Distilling Mobile Privacy Requirements from Qualitative Data

Thesis

Distilling Mobile Privacy Requirements from Qualitative Data

Keerthi Thomas

Computing Department

The Open University

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

July 2013

Abstract

As mobile computing applications have become commonplace, it is increasingly important for them to address end-users' privacy requirements. Mobile privacy requirements depend on a number of contextual socio-cultural factors to which mobility adds another level of contextual variation. However, traditional requirements elicitation methods do not sufficiently account for contextual factors and therefore cannot be used effectively to represent and analyse the privacy requirements of mobile end users. On the other hand, methods that investigate contextual factors tend to produce data which can be difficult to use for requirements modelling. To address this problem, we have developed a *Distillation* approach that employs a problem analysis model to extract and refine privacy requirements for mobile applications from raw data gathered through empirical studies involving real users. Our aim was to enable the extraction of mobile privacy requirements that account for relevant contextual factors while contributing to the software design and implementation process. A key feature of the distillation approach is a problem structuring framework called *privacy facets (PriF)*. The facets in the PriF framework support the identification of privacy requirements from different contextual perspectives namely - actors, information, information-flows and places. The PriF framework also aids in uncovering privacy determinants and threats that a system must take into account in order to support the end-user's privacy. In this work, we first show the working of distillation using qualitative data taken from an empirical study which involved social-networking practices of mobile users. As a means of validating distillation, another distinctly separate qualitative dataset from a location-tracking study is used, in both cases, the empirical studies relate to privacy issues faced by real users observed in their mobile environment.

To
my parents Thomas & Mariarose,
my wife Rebeca and son Timothy

Acknowledgements

I am grateful for the opportunity I was given to pursue my PhD on the Privacy Rights Management for Mobile Applications (PRiMMA) project, for this I thank my supervisors Bashar Nuseibeh, Arosha Bandara and Blaine Price. Without their help I couldn't have completed this thesis. I owe much to Bashar's strategic guidance, Arosha's challenges and motivations and Blaine's support and encouragement.

It was an immense privilege to have received regular feedback and help from Michael Jackson who was kind enough to bear with me while I learnt from him. I am thankful to my colleague Clara Mancini for shaping my ideas on the use of qualitative methods. I thank Yijun Yu, Thein Than Thun and Charles Haley for the their critical feedback and helpful suggestions. Many thanks to Marian Petre for her expert advice and words of wisdom which helped me cope with the stresses and strains of doing a PhD. I also thank Robin Laney from PG-Forum who provided guidance on research methods. Although not directly related to my thesis, I must acknowledge the support I received from Enrico Motta and Mathieu d'Aquin to write my thesis while also working at KMi. I am also grateful to Paul Mulholland from KMi who provided insights on improving reliability of my approach.

I owe much to my wife Rebeca who has been steadfast and unwavering in her support for me. I am deeply grateful to my mother Maria Rose and my late dad Anthony Thomas for raising my aspirations while I was at home. It was my first sister Lillian Mary and my brother-in-law R. Prabhagaran who inspired me from an early age to pursue a PhD, I thank them both.

This work was done as part of the Privacy Rights Management for Mobile Applications (PRiMMA) project, which was funded by the UK EPSRC (Grant # EP/F024037/1)

Contents

List of Figures

List of Tables

LIST OF TABLES

1

Introduction

In any society, individuals often play different and sometimes conflicting roles. Managing these roles requires individuals to continuously regulate their social interactions by either presenting a different *'self'* at various times or by restricting information about themselves and their emotions. This regulation of 'self' is critical to how individuals protect themselves from the stresses and strains of social interactions (Westin, 1967, p.13). Privacy, which is centered in personal liberty and autonomy, is therefore considered essential for one's physical and emotional well being [1]. Modern technologies can, on one hand, enable and facilitate social interactions, but on the other hand, when they are poorly designed they encroach on the privacy of users. One such modern innovation is the mobile phone. In the last few years, mobile phones have increased in popularity and sophistication, embedding several sensors and instruments on board. The software applications that run on mobile phones are not lagging behind either: they now support sophisticated user interaction and functionality which keeps improving with every new release. New generations of mobile phones, starting with the iPhone [2] from Apple have dramatically changed the way people perceive and interact with their mobile devices. By mid 2012, Apple claimed that its iPhone 'App Store' (software application

