 relationshipFilters = [];
17
18 var activeFilters = {};
19
20 var gravity = 0.08;
21 var charge = -700;
22 var linkDistance = 190;
23 var linkStrength = 0.5;
24
25 var defaultNodeSize = 34;
26 var padding = 20; // separation between same-color circles
27 var clusterPadding = 150; // separation between different-color
       circles
28 var maxRadius = 12;
29
30 activeFilters["Location"] = [];
31 activeFilters["Affiliation"] = [];
32 activeFilters["Relationship"] = [];
33
34 console.dir(data);
35
36 initialiseFilter("Location", data.locations, locationFilters);
37 initialiseFilter("Affiliation", data.affiliations,
       affiliationFilters)
38 initialiseFilter("Relationship", data.relationships,
       relationshipFilters)
39
40 // For each person in the data set check for unique locations and
        store them later to use as filters, and append them to the
       filter menu.
41 $.each(data.people, function (index, value) {
42
43 // Identify minimum date
44 if  (new Date(value.Recency) < minDate) {
45
46 minDate = new Date(value.Recency);
47
48 }
49
```

```
50 // Identify maximum date
51 if  (new Date(value.Recency) > maxDate) {
52
53 maxDate = new Date(value.Recency);
54
55 }
56
57 });
58
59 $(document).ajaxSend(function(event, request, settings) {
60 $('#loading-indicator').show();
61 });
62
63 $(document).ajaxComplete(function(event, request, settings) {
64 $('#loading-indicator').hide();
65 });
66
67 $("#gravity-slider").slider({
68 orientation: "vertical",
69 min: 0.0,
70 max: 0.15,
71 step: 0.01,
72 value: gravity
73 });
74
75 $("#charge-slider").slider({
76 orientation: "vertical",
77 min: -1,
78 max: -3000,
79 step: 50,
80 value: -500
81 });
82
83 $("#distance-slider").slider({
84 orientation: "vertical",
85 min: 0,
86 max: 400,
87 step: 5,
88 value: linkDistance
89 });
```

```
90
91  $("#link-strength-slider").slider({
92  orientation: "vertical",
93  min: 0,
94  max: 1,
95  step: 0.01,
96  value: linkStrength
97  });
98
99  $("#recencyslider").dateRangeSlider({
100 bounds: {min: minDate, max: maxDate},
101 defaultValues:{min: minDate, max: maxDate}
102 });
103
104 $("#closenessslider").rangeSlider({
105 bounds: {min: 1, max: 10},
106 defaultValues:{min: 1, max: 10}
107 });
108
109 // Colours
110 var steelblue = d3.rgb( 70, 130, 180);
111 var backgroundBlue = d3.rgb(215, 231, 245);
112 var groupBlue = d3.rgb(173, 186, 245);
113 var groupPurple = d3.rgb(235, 215, 245);
114 var groupRed = d3.rgb(255, 135, 135);
115 var groupGreen = d3.rgb(215, 245, 226);
116 var groupYellow = d3.rgb(255, 255, 173);
117 var groupPink = d3.rgb(255, 179, 218);
118 var groupOrange = d3.rgb(255, 201, 140);
119 var groupTurquoise = d3.rgb(173, 251, 255);
120 var groupGrey = d3.rgb(230, 225, 225);
121
122 var groupColourList = [groupBlue, groupPurple, groupRed,
       groupGreen, groupYellow, groupPink, groupOrange,
       groupTurquoise, groupGrey];
123
124 var currentDataNodes = [],
125 currentDataLinks = [],
126 tempDataLinks = [],
127 currentGroupLinks = [];
```

```
128
129 var currentGroupList = [];
130
131 var removedDataNodes = [],
132 removedDataLinks = [],
133 removedGroupLinks = [],
134 removedGroups =[],
135 emptyGroups = [],
136 filteredGroups = [];
137
138 var dimensionText;
139
140 var allPrimaryLinks = [];
141
142 var selectedNode;
143 var selectedList = [];
144
145 var currentLayout = "None";
146 var currentStructure = "Primary";
147
148 var width = 1000,
149 height = 900;
150
151 var zoom = d3.behavior.zoom().on("zoom", redraw)
152
153 var svg = d3.select("#visualisation")
154 .append("svg")
155 .attr("width", width)
156 .attr("height", height)
157 .attr("pointer-events", "all")
158 .append('svg:g')
159 .call(zoom)
160 .append('svg:g');
161
162 svg.append('svg:rect')
163 .attr('x', -1000)
164 .attr('y', -700)
165 .attr('width', width * 3)
166 .attr('height', height * 3)
167 .attr('fill', 'white');
```

```
168
169 currentDataNodes = data.people;
170 currentDataLinks = data.links;
171 currentGroupLinks = data.groupsLinks;
172
173 var force = d3.layout.force();
174 var linkStyle;
175 var nodeStyle;
176
177 var nodes;
178
179 $.each(currentGroupLinks, function (index, value) {
180
181 allPrimaryLinks.push(value);
182
183 });
184
185 currentLayout = "group";
186
187 currentDataNodes = currentDataNodes.concat(data.groups);
188 console.dir(currentDataNodes);
189 console.dir(currentGroupLinks);
190 associateGroups(currentDataNodes, currentGroupLinks);
191
192 nodes = d3.values(currentDataNodes);
193
194 force.nodes(currentDataNodes)
195 .links(currentGroupLinks).start();
196
197 runGroupLayout(currentDataNodes, currentGroupLinks);
198
199
200 function start() {
201
202 linkStyle = svg.selectAll(".link");
203 nodeStyle = svg.selectAll(".node");
204
205 linkStyle = linkStyle.data(force.links(), function(d) { return d.
    source.ID + "-" + d.target.ID; });
206 linkStyle.enter().insert("line", ".node").attr("class", "link");
```

```
207 linkStyle.exit().remove();
208
209 nodeStyle = nodeStyle.data(force.nodes(), function(d) { return d.
    ID;});
210 nodeStyle.attr("r", function(d) { return d.radius; });
211 var nodeAppend = nodeStyle.enter().append("g")
212 .attr("class", function(d) { return "node " + d.ID + " " + d.type
    ; })
213 .call(force.drag)
214 .on("mouseover", function(d) {
215 svg.selectAll(".outer-circle").transition()
216 .attr("r", function(d, i) { return d.radius; });
217
218 // add the nodes
219 nodeAppend.append("circle")
220 .attr("r", function(d, i) { return d.radius; }) // Node Size
221 .attr("class", "outer-circle")
222 .style("fill", function(d, i) { if (d.firstName == "ME") { return
     "orange" } else if ((d.type == "group") && (currentLayout ==
    "link")) { return "red"; }else if ( d.type == "group") {
    return defaultNodeColour(d) }  else { return defaultNodeColour
    (d);} })
223 .style("stroke", function(d, i) {
224 var currentColour = defaultNodeColour(d);
225 currentColour.darker(1);
226 return currentColour;
227 });
228
229 svg.selectAll(".outer-circle").transition()
230 .attr("r", function(d, i) { return d.radius; });
231
232 nodeAppend.append("clipPath")
233 .attr('id', function(d) { return "clip" + d.index })
234 .append('circle')
235 .attr("r", 30);
236
237 // image circle
238 nodeAppend.append("image")
239 .attr("clip-path", function(d) { return "url(#clip" + d.index + "
    )"})
```

```
240 .attr("xlink:href", function(d) { if (d.type == "person" || d.
        type == "duplicate") { return  "static/images/" + d.fill } })
241 .attr("x", -42)
242 .attr("y", -39)
243 .attr("width",80)
244 .attr("height", 80);
245
246 // add the text
247 nodeAppend.append("text")
248 .attr("class", "nodetext")
249 .attr("x", 0)
250 .attr("y", function(d) {
251
252 if(d.type == "group") { return 0; }
253 else { return 48 }
254
255 })
256 .attr("text-anchor", "middle")
257 .style("font", function(d) { if (d.type == "group") {return "16px
        sans-serif";} else { return "12px sans-serif" }})
258 .style("font-weight", function(d) { if (d.type == "group") {
        return "bold";}})
259 .style("fill", function(d) { if (d.type == "group") {return "
        white";}})
260 .text(function(d) {
261 return d.firstName; });
262
263 nodeStyle.exit().remove();
264
265 force.start();
266
267 }
```

Extract of Code from Facebook Scraper

Listed below is an extract containing several functions from the Facebook Scraper code. The scraper requests data from the profile and inbox of the user, and then identifies

relevant information contained in the messages of the user, such as links that have been sent to or from particular contacts.

```
1
2  import facebook
3  import re
4  from Message import Message
5  from Person import Person
6  import sys
7  import os
8  import time
9  import urllib
10 import urllib2
11 import json
12 import traceback
13
14 class FacebookScraper:
15
16 def __init__(self, token):
17 self.token = token
18 self.messageList = []
19 self.reciprocatedSet = set([])
20 self.graph = facebook.GraphAPI(self.token)
21 self.attachmentList = []
22 self.profile  = None
23
24 def retrieveProfilePictures(self, people):
25
26 for person in people:
27 if person.facebookID != None:
28 friend = self.graph.get_connections(person.facebookID, "picture",
       width=200)
29
30 dirs = os.path.dirname(__file__)
31 filePath = os.path.join(dirs, "static/images/"+ str(person.
     personID) +'.jpg')
32
33 fd = open(filePath, 'wb')
```

```
fd.write(friend["data"])
fd.close()

def retrieveFacebookInbox(self):

self.profile = self.graph.get_object("v1.0/me")
#friends = graph.get_connections("me", "friends")
messages = self.graph.get_connections("v1.0/me","inbox")
i = 0

paging = messages["paging"]

self.examineMessages(self.profile, messages)

for x in range(0, 1):
nextPage =  paging["next"].split(self.profile["id"] + "/")
messages = self.graph.get_connections("v1.0/me", nextPage[1])
paging = messages["paging"]
self.examineMessages(self.profile, messages)

for contact in self.messageList:
for fromContact in contact.fromPeople:
#print("From: " + fromContact.name)
self.reciprocatedSet.add(fromContact)
for toContact in contact.toPeople:
#print("To: " + toContact.name)
self.reciprocatedSet.add(toContact)

for instantMessage in self.messageList:
messageDate = instantMessage.time
messageBody = instantMessage.message
links = re.findall(r'((http[s]?://|www.)(?:[a-zA-Z]|[0-9]|[$-_@
    .&#+]|[!*() ,]|(?:%[0-9a-fA-F][0-9a-fA-F]))+)', messageBody)
if len(links) > 0:
messageSenders = instantMessage.fromPeople
for sender in messageSenders:
for contact in self.reciprocatedSet:
if sender.name == contact.name:
for link in links:
if link[1] == "www.":
```