[1] The Supreme Court of Canada had stated: 'society has come to realize that privacy is at the heart of liberty in a modern state...Grounded in man's physical and moral autonomy, privacy is essential for the well-being of the individual' R. v. Dyment (1988), 55 D.L.R. (4th) 503 at 513 (S.C.C)

[2] http://www.apple.com/uk/iphone/

store) served 40 billion application downloads [1]. Mobile applications not only provide rich services such as news, stock information, weather, games etc., but also perform functions not envisaged before. For example, a mobile phone can be converted into an emergency flashlight, a barcode-scanner or even a navigational compass depending on the type of mobile application used. In order to support such rich functionality, many applications rely on having full access to services and data on board a mobile device. Since mobile applications have the capability to collect, process and exploit heterogeneous data within a mobile phone, it is vital that developers of such systems address issues of privacy, especially as increasing numbers of mobile users connect and network with each other to share information.

Privacy is a difficult notion to grasp because there is no single definition for it. As Solove (2008) observes, privacy is rather an umbrella term referring to several concepts which have both similarities and differences. Some have defined privacy as *'the right to be left alone'* (Warren & Brandeis, 1890) or a *'claim to free self-determination'* (Westin, 1967); others have described it in terms of *'control of access to the self within a social environment'* (Altman, 1975) or in terms of *'degrees of social freedom through selective social engagement'* (Schoeman, 1984). More recently, Petronio (2002) describes privacy as *'a dialectic process of boundary management'*. Palen & Dourish (2003) state:

> *Privacy management is not about setting rules and enforcing them; rather, it is the continual management of boundaries between different spheres of action and degrees of disclosure within those spheres. Boundaries move dynamically as the context changes. These boundaries reflect tensions between conflicting goals; boundaries occur at points of balance and resolution. The significance of information technology in this view lies in its ability to disrupt or destabilize the regulation of boundaries. Information technology plays multiple roles. It can form part of the context in which the process of boundary maintenance is conducted; transform the boundaries; be a means of managing boundaries; mediate representations of action across boundaries; and so forth.*

[1] http://www.apple.com/pr/library/2013/01/07App-Store-Tops-40-Billion-Downloads-with-Almost-Half-in-2012.html

To design privacy-aware mobile applications, software engineers should be able to decipher this concept of privacy, understand the users' privacy expectations and perceptions and then relate them to the existing functional requirements of the software system (Spiekermann & Cranor, 2009).

From Zave & Jackson (1997), we know requirements are optative properties that relate to what we want to achieve or control in an environment through the use of a software system. However, these functional requirements must be augmented with privacy and security requirements if they are to support public safety and individual rights. For mobile applications, privacy requirements can be broadly stated as a set of requirements which regulate the collection, storage, processing and dissemination of a mobile user's data, to protect the user from privacy-related harms.

Previous research by Adams & Sasse (1999) has highlighted how system designers, policy makers, and organisations can easily become isolated from end-users' perceptions of privacy in different contexts. For mobile applications, the end-users' context changes frequently and in unpredictable ways. Observations of such users suggest that changes in context result in changes in the users' privacy requirements (Mancini *et al.*, 2009). Omitting these privacy requirements affects the user's privacy and consequently has an impact on how well the system is adopted or utilised. All too often the design of technologies influencing privacy management is considered and addressed as an afterthought (Anton & Earp, 2000) when it is too late to ensure that the necessary guarantees and assurances of privacy are met.

Eliciting end-user privacy requirements for mobile applications is both sensitive and difficult. Questionnaires do not elicit rich enough information about users' decisions and how they have been influenced by the emerging context in any particular situation. Beyer & Holtzblatt (1995) have shown that shadowing of users is useful in capturing contextual requirements to design and build new systems. However, when it comes to privacy this approach is problematic, since the experience of being under constant observation is likely to change the behaviour of the users in ways that invalidate any observed behaviours with respect to privacy. To overcome such limitations, Goguen & Linde (1993) propose the use of enthomethodology to gain deeper understanding of the issues involved. They suggest the use of conversation, discourse and interaction analyses to obtain tacit knowledge of what users actually do in different work situations. While there have been ethnographic studies conducted by the HCI community to study end-user privacy (Adams, 2000; Bellotti & Sellen, 1993), including our own user studies (Mancini *et al.*, 2009, 2011), the qualitative data from such studies do not readily translate into requirements. Often, the qualitative data in the form of interview transcripts may contain privacy requirements that are embedded and tightly entwined with user's contextual experiences. The technical challenge is therefore to extract these requirements from the raw qualitative data using a systematic approach. In view of the above, we state our research aims and objectives below.

1.1 Research Problem, Scope and Contribution

The main aim of this research is to provide a suitable approach for software engineers to derive privacy requirements for mobile applications from user experience reports and empirical studies. The research question can be stated as:

Research question
How can we derive mobile privacy requirements from qualitative data?

For the above question, a new requirements engineering approach will have to address several sub questions: (a) *how can we structure and separate privacy relevant information from qualitative data?* (b) *how can we identify and extract mobile privacy requirements from this data?* (c) *how can we model and represent the extracted mobile privacy requirements*, and (d) *what type of analyses must the approach support?* These are the questions this research aims to answer.

Research Aims & Objectives

The research objectives are to assist software engineers to:

(a) Structure the qualitative data to be able to analyse them in a systematic way

(b) Identify the end-user's privacy goals and expectations in different contexts using the structured data

(c) Isolate the privacy threats faced by the user

(d) Derive and model privacy requirements which support reasoning and analyses

(e) Represent the privacy requirements in such a way that designers, engineers and architects can make use of them.

Problem scope

A key assumption is that qualitative data will be taken from empirical studies whose focus is on the privacy of mobile users. This is a key requirement, otherwise the qualitative data will be orthogonal to the objectives of the requirements engineering approach, rendering it ineffective.

Privacy is a broad topic and the focus of this research is to derive *personal privacy* requirements for mobile application users. Personal privacy refers to how people manage their privacy with respect to other individuals, as opposed to large organisations (Iachello & Hong, 2007). Further, personal privacy is a more 'intimate' concern than *organisational privacy* which relates to an enterprise's requirement to meet existing legislation (Bagues *et al.*, 2007). This being the focus of our research, we do not concentrate on privacy relating to application service providers who capture and use data (or transmit to 3rd parties) in ways which affect the privacy of the end-users; this has been studied by others (Cranor, 1998; Deng *et al.*, 2011; Kalloniatis *et al.*, 2007; Spiekermann & Cranor, 2009; Yu & Cysneiros, 2002) and is beyond the scope of our work. The main aim of our work is to identify personal privacy requirements that exclusively support mobile users and their interactions with other mobile users in a social network or group.

Novel contribution

To achieve our stated research objectives, we propose *distillation*, a novel systematic approach which employs analysis models and patterns to extract and refine privacy requirements from qualitative data gathered from empirical studies. Distillation makes use of a framework called *Privacy Facets (PriF)* to (a) structure the qualitative data (b) model information-flow problem patterns that relate to user's privacy, and (c) perform problem analysis to uncover new privacy requirements for mobile applications. To structure the qualitative data, we provide *negative behaviour patterns (NBPs)* and *negative emotional indicators (NEIs)* which identify *privacy-sensitive contexts (PS-contexts)* which contain *privacy norms* and *privacy harms*. While a set of *facet ques-*

tions and codes are provided to analyse these privacy norms, a set of *extraction rules* are provided to help isolate *privacy threats* and *privacy concerns* associated with them. Further, *information-flow patterns* assist in modelling and analysing new privacy requirements. Finally, we provide *language extensions* for Privacy Arguments (Tun *et al.*, 2012) to represent these newly discovered privacy requirements. These are explained, demonstrated and evaluated in the following chapters.