```
73  url = "http://" + link[0]
74  else:
75  url = link[0]
76
77  try:
78  source = urllib2.urlopen(url)
79  source = source.read()
80
81  titleRE = re.compile("<title>(.+?)</title>")
82  title = titleRE.search(source).group(1)
83  print "hiho"
84  except:
85
86  title = "Not Found"
87
88  if sender.facebookID == self.profile['id']:
89  direction = "IN"
90  else:
91  direction = "OUT"
92
93  contact.links.append({"attachmentName": title, "URL": url, "date"
        : messageDate, "direction": direction, "type": "facebook"})
94
95
96  for contact in self.reciprocatedSet:
97  #contact.name = re.escape(contact.name)
98  for attachmentItem in self.attachmentList:
99  print attachmentItem["facebookID"]
100 print contact.facebookID
101 if int(attachmentItem["facebookID"]) == int(contact.facebookID):
102 print "it works"
103 contact.links.append({"attachmentName": attachmentItem["
        attachmentName"], "URL": attachmentItem["URL"], "date":
        attachmentItem["date"], "type": "facebook"})
104
105 contact.name = str(contact.name).replace("'", "''")
106
107 return
```

Appendix C – Ethics Application Example

Study 1 – Ethics Application

The following images are copies of the ethics application for the information re-finding diary study (Chapter 4).

Faculty of Engineering
Application for approval of research study involving human participants

ALL applicants must provide the following information
The applicant must be the person who will conduct the investigations; each application must be made by one applicant:

- usually the student in the case of taught or research courses,
- usually the researcher (the member of university research or academic staff) who will conduct the study in the case of funded research projects,
- usually the principal investigator in the case of applications for ethics approval in advance of submission of a research proposal

If the applicant is an Undergraduate or Postgraduate taught or research student please complete the information below. The application must be approved by a Supervisor.

Name of student:	Paul Holmes	Student No:	4151243
Course of study:	Horizon DTC	Email address:	psxph1@nottingham.ac.uk
Supervisor:	James Goulding		

If the applicant is a member of university research or academic staff, please complete the information below: For research staff, the application must be approved by the Principal Investigator

Name:		Principal Investigator (Budget Holder)	
Email address:		PI Signature:	

Title of investigation : Personal Information Retrieval Tasks Diary Study

Planned date for study to begin: 12/01/2011 Duration of Study: 1 Week

Please state whether this application is:

☐ New ☐ Revised ☐ A renewal ☐ For a continuation study

Selection of review process

Please indicate whether the application is required to go forward to the ethics committee for formal review, or, in the case of projects completed by ***taught undergraduate and postgraduate students only***, whether the application can be approved by the supervisor under the expedited review process.

☐ Formal review, application will be submitted to ethics committee

☐ Expedited review, application is approved by supervisor

Approval by supervisor: expedited review

I approve the application as supervisor of this project, under the expedited review procedure.

Name of supervisor..... Signature..... Date.....

Office use only
Date form received:
Passed to reviewers: 1. Name..... Date.....
(formal review only) 2. Name..... Date.....
Date decision returned to applicant:

Figure 3: Ethics Application Form 1 of 3

Ethical Issues Checklist

The purpose of this Checklist is to facilitate the review process and to identify any ethical issues that may concern the Committee. It is meant to be an aid to both the researcher and the Committee. Listed below are areas which require some justification and attention on your part in specifying your study protocol. Please answer each question honestly, giving full details where required. Answering "YES" to any of the questions will not necessarily lead to a negative response to your application but it will draw issues to your attention and give the reviewers the opportunity to ensure appropriate steps are being taken. In expedited review, supervisors should ensure that for any questions where the answer "YES" has been given, appropriate measures have been taken to maintain ethical compliance.

Applicant's full name: Paul Holmes

You must complete ALL of this section before submitting your application

1 Who is the population to be studied?

Researchers within Horizon and the Human Factors Research Group.

2 Please give details of how the participants will be identified, approached and recruited. (Include any relationship between the investigator and participants e.g. instructor-student).

Participants will be recruited through the mailing lists of the respective research groups, both of which the investigator is a member of.

3 Will the population studied include any vulnerable members of the public? **YES** ☐ **NO** ☐
Note: for the purpose of ethics approval this includes participants who are under 18, people who are disabled or in poor health, and also those who are non-English speakers and may not be able to understand the consent forms. (If YES, please give further details)

.....

4 Will it be possible to associate specific information in your records with specific participants on the basis of name, position or other identifying information contained in your records? **YES** ☐ **NO** ☐

5 What steps have you taken to ensure confidentiality of personal information and anonymity of data both during the study and in the release of its findings?

All data will be recorded anonymously and any information or comments that might identify an individual will be removed during the collation and analysis of the data.

6 Describe what data will be stored, where, for what period of time, the measures that will be put in place to ensure security of the data, who will have access to the data, and the method and timing of disposal of the data.

Participants in this project will be asked to record the details of occasions when they have to perform personal information retrieval tasks, and what steps these tasks consist of, throughout each day. An example of a personal information retrieval task might be locating a webpage you've previously seen that contains some information you require, but you can't remember the URL or name of. This data will be used to identify common types of retrieval tasks people perform or to examine features of particular types of tasks, such as those involving the fusion of data from multiple sources.

The data recorded as part of this project will be either collected via a web form or a dictaphone, dependant on the preferences of the participants. All data collected will be kept securely within the university, or during collection reside within a secure server in the UK. Access to the data collected will be restricted to those immediately involved in the research, Paul Holmes, James Goulding and Sarah Sharples.

7 Will persons participating in the study be subjected to physical or psychological discomfort, pain or aversive stimuli which is more than expected from everyday life? (If YES, please give further details) **YES** ☐ **NO** ☐

.....

Figure 4: Ethics Application Form 2 of 3

8	Will participants engage in strenuous or unaccustomed physical activity? (If YES, please give further details)	**YES** ☐	**NO** ☐
9	Will the investigation use procedures designed to induce participants to act contrary to their wishes? (If YES, please give further details)	**YES** ☐	**NO** ☐