The thesis is structured as follows. In the next chapter we review the literature and provide a background for the specific challenges that relate to requirements engineering for mobile privacy. Chapter 3 provides a description of the PriF framework. Chapter 4 provides a detailed description of how the distillation approach can be used to extract and refine privacy requirements using qualitative data. Chapter 5 validates distillation using a case study based on the qualitative data taken from an investigation of a mobile tracking application. Chapter 6 provides practical guidance in exploiting existing software tools to support distillation. Finally, in Chapter 7 we present our conclusions and future research directions.

2

Background and Related Work

This chapter provides background information on key concepts that are relevant to the research aims and objectives. The first section describes a notion of privacy and a taxonomy of privacy threats. The next section provides a general overview of requirements engineering techniques and methods. The third section introduces mobile applications and the current requirements engineering approaches used for eliciting mobility requirements. This is followed by a discussion of privacy requirements and the approaches used for its elicitation. The last section identifies gaps in the literature to show why a new approach is required for mobile privacy requirements.

2.1 Privacy

In the following subsections we first explore a notion of privacy from a philosophical point of view and then relate it to modern information systems, especially cataloging the several types of privacy issues that emerge.

2.1.1 The notion of privacy

Privacy has been researched over many years and there is no single definition for it. In Warren & Brandeis (1890), privacy is described as the *'right to be left alone'*, while Westin (1967, p.7) sees it as a *claim* to free self-determination where he states:

> *Privacy is a claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated to others. Viewed in terms of the relation of the individual to social participation, privacy is the voluntary and temporary withdrawal of a person from the general society through physical or psychological means, either in a state of solitude or small group intimacy or, when among large groups, in a condition of anonymity or reserve.*

This definition of privacy has influenced many works that followed later and continues to shape our thinking even today. Margulis (2003) while providing a synopsis of Westin's privacy theory succinctly describes the four *privacy states* as the 'hows' of privacy, namely (i) *Solitude* - is being free from observation of others, (ii) *Intimacy* - refers to small group seclusion for members to achieve a close, relaxed, frank relationship, (iii) *Anonymity* - refers to freedom from identification and from surveillance in public places and for public acts, and (iv) *Reserve* - is based on a desire to limit disclosures to others; it requires others to recognise and respect that desire. In addition to this, Westin also posits four specific functions or purposes of privacy, which Margulis refers to as the 'whys' of privacy, namely (i) *Personal autonomy* refers to the desire to avoid being manipulated, dominated, or exposed by others (ii) *Emotional release* refers to release from tensions of social life such as role demands, emotional states, minor deviances, and the management of losses and of bodily functions. Privacy provides 'time-out' from social demands. (iii) *Self-evaluation* refers to integrating experience into meaningful patterns and exerting individuality on events. It includes processing information, supporting the planning process, integrating experiences and allowing moral and religious contemplation. (iv) *Limited and protected communication* - where limited communication refers to setting interpersonal boundaries and protected communication provides for sharing personal information with trusted others.

In addition to Westin's theory, the notion of privacy is heavily influenced by another theory by Altman (1975, p.24) where he describes privacy as *'the selective control of access to the self'*. Although it is similar to Westin, the difference here is that Altman suggests privacy as a socially *interactive process* where individuals actively control and *'regulate access to the self within social environments'*. In this regard, it is worth mentioning Margulis's (2003) summary of five *privacy properties* from Altman's theory.

The privacy properties are (i) privacy is a *temporal/dynamic process* of interpersonal boundary control: a process that regulates interactions with others i.e. how open or closed an individual becomes in response to changes in internal and external conditions (ii) privacy has *desired and actual levels of privacy*, (iii) Privacy is a *non-monotonic function* with an optimal level of privacy (desire=actual level), too little privacy (actual >desired level; e.g. crowding) and too much privacy (desired >actual level; e.g. social isolation), (iv) Privacy is bi-directional, involving inputs from others (e.g. noise) and outputs to others (e.g. oral communication), and (v) Privacy applies to both individuals and groups.

Several others have expanded on these two foundational works, for example Schoeman (1984) refers to privacy as *'degrees of social freedom through selective social engagement'*. More recently, Petronio (2002) refers to privacy as *'a dialectic process of boundary management'* dependent on a number of contextual factors, such as the background, age, circumstances, aptitudes, etc., of the parties involved. The dynamic nature and contextual dependency of privacy boundaries has been articulated from a socio-technological perspective by Palen & Dourish (2003). But they also state that privacy management is not about setting rules and enforcing them but rather, it is *'a continuous process where the users manage their boundaries between different spheres of action and degrees of disclosure within those spheres'*. These boundaries move dynamically as and when the context changes. In addition, these boundaries reflect tensions between conflicting goals and these boundaries occur at points of balance and resolution. Understanding privacy as a boundary regulation process is useful because current information systems have the ability to disrupt or destabilize the regulation of these boundaries.

The main difficulty with these philosophical concepts of privacy is that they are difficult to map to the requirements of software systems. Attempting to address this gap, Nissenbaum (2010) proposes a framework called *contextual-integrity* to describe privacy. This notion of privacy is particularly suited to context-aware information systems because privacy is stated as *'a right to appropriate flow of personal information'* where the social context and its related informational norms play an important role in protecting the privacy of a person.

Since privacy is a complex notion having several descriptions and interpretations, any privacy requirements engineering approach (including the ones that target mobile applications) will have to take this into account and address it appropriately. In the following section, we explore several types of privacy threats.

2.1.2 Privacy threats

In the real world, more often than not, privacy manifests itself as *violations* and *breaches*. While acknowledging privacy as a network of terms that have both similarities and differences, Solove (2008) has synthesized a *'taxonomy of privacy'* based on numerous real-life court cases which were filed for privacy violations in the US and other courts in the world. An interesting aspect of Solve's taxonomy is that it directly links to four basic groups of 'harmful activities', namely:

(i) information collection

(ii) information processing

(iii) information dissemination, and

(iv) invasion

Each of these groups consists of different related subgroups of harmful activities as shown in Table 2.1. The first group of activities that affect privacy is information collection. Under this group, *surveillance* is described as 'the watching, listening to, or recording of an individual's activities' and *interrogation* is described as 'an activity which consists of various forms of questioning or probing for information'.

The second group of activities involves information processing i.e. the way information is stored, manipulated and used. In this activity group, *aggregation* involves the combination of various pieces of data about a person. *Identification* is linking information to particular individuals. *Insecurity* involves carelessness in protecting stored information from leaks and improper access. *Secondary use* is the use of collected information for a purpose different from the use for which it was collected without the data subject's consent. *Exclusion* concerns the failure to allow the data subject's consent. Exclusion concerns the failure to allow the data subject to know about the data that others have about her and participate in its handling and use. None of the activities mentioned under this group involve gathering of data (because it is assumed to have been collected), focusing instead on the way data is maintained and used.

The third group of activities involves the dissemination of information. *Breach of confidentiality* is breaking a promise to keep a person's information confidential. *Disclosure* involves the revelation of truthful information about a person that affects the way the others judge her reputation. *Exposure* involves revealing things like nudity, grief, or bodily functions. *Increased accessibility* is amplifying the accessibility of information. *Blackmail* is the threat to disclose personal information. *Appropriation* involves the use of the data subject's identity to serve another's aims and interests. *Distortion* consists of disseminating false or misleading information about individuals.