10	Will the investigation use procedures designed to induce embarrassment, humiliation, lowered self esteem, guilt, conflict, anger, discouragement or other emotional reactions? (If YES, please give further details)	**YES** ☐	**NO** ☐
11	Will participants be induced to disclose information of an intimate or otherwise sensitive nature? (If YES, please give further details)	**YES** ☐	**NO** ☐
12	Will participants be deceived or actively misled in any manner? (If YES, please give further details)	**YES** ☐	**NO** ☐
13	Will information be withheld from participants that they might reasonably expect to receive? (If YES, please give further details)	**YES** ☐	**NO** ☐
14	Will the research involve potentially sensitive topics? (If YES, please give further details)	**YES** ☐	**NO** ☐

If you require space for additional information, please add it here and identify the question to which it refers:

Figure 5: Ethics Application Form 3 of 3

The following images are from an application to amend the original ethics approval for the information re-finding diary study. This amendment was necessary as the decision was made to use a mobile phone based diary, whereas the original method envisaged using paper or a webform to record the diary entries.

The University of Nottingham

Faculty of Engineering

Application for *amendment* to approved research study involving human participants

ALL applicants must provide the following information

The applicant must be the person who will conduct the investigations; each application must be made by one applicant:

- usually the student in the case of taught or research courses,
- usually the researcher (the member of university research or academic staff) who will conduct the study in the case of funded research projects,
- usually the principal investigator in the case of applications for ethics approval in advance of submission of a research proposal.

If the applicant is an Undergraduate or Postgraduate taught or research student please complete the information below. The application must be approved by a Supervisor.

Name of student:	Paul Holmes	Student No:	4151243
Course of study:	Horizon DTC 4 Year PhD	Email address:	psxph1@nottingham.ac.uk
Supervisor:	Sarah Sharples + James Goulding		

If the applicant is a member of university research or academic staff, please complete the information below: For research staff, the application must be approved by the Principal Investigator.

Name:		Principal Investigator (Budget Holder)	
Email address:		PI Signature:	

Title of investigation: .....

Planned date for study to begin ..... Duration of Study .....

This application is an amendment to a previously approved ethical application. Please provide the following information concerning the *original application*.

Title of original application:	Personal Information Retrieval Tasks Diary Study	
Reference of original application:	2011-92	
Decision of original application:	Approval awarded subject to changes	Date of decision: 07/12/11

Please confirm that you agree that the following statements regarding the amendments are true:

☐ The amendments made to this study do not substantively alter the responses given to the Ethical Issues Checklist in the original application.

☐ There have been no changes in the substance of the consent form(s) where used.

Office use only
Date form received:
Passed to reviewer: 1. Name..... Date.....
Date decision returned to applicant:

Figure 6: Study 1 – Ethics Application Amendment Form 1 of 2

Details of amendments made:

Please use this space to provide details of the amendments made to the original ethics application and how these changes will impact on the ethical nature of the study.

- Data collection will now be solely through the use of android mobile phones. These will either be supplied to the participants or the participants can use their personal android mobile phones if they have one and wish to use it. If participants use their personal mobile phones then extraction of the data will require the researcher to interact with their personal phone. To avoid any privacy or sensitivity about this the participant will be instructed how to transfer the data from their phone for the researcher to collect via a USB memory stick.
- Participants will be asked to record basic demographic information (gender and age), and perform a self-assessed computer literacy questionnaire to determine their fluency with computers.
- Interviews will be conducted at the end of the study, to review the answers recorded in the diary concerning the process undertaken in locating the information the participant required, and their memory of the information they were looking for. These interviews will be digitally recorded, and any transcripts made will ensure any identifying information is made anonymous.
- All of the information new information collected will be stored securely in keeping with the original application.

As the applicant I confirm that I have read and understand the Ethical requirements for my study and have read and complied with the University of Nottingham Code of Research Conduct and Research Ethics.

Signature of applicant ..... Date .....

As supervisor, I confirm that I have checked the details of this application.

Signature of supervisor ..... Date .....

Page **2** of **3**

Figure 7: Study 1 – Ethics Application Amendment Form 2 of 2

Below are images of the information sheet and consent form that were provided to participants prior to beginning the diary study.

Participant Information Sheet

The purpose of this study is to examine what people remember about their information, and what they do to locate it, by recording a cross section of their daily information activities. The results of the study will ideally provide insights into how people behave with their information and from this how we might therefore better design interfaces for managing our information.

In exchange for participating in this study you will be given £20 of highstreet vouchers to compensate you for your time. As a participant you will be asked to record on an Android mobile phone the details of the occasions when you perform personal information retrieval tasks (on any electronic device), as part of an electronic "diary".

Personal information retrieval tasks are those occasions when you need some information, which you have seen before, and then you try to locate that information again. They might be, for example, locating a webpage you've previously seen, finding the folder you put the document your Christmas list is in, or locating an old email where you only know the sender's name. These tasks might also be multi-stage in their nature, such as finding the record of an old instant message that mentions a nice restaurant in an area unfamiliar to you that you then search for on the Internet. Each of these diary entries will ask you to record the following aspects of what you did when looking for the information:

- What was your information need? (What was it you wanted to find out?)
- What did you remember about the information before looking for it?
- What did you do to find the information you needed?
- Were you successful or unsuccessful in locating the required information?
- How difficult was it to locate?

You would be asked to record the tasks you perform in this manner for two weeks using an application developed for this study. If you do not have an Android mobile phone then a phone can be loaned to you for the period of the study. Preferably entries made in the diary would be made shortly after completing the task, but may alternatively be completed at a later more convenient time.

At the end of the study the data you've recorded will be collected from the phone, and a short interview (30 minutes or less) conducted with you. You will also be given your gift vouchers for participating assuming diary entries have been recorded.

If you participate you would be free to withdraw from this user trial at any stage without giving a reason, and any data you record will be stored securely and anonymously. This Participant Information Sheet should be read in conjunction with the Participant Consent Form. If you have any further questions please contact me (details below).

Thank you for your participation.

Paul Holmes, PhD Student, Horizon Doctoral Training Centre, University of Nottingham

Email: psxph1@nottingham.ac.uk
Mobile: 07796165053
Web: https://www.horizon.ac.uk/students/horizon/students/107student-template.html

Figure 8: Study 1 – Participant Information Sheet

PARTICIPANT CONSENT FORM

Project title .....

Researcher .....Paul Holmes.....

- I have read the Participant Information Sheet andthe nature and purpose of the research project ha been explained to me. I understand and agree to take part.
- I understand the purpose of the research project and my involvement in it.
- I understand that I may withdraw from the research project at any stage and that this will not affect my status now or in the future.
- I understand that while information gained during the study may be published, I will not be individually identified and my personal results will remain confidential.
- I understand that all data will be stored securely in electronic form and hard copy.
- I understand that I may contact the researcher or supervisor if I require further information about t research, and that I may contact the Chair of the Research Ethics Commitee for Engineering, University of Nottingham, if I wish to make a complaint relating to my involvement in the research.

Signed ..... (research participant)

Print name ..... **Date** .....

Contact details

Researcher: Paul Holmes - psxph1@nottingham.ac.uk

Supervisor: Sarah Sharples – Sarah.Sharples@nottingham.ac.uk
Supervisor: James Goulding – James.Goulding@nottingham.ac.uk

Figure 9: Study 1 – Participant Consent Form

Study 2 – Ethics Application

The following images are copies of the ethics application for the information repertory grid and free diagramming study (Chapter 5).

Faculty of Engineering
Application for approval of research study involving human participants

ALL applicants must provide the following information
The applicant must be the person who will conduct the investigations; each application must be made by one applicant:

- usually the student in the case of taught or research courses,
- usually the researcher (the member of university research or academic staff) who will conduct the study in the case of funded research projects,
- usually the principal investigator in the case of applications for ethics approval in advance of submission of a research proposal

If the applicant is an Undergraduate or Postgraduate taught or research student please complete the information below. The application must be approved by a Supervisor.

Name of student:	Paul Holmes	Student No:	4151243
Course of study:	Horizon DTC	Email address:	p.holmes@nottingham.ac.uk
Supervisor:	Sarah Sharples		

If the applicant is a member of university research or academic staff, please complete the information below: For research staff, the application must be approved by the Principal Investigator

Name:		Principal Investigator (Budget Holder)	
Email address:		PI Signature:	

Title of investigation: Relationships and Information Groupings Between People

Planned date for study to begin: 01/04/2013 Duration of Study: 2 Weeks

Please state whether this application is:

☐ New ☐ Revised ☐ A renewal ☐ For a continuation study

Selection of review process

Please indicate whether the application is required to go forward to the ethics committee for formal review, or, in the case of projects completed by ***taught undergraduate and postgraduate students only***, whether the application can be approved by the supervisor under the expedited review process.

☐ Formal review, application will be submitted to ethics committee

☐ Expedited review, application is approved by supervisor

Approval by supervisor: expedited review

I approve the application as supervisor of this project, under the expedited review procedure.

Name of supervisor..... Signature..... Date.....

Office use only
Date form received:
Passed to reviewers: 1. Name..... Date.....
(formal review only) 2. Name..... Date.....
Date decision returned to applicant:

Figure 10: Study 2 – Ethics Application Form 1 of 2

Ethical Issues Checklist

The purpose of this Checklist is to facilitate the review process and to identify any ethical issues that may concern the Committee. It is meant to be an aid to both the researcher and the Committee. Listed below are areas which require some justification and attention on your part in specifying your study protocol. Please answer each question honestly, giving full details where required. Answering "YES" to any of the questions will not necessarily lead to a negative response to your application but it will draw issues to your attention and give the reviewers the opportunity to ensure appropriate steps are being taken. In expedited review, supervisors should ensure that for any questions where the answer "YES" has been given, appropriate measures have been taken to maintain ethical compliance.

Applicant's full name: Paul Holmes

You must complete ALL of this section before submitting your application

1 Who is the population to be studied?

This research requires people from a variety of backgrounds and experiences to examine the different structures in relationships between people and their contacts. Some of the people recruited will be students/staff at the University of Nottingham, but a variety of adult professionals will also be recruited across Nottingham.

2 Please give details of how the participants will be identified, approached and recruited. (Include any relationship between the investigator and participants e.g. instructor-student).

Participants will be identified and recruited primarily through recruitment emails or flyers distributed in the Human Factors Research Group and also external organisations, such as the Antenna business hub in Nottingham.

3 Will the population studied include any vulnerable members of the public? **YES** ☐ **NO** ☐
Note: for the purpose of ethics approval this includes participants who are under 18, people who are disabled or in poor health, and also those who are non-English speakers and may not be able to understand the consent forms. (If YES, please give further details)

.....

4 Will it be possible to associate specific information in your records with specific participants on the basis of name, position or other identifying information contained in your records? **YES** ☐ **NO** ☐

5 What steps have you taken to ensure confidentiality of personal information and anonymity of data both during the study and in the release of its findings?

All results of the study will be recorded and filed anonymously. Any information that might obviously identify an individual in the transcription of the interviews and questionnaires will be removed during the collation and analysis of the data.

6 Describe what data will be stored, where, for what period of time, the measures that will be put in place to ensure security of the data, who will have access to the data, and the method and timing of disposal of the data.