The fourth and final group of activities involves invasions into people's private affairs. Invasion, unlike the other groupings, need not involve personal information. *Intrusion* concerns invasive acts that disturb one's tranquility or solitude. *Decisional interference* involves incursion into the data subject's decision regarding her private affairs.

Table 2.1: Solve's taxonomy of privacy threats

Information Collection	**Information Processing**
Surveillance	Aggregation
Interrogation	Identification
	Insecurity
Information dissemination	Secondary use
Breach of confidentiality	Exclusion
Disclosure	
Exposure	**Invasion**
Increased Accessibility	Intrusion
Blackmail	Decisional interference
Appropriation	
Distortion	

When describing his privacy taxonomy, Solove states that, in comparison to other groups, the broadest grouping of privacy problems is that of information dissemination. All groups of privacy problems, however, invariably relate to users' personal data and their interacting with other users of the software system. Therefore, the next section explores the nature of mobile users.

2.2 Mobile users

Every software system has *stakeholders*. System stakeholders are people or organisations who will be affected by the system and who have a direct or indirect influence on the system requirements (Kotonya & Sommerville, 1997). However, to elicit requirements, all *stakeholders* of the system have to be identified. One of the approaches to identify stakeholders is proposed by Sharp *et al.* (1999). System stakeholders can be of different types with very different goals, needs and expectations. They are broadly grouped into four categories as proposed by Macaulay (1994):

(a) Those who are responsible for its design and development (e.g. project managers, software designers etc.)

(b) Those with a financial interest, responsible for its sale or for its purchase (e.g. business analyst, marketing manager, buyer etc.)

(c) Those responsible for its introduction and maintenance within an organisation (e.g. software developers, training and support staff)

(d) Those who have an interest in its use (e.g. all classes of users - primary, secondary and tertiary).

In the context of mobile applications, the categories (a) and (c) refer to *software designers and developers* who are responsible for the design and implementation of the mobile system. From Lessig (1999), it is known that the technical architecture of system in one of the factors that influences the privacy of the end-users. Moreover, lack of facilities for privacy management in a software system can limit its use. Therefore, it is important that the design and architecture of new or existing software systems support the needs and expectation of their end-users.

While those in category (b) have responsibility for the success of the system, in the context of mobile applications, this could refer to business entities or *service providers* who provide mobile application infrastructure for their customers to use. In this case, customers are mobile users and the product being sold is the mobile service. Lessig (1999), points out that the 'market' is one of the factors that affect the privacy of end-users. The service providers have a vested interest in preserving their reputation in the market and this reputation is contingent upon ensuring that their customers or end-users do not suffer privacy violations when using their services. This again highlights the importance for service providers to build software systems that support the privacy of its end-users or mobile users.

The stakeholders in category (d) are users who perform tasks using the system and who are likely to experience the effects of the new system. The classes of users referred to in this category are of three types as defined by Eason (1988):*primary*, *secondary* and *tertiary* users. Primary users are those who are likely to be frequent hands-on users of the proposed system, while secondary users are occasional users or those who use the system through an intermediary and tertiary users are those who are affected by the introduction of the system or who will purchase it.

In the case of mobile applications, the user of the mobile device or *mobile user* is the primary user. Unlike desktop systems, a mobile phone is considered to be a personal device, carried and used exclusively by its owner. While the mobile users may not readily offer their phones to be accessed and used by others (secondary users), they may occasionally share information on their mobile phones with others who may be in close proximity to them (Mancini *et al.*, 2009).

In addition to being the primary user, the mobile user, is also the customer who pays for the software service or product (tertiary user). This places the mobile user in the unique position of being both an end-user and a customer of the software system and emphasizes the central role mobile users play in requirements elicitation and analysis.