The data that will be stored includes participant's descriptors of their relationships with their contacts (the contacts themselves will not be stored), diagrams of how they perceive groupings of their contacts, relevant basic demographic information recorded on a paper questionnaire, digitally recorded interviews regarding their relationships with their contacts and the information they associate with them. All data collected will be kept securely within the university. Access to the raw data collected will be restricted to those immediately involved in the research, Paul Holmes, Sarah Sharples & James Goulding.

7 Will persons participating in the study be subjected to physical or psychological discomfort, pain or aversive stimuli which is more than expected from everyday life? (If YES, please give further details) **YES** ☐ **NO** ☐

.....

Figure 11: Study 2 – Ethics Application Form 2 of 2

Study 3 – Ethics Application

The following images are copies of the ethics application for the information 'person-centric' prototype evaluation study (Chapter 7).

Faculty of Engineering
Application for approval of research study involving human participants

ALL applicants must provide the following information
The applicant must be the person who will conduct the investigations; each application must be made by one applicant:

- usually the student in the case of taught or research courses,
- usually the researcher (the member of university research or academic staff) who will conduct the study in the case of funded research projects,
- usually the principal investigator in the case of applications for ethics approval in advance of submission of a research proposal

If the applicant is an Undergraduate or Postgraduate taught or research student please complete the information below. The application must be approved by a Supervisor.

Name of student:	Paul Holmes	Student No:	4151243
Course of study:	Horizon DTC	Email address:	p.holmes@nottingham.ac.uk
Supervisor:	James Goulding & Sarah Sharples		

If the applicant is a member of university research or academic staff, please complete the information below: For research staff, the application must be approved by the Principal Investigator

Name:		Principal Investigator (Budget Holder)	
Email address:		PI Signature:	

Title of investigation : .....

Planned date for study to begin ..... Duration of Study .....

Please state whether this application is:

☐ New ■ Revised ☐ A renewal ☐ For a continuation study

Selection of review process

Please indicate whether the application is required to go forward to the ethics committee for formal review, or, in the case of projects completed by ***taught undergraduate and postgraduate students only***, whether the application can be approved by the supervisor under the expedited review process.

■ Formal review, application will be submitted to ethics committee

☐ Expedited review, application is approved by supervisor

Approval by supervisor: expedited review

I approve the application as supervisor of this project, under the expedited review procedure.

Name of supervisor..... Signature..... Date.....

Office use only
Date form received:
Passed to reviewers: 1. Name..... Date.....
(formal review only) 2. Name..... Date.....
Date decision returned to applicant:

Figure 12: Study 3 – Ethics Application Form 1 of 2

Ethical Issues Checklist

The purpose of this Checklist is to facilitate the review process and to identify any ethical issues that may concern the Committee. It is meant to be an aid to both the researcher and the Committee. Listed below are areas which require some justification and attention on your part in specifying your study protocol. Please answer each question honestly, giving full details where required. Answering "YES" to any of the questions will not necessarily lead to a negative response to your application but it will draw issues to your attention and give the reviewers the opportunity to ensure appropriate steps are being taken. In expedited review, supervisors should ensure that for any questions where the answer "YES" has been given, appropriate measures have been taken to maintain ethical compliance.

Applicant's full name …Laura Lewis ……

You must complete ALL of this section before submitting your application

1 Who is the population to be studied?

People recruited will be students/staff at the University of Nottingham, but a variety of adult professionals will also be recruited across Nottingham. The only requirement will be that they have an email address or Facebook account they have actively used for more for a few months or more.

2 Please give details of how the participants will be identified, approached and recruited. (Include any relationship between the investigator and participants e.g. instructor-student).

Participants will be identified and recruited primarily through recruitment emails or flyers distributed in the Human Factors Research Group and also externally.

3 Will the population studied include any vulnerable members of the public? **YES** ☐ **NO** ■
Note: for the purpose of ethics approval this includes participants who are under 18, people who are disabled or in poor health, and also those who are non-English speakers and may not be able to understand the consent forms. (If YES, please give further details)

……

4 Will it be possible to associate specific information in your records with specific participants on the basis of name, position or other identifying information contained in your records? **YES** ☐ **NO** ■

5 What steps have you taken to ensure confidentiality of personal information and anonymity of data both during the study and in the release of its findings?

Participant names will not be recorded in the study. All data collected will be associated with a participant ID number rather than a name. Consent forms (which will contain the participant names) will be stored separately from study data.

The data stored from the participant's email and Facebook account during the initial data collection phase is limited to the message headers (the time, subject, people sent to/from), which are necessary to present to the participant all of the messages related to each contact and are only accessed and used by the participant. Besides this the only other data collected would be the URLs and names of attachments people have sent them or they have sent, and the Facebook profile images of their Facebook contacts that have messaged them. After the participant completes the study this data will be deleted, and only the named groupings of contacts the participants create will be retained. While the creator of the groupings will be anonymised it is possible that a participant could be identified from their contacts and groupings as a whole by someone familiar with them. None of the names of any contacts would be used in images or otherwise within publications, or reproduced in any further documents during the analysis, and any names of groups created would be anonymised if containing identifying information.

During the study participants will be asked to access some of their personal messages through the interface. As previously described these are not held locally and only when the participant clicks on a particular message they wish to view will the message content be downloaded and displayed. The downloaded messages are not stored and are discarded when the participant closes the message.

Screen capture and verbal protocol will be used throughout the study. This data will be used to codify how

Figure 13: Study 3 – Ethics Application Form 2 of 3

Ethical Issues Checklist

The purpose of this Checklist is to facilitate the review process and to identify any ethical issues that may concern the Committee. It is meant to be an aid to both the researcher and the Committee. Listed below are areas which require some justification and attention on your part in specifying your study protocol. Please answer each question honestly, giving full details where required. Answering "YES" to any of the questions will not necessarily lead to a negative response to your application but it will draw issues to your attention and give the reviewers the opportunity to ensure appropriate steps are being taken. In expedited review, supervisors should ensure that for any questions where the answer "YES" has been given, appropriate measures have been taken to maintain ethical compliance.

Applicant's full name …Laura Lewis ……

You must complete ALL of this section before submitting your application

1 Who is the population to be studied?

People recruited will be students/staff at the University of Nottingham, but a variety of adult professionals will also be recruited across Nottingham. The only requirement will be that they have an email address or Facebook account they have actively used for more for a few months or more.

2 Please give details of how the participants will be identified, approached and recruited. (Include any relationship between the investigator and participants e.g. instructor-student).

Participants will be identified and recruited primarily through recruitment emails or flyers distributed in the Human Factors Research Group and also externally.

3 Will the population studied include any vulnerable members of the public? **YES** ☐ **NO** ■
Note: for the purpose of ethics approval this includes participants who are under 18, people who are disabled or in poor health, and also those who are non-English speakers and may not be able to understand the consent forms. (If YES, please give further details)

……

4 Will it be possible to associate specific information in your records with specific participants on the basis of name, position or other identifying information contained in your records? **YES** ☐ **NO** ■

5 What steps have you taken to ensure confidentiality of personal information and anonymity of data both during the study and in the release of its findings?

Participant names will not be recorded in the study. All data collected will be associated with a participant ID number rather than a name. Consent forms (which will contain the participant names) will be stored separately from study data.

The data stored from the participant's email and Facebook account during the initial data collection phase is limited to the message headers (the time, subject, people sent to/from), which are necessary to present to the participant all of the messages related to each contact and are only accessed and used by the participant. Besides this the only other data collected would be the URLs and names of attachments people have sent them or they have sent, and the Facebook profile images of their Facebook contacts that have messaged them. After the participant completes the study this data will be deleted, and only the named groupings of contacts the participants create will be retained. While the creator of the groupings will be anonymised it is possible that a participant could be identified from their contacts and groupings as a whole by someone familiar with them. None of the names of any contacts would be used in images or otherwise within publications, or reproduced in any further documents during the analysis, and any names of groups created would be anonymised if containing identifying information.

During the study participants will be asked to access some of their personal messages through the interface. As previously described these are not held locally and only when the participant clicks on a particular message they wish to view will the message content be downloaded and displayed. The downloaded messages are not stored and are discarded when the participant closes the message.

Screen capture and verbal protocol will be used throughout the study. This data will be used to codify how

Figure 14: Study 3 – Ethics Application Form 3 of 3

Appendix D – Coding of Interface Observations

The tables and images in this appendix are extracts of the data analysis performed on the screen capture of participants grouping their contacts using two different grouping interfaces and according to specified grouping criteria in Chapter 7. The analysis was performed using Noldus Observer XT 10.5, and the figures below show images of this software and the tables provide output of the coding exported from it.

Figure 15 shows an example of the coding against the screen capture of Participant L1, part way through the ego-centric interface condition. It illustrates an example of a participant creating a series of groups prior to adding contacts to them as described by Figure 7.4. Table 1 shows the complete coding output of this condition.

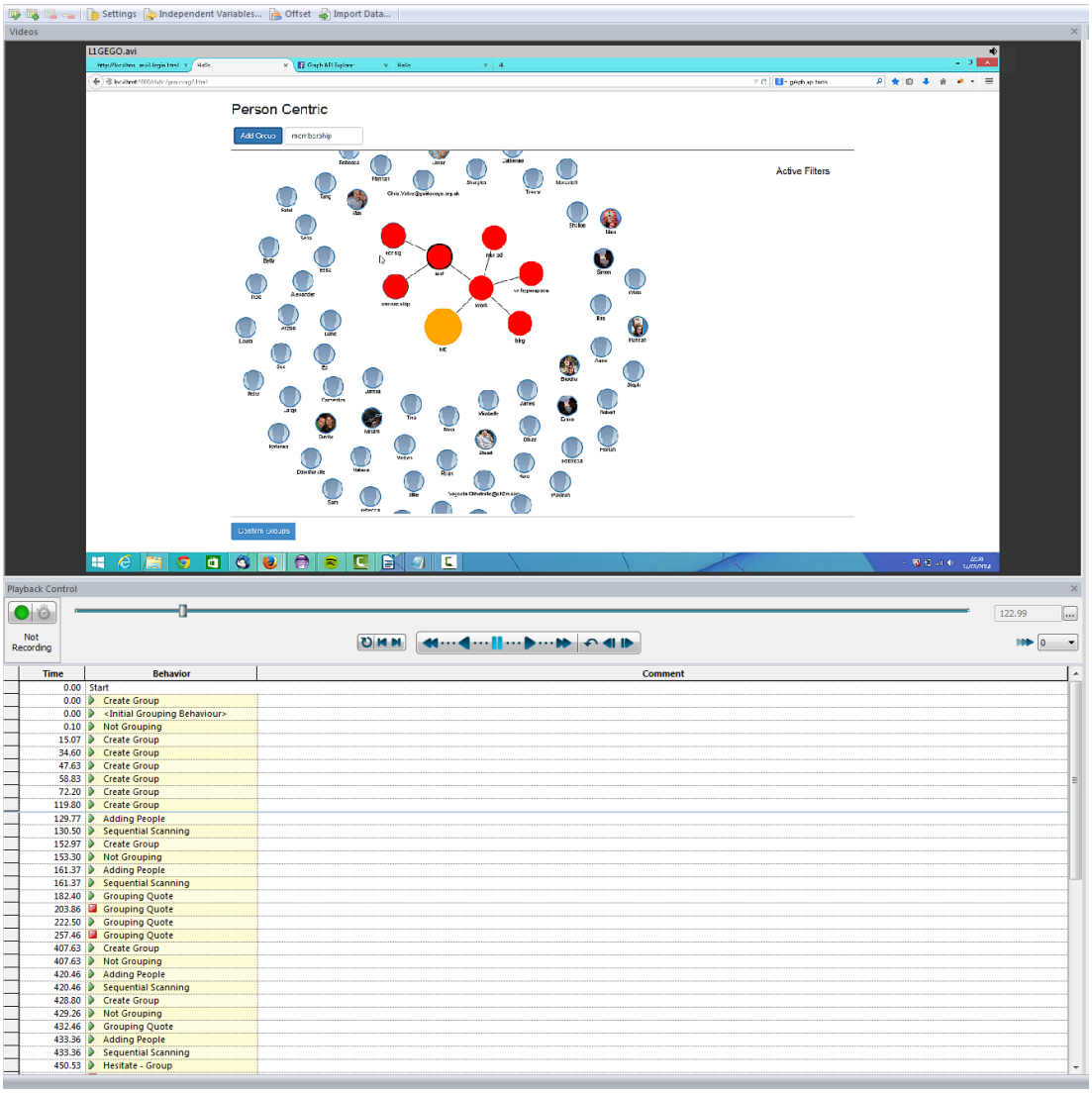


Figure 15: Participant L – Sequential Group Creation Coding Example

Time	Duration	Behaviour	Event
00:00:00	0.099999	Initial Grouping Behaviour	State start
00:00:00	15.066516	Create Group	State start
00:00:00	0	Initial Grouping Behaviour	State stop
00:00:00	130.398696	Not Grouping	State start
00:00:15	0	Create Group	State stop
00:00:15	19.533138	Create Group	State start
00:00:34	0	Create Group	State stop
00:00:34	13.033203	Create Group	State start
00:00:47	0	Create Group	State stop
00:00:47	11.199888	Create Group	State start
00:00:58	0	Create Group	State stop
00:00:58	13.366533	Create Group	State start
00:01:12	0	Create Group	State stop
00:01:12	47.599524	Create Group	State start
00:01:59	0	Create Group	State stop
00:01:59	9.966567	Create Group	State start
00:02:09	0	Create Group	State stop
00:02:09	23.199768	Adding People	State start
00:02:10	0	Not Grouping	State stop
00:02:10	22.799772	Sequential Scanning	State start
00:02:32	0	Adding People	State stop
00:02:32	8.399916	Create Group	State start
00:02:33	0	Sequential Scanning	State stop
00:02:33	8.066586	Not Grouping	State start
00:02:41	0	Create Group	State stop
00:02:41	0	Not Grouping	State stop
00:02:41	246.264204	Adding People	State start
00:02:41	246.264204	Sequential Scanning	State start
00:06:47	0	Adding People	State stop
00:06:47	0	Sequential Scanning	State stop
00:06:47	12.833205	Create Group	State start
00:06:47	12.833205	Not Grouping	State start
00:07:00	0	Create Group	State stop

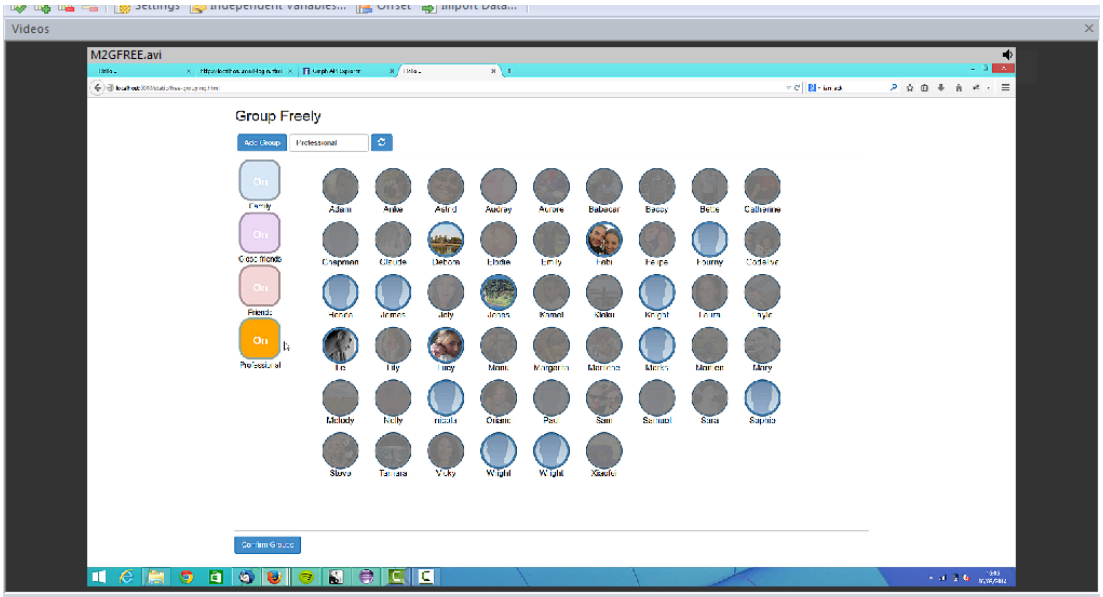
Time	Duration	Behaviour	Event
00:07:00	0	Not Grouping	State stop
00:07:00	8.33325	Adding People	State start
00:07:00	8.799912	Sequential Scanning	State start
00:07:08	0	Adding People	State stop
00:07:08	4.566621	Create Group	State start
00:07:09	0	Sequential Scanning	State stop
00:07:09	4.099959	Not Grouping	State start
00:07:12	18.199818	Grouping Quote	State start
00:07:13	0	Create Group	State stop
00:07:13	0	Not Grouping	State stop
00:07:13	17.166495	Adding People	State start
00:07:13	18.833145	Sequential Scanning	State start
00:07:30	0	Adding People	State stop
00:07:30	1.66665	Hesitate - Group	State start
00:07:30	0	Grouping Quote	State stop
00:07:32	0	Hesitate - Group	State stop
00:07:32	0	Sequential Scanning	State stop
00:07:32	1.166655	Create Group	State start
00:07:32	22.366443	Not Grouping	State start
00:07:33	0	Create Group	State stop
00:07:33	10.233231	Hesitate - Group	State start
00:07:43	0	Hesitate - Group	State stop
00:07:43	5.699943	Create Group	State start
00:07:49	0	Create Group	State stop
00:07:49	5.266614	Create Group	State start
00:07:54	0	Create Group	State stop
00:07:54	0	Not Grouping	State stop
00:07:54	15.833175	Adding People	State start
00:07:54	15.833175	Sequential Scanning	State start
00:08:10	0	Adding People	State stop
00:08:10	0	Sequential Scanning	State stop
00:08:10	13.499865	Create Group	State start
00:08:10	13.499865	Not Grouping	State start

Time	Duration	Behaviour	Event
00:08:23	0	Create Group	State stop
00:08:23	0	Not Grouping	State stop
00:08:23	6.499935	Adding People	State start