Previous research by van Biljon *et al.* (2008) shows that mobile phones are used for various purposes such as *increasing safety and security* (e.g. calling for help in emergencies), *maintenance of social relationships* (e.g. sending messages and making phone calls to friends and relatives), *micro-coordination* (i.e. to organise personal and social activities), *accessing of information* (e.g. calendar, phone book, diary or Internet browsing), *entertainment* (e.g. games, chat, listening to music etc.) and *status enhancement* via brand or mode of the phone. This study and others (Mahatanankoon *et al.*, 2005) indicate that users of mobile phones considered maintenance of personal relationships and reporting of emergencies (safety & security) based on location as the most important functions afforded by the mobile system. This highlights the importance of eliciting requirements to protect the privacy of mobile users.

However, privacy attitudes can vary from one mobile use to another (i.e. privacy is very subjective and is based on individual perceptions and values). Broadly speaking, all system users can be considered to belong to either one of the three categories described in Westin's privacy surveys (Kumaraguru & Cranor, 2005):

(i) *The privacy fundamentalists* - these users view privacy as having an especially high value which they feel very strongly about and they usually have high levels of distrust. They tend to feel that they have lost a lot of their privacy and are strongly resistant to any further erosion of it. (Westin's survey in 2003 indicated that 26% of all American adults belong to this group).

(ii) *The privacy pragmatists* - these users also have strong feelings about privacy and tend to have medium to high levels of distrust. They are very concerned to protect themselves from the misuse of their personal information by other people and organizations. They weigh the value to them and society of providing personal information and they are often willing to allow people to have access to, and to use, their personal information - where they understand the reasons for its use, can see the benefits for so doing and when they believe care is taken to prevent the misuse of this information. (Westin's survey in 2003 indicated that 64% of all American adults belong to this group).

(iii) *The privacy unconcerned* - these have no real concerns about privacy or about how other people and organizations are using information about them. They usually have low to no levels of distrust. (Westin's survey in 2003 indicated that 10% of all American adults belong to this group).

While the above categories relate to system users in general, it is equally applicable to users of mobile systems. Although it is the privacy 'unconcerned' who are most vulnerable to privacy violations, it is the other two privacy groups that add up to 90% of users. Therefore, it is essential to elicit privacy requirements to manage the privacy needs and expectations of all users. In the following sections, requirements engineering approaches for mobile applications and privacy are explored.

2.3 Requirements Engineering

The term ‘Requirement’ has many meanings. Sommerville (2010) describes requirements as a statement of ‘what user services the system is expected to provide’, Young (2004, p. 1) enriches this initial definition by describing requirements as a ‘necessary attribute in a system, a statement that identifies a capability, characteristic, or quality factor of a system in order for it to have value and utility to a customer or user’. Goguen (1992) takes into account that systems are often placed in an operating environment which needs to be factored in, and therefore suggests requirements to be ‘properties that a system should have in order to succeed in the environment in which it will be used’. Unlike, these definitions which focus on the system, Jackson (1995b) states ‘requirements are about the phenomena of the application domain, not about the machine’, where *phenomena* refers to states and events, the *application domain* refers to real-world objects that will have to be observed and controlled, and the *machine* refers to both the hardware and software parts of the computer system we wish to build. This definition of requirement is attractive because it allows us to focus on ‘what’ control or behaviour is needed in the application domain without being restricted or constrained by ‘how’ it will be accomplished - descriptions which relate to *specifications* of the machine (or system). One of the main benefits in using Jackson’s view of requirements is that because it discourages any prejudging of potential solutions while stating the requirements, it fosters creativity without being overtly restricted by the limitations of the computer. In addition to these, it also provides a means to analyse and spot any mismatch between what the system was expected to achieve and how it is achieved.