00:08:23	6.499935	Sequential Scanning	State start
00:08:30	0	Adding People	State stop
00:08:30	0	Sequential Scanning	State stop
00:08:30	0.933324	Create Group	State start
00:08:30	0.933324	Not Grouping	State start
00:08:31	0	Create Group	State stop
00:08:31	0	Not Grouping	State stop
00:08:31	8.699913	Adding People	State start
00:08:31	10.133232	Sequential Scanning	State start
00:08:40	0	Adding People	State stop
00:08:40	5.799942	Create Group	State start
00:08:41	0	Sequential Scanning	State stop
00:08:41	4.733286	Not Grouping	State start
00:08:45	0	Create Group	State stop
00:08:45	6.399936	Hesitate - Person	State start
00:08:46	0	Not Grouping	State stop
00:08:46	32.599674	Sequential Scanning	State start
00:08:52	0	Hesitate - Person	State stop
00:08:52	25.433079	Adding People	State start
00:09:17	0	Adding People	State stop
00:09:17	8.699913	Create Group	State start
00:09:18	0	Sequential Scanning	State stop
00:09:18	7.566591	Not Grouping	State start
00:09:26	0	Create Group	State stop
00:09:26	0	Not Grouping	State stop
00:09:26	189.531438	Adding People	State start
00:09:26	277.763889	Sequential Scanning	State start
00:12:35	0	Adding People	State stop
00:12:35	10.166565	Hesitate - Person	State start
00:12:46	0	Hesitate - Person	State stop

Time	Duration	Behaviour	Event
00:12:46	77.66589	Adding People	State start
00:14:03	0	Adding People	State stop
00:14:03	10.433229	Create Group	State start
00:14:04	0	Sequential Scanning	State stop
00:14:04	10.033233	Not Grouping	State start
00:14:14	0	Create Group	State stop
00:14:14	0	Not Grouping	State stop
00:14:14	179.375208	Adding People	State start
00:14:14	179.375208	Sequential Scanning	State start
00:17:13	0	Adding People	State stop
00:17:13	0	Sequential Scanning	State stop

Table 1: Participant L Ego-centric condition coding output

Figure 16 and Table 2 provide an example of the coding of Participant P. They illustrate an example of a participant who has followed a repeated pattern of creating a contact group and followed by adding many contacts to that group as illustrated by Figure 7.5.



Time	Behavior	Comment
0.00	Start	
0.00	Create Group	
0.00	<Initial Grouping Behaviour>	
6.43	Not Grouping	
31.73	Adding People	
32.43	Searching	
43.23	Create Group	
44.17	Not Grouping	
48.97	Adding People	
49.23	Sequential Scanning	
119.10	Create Group	
120.77	Not Grouping	
125.57	Adding People	
125.77	Sequential Scanning	
253.66	Create Group	
253.86	Not Grouping	
273.66	Adding People	
273.73	Sequential Scanning	
383.56	Create Group	
383.63	Not Grouping	
413.63	Adding People	
413.63	Sequential Scanning	
453.50	Stop	

Figure 16: Participant P – Non-sequential Group Creation Coding Example

Time	Duration	Behaviour	Event
00:00:00	6.433269	Grouping Behaviour>	State start
00:00:00	31.733016	Create Group	State start
00:00:06	0	Grouping Behaviour>	State stop
00:00:06	25.99974	Not Grouping	State start
00:00:31	0	Create Group	State stop
00:00:31	11.499885	Adding People	State start
00:00:32	0	Not Grouping	State stop
00:00:32	11.733216	Searching	State start
00:00:43	0	Adding People	State stop
00:00:43	5.733276	Create Group	State start
00:00:44	0	Searching	State stop
00:00:44	5.066616	Not Grouping	State start
00:00:48	0	Create Group	State stop
00:00:48	70.132632	Adding People	State start
00:00:49	0	Not Grouping	State stop
00:00:49	71.532618	Sequential Scanning	State start
00:01:59	0	Adding People	State stop
00:01:59	6.466602	Create Group	State start
00:02:00	0	Sequential Scanning	State stop
00:02:00	4.99995	Not Grouping	State start
00:02:05	0	Create Group	State stop
00:02:05	128.098719	Adding People	State start
00:02:05	0	Not Grouping	State stop
00:02:05	128.098719	Sequential Scanning	State start
00:04:13	0	Adding People	State stop
00:04:13	19.9998	Create Group	State start
00:04:13	0	Sequential Scanning	State stop
00:04:13	19.866468	Not Grouping	State start
00:04:33	0	Create Group	State stop
00:04:33	109.898901	Adding People	State start
00:04:33	0	Not Grouping	State stop
00:04:33	109.898901	Sequential Scanning	State start
00:06:23	0	Adding People	State stop

Time	Duration	Behaviour	Event
00:06:23	30.066366	Create Group	State start
00:06:23	0	Sequential Scanning	State stop
00:06:23	29.9997	Not Grouping	State start
00:06:53	0	Create Group	State stop
00:06:53	0	Not Grouping	State stop
00:06:53	39.870803	Adding People	State start
00:06:53	39.870803	Sequential Scanning	State start
00:07:33	0	Adding People	State stop
00:07:33	0	Sequential Scanning	State stop

Table 2: Participant P List condition coding output

Appendix E – Study 3 Interface Evaluation Interview Questions

In the semi-structured interviews conducted as part of the evaluation in Chapter 7 each of the participants were asked the questions listed in this appendix. The questions were as follows:

1. Groups
    - How satisfied were you with the groups you created?
    - Would you want make any changes to the groupings now retrospectively? What changes?
    - How did you find using the two grouping mechanisms to create groups?
        (a) Which did you think created more useful groups?
        (b) Which did you prefer to use to group people?
2. Filters
    - Importance filters.
        (a) Did you filter people based on importance?
        (b) When and why did you use this? Why did you not?
        (c) Who did you want to filter out or keep?
    - Dimension filters.
        (a) Did you use the dimension filters?
        (b) When and why did you choose to do this?
        (c) Who did you want to filter out or keep?

- Did the filters assist you in performing the tasks or not?

3. Structure
    - Did you restructure the groupings?
    - When and why did you choose to do this?
    - Did the restructuring assist you in performing the tasks or not?
4. Person details screen.
    - How did you feel about the way the interface associated a variety of information with people?
    - Was exposing messages like this useful or not?
    - Was exposing files like this useful or not?
    - Was exposing links like this useful or not?
5. Tasks.
    - Did you employ any particular strategies when engaging in the tasks?
        (a) For sharing?
        (b) For re-finding?
    - Which task did you find easiest/hardest, why?
    - How easy or difficult was it to locate certain people you wanted?
    - How easy or difficult was it to locate information you wanted?
    - Was the system more suited to sharing or re-finding information? Why?
6. Sharing.
    - Would you use a system like this to control who you share information with? Why (not)?
    - Are there any specific circumstances you might wish to view information in this way over a more traditional view of your data?
    - How does this system compare with the way you currently share/control information? What advantages/disadvantages are there to each?
7. Re-finding.
    - Would you use a system like this to re-find information? Why (not)?

- Are there any specific circumstances you might wish to view information in this way over a more traditional view of your data?
- How does this system compare with the way you currently re-find information? What advantages/disadvantages are there to each?

8. Interface designs.
   - Can you rate each of the interfaces from best to worst?
   - What do do you think the benefits and drawbacks to each design are?

9. Experience.
   - How did using the system make you feel?
   - What did you think to the approach this system took to presenting information for sharing or re-finding tasks?
   - Would you ever use a system like this? If so, in what circumstances would you use it?
   - If it was an option, would switch between different views of your information to perform tasks?
   - If you could make changes to this system to help you perform information sharing or re-finding tasks what might they be?