Similar to requirements, there are divergent views on what *Requirements Engineering (RE)* actually involves and why it is considered to be a separate discipline. Some (Sommerville, 2010) state it as ‘the process of establishing the services the system should provide and the constraints under which it must operate’, others such as Davis (1990) suggest RE to be the analysis and documentation of both the user needs and the external behaviour of the system to be built. Zave (1997) describes RE as ‘a branch of software engineering which is concerned with the real-world goals for, functions of, and constraints on software systems. It is also concerned with the relationship of these factors to precise specifications of software behaviour'. Therefore, RE is a process of discovering the requirements for the system we wish to build and it is not so straightforward. Requirements must be gathered from the stakeholders - those who have an active interest in building and using the software system. But stakeholders (financial investors, software engineers, end-users, support staff, operators etc.) can be from diverse domains with divergent views on what the software system should achieve. Since the stakeholders may or may not be familiar with building software systems themselves, they might find it hard to articulate their requirements in a way which can be easily understood by the software engineers who must build them. Therefore, the process of systematically eliciting, gathering, describing and analysing requirements is considered to be an engineering process, rightly justified by Jackson (1995a) as:

> *In software development we build machines by describing them: the medium is text, but the products are engineering products...because the machines, and the useful purposes they serve, are complex, we must master the complexity by appropriate structuring and design of our descriptions. We must engineer not only the machines, but also the texts by which we describe them..*

At a lower level, this 'mastering of complexity', through the use of 'appropriate structuring and design of descriptions' translates into concrete steps in *processes* consisting of *techniques* and *methods* employed by software engineers for the type of system they are contracted to build. Nuseibeh & Easterbrook (2000) state: 'a process or an instance of a *process model* denotes an abstract description of how to conduct a collection of activities, where the behaviour of one or more agents and their management of resources are described. A *technique* prescribes how to perform a single activity, and also describes its outcome using a particular notation. A *method* not only provides a prescription for how to perform a collection of activities but also provides guidance on their use'. A method is a collection of techniques which are related and integrated to work coherently.

Further, the activities in RE are broadly classified by Nuseibeh & Easterbrook (2000) as falling under one of the following categories:

1. *Eliciting requirements:* finding out what problem needs to be solved for the stakeholders.
2. *Modelling and analysing requirements:* constructing abstract descriptions of requirements that are amenable to interpretation and analysis.
3. *Communicating requirements:* facilitating effective communication of the requirements among different stakeholders.
4. *Agreeing requirements:* convincing stakeholders that the represented requirements are appropriate.
5. *Evolving requirements:* managing changes in requirements when the environment and stakeholder requirements change.

The following sections describe how these activities are supported within the RE process.

2.3.1 Elicitation techniques

RE provides several techniques and methods for requirements elicitation, among them some appear to be more effective than others and have gained acceptance. Although not exhaustive, here we briefly mention the various techniques and methods that are available for software engineers to use. [Adapted from Goguen & Linde (1993); Maiden & Rugg (1996)].

Observation: The user's actual practices in the domain are observed. One example is requirements apprenticing (Reubenstein & Waters, 1991) which is based on this technique.

Introspection: This amounts to imagining what kind of system 'I would want if I were doing this job, using this equipment, etc.' (Goguen & Linde, 1993). Maiden *et al.* (2004) suggest that novel and creative requirements can be elicited when users are allowed to 'imagine' futuristic requirements.

Interviews (structured/unstructured): In *structured interviews*, the requirements engineer asks the stakeholder a list of prepared questions. *Unstructured interviews* on the other hand, do not make use of any prepared questions and offer greater flexibility.

As a result, they have been found to be good in eliciting relevant requirements from stakeholders.

Protocol analysis: Here, a subject is asked to engage in some task and concurrently talk aloud, explaining their thought process.

Card sorting: The stakeholder is asked to sort a set of cards into groups representing the name of some domain entity, and explain on the criterion used for sorting.

Laddering: A small set of probes are used to elicit the structure and content of stakeholders' knowledge. This technique makes an assumption that knowledge can be arranged in a hierarchical fashion (e.g. goal trees). The hierarchical knowledge elicited through laddering is represented in a standard format and is suitable for automated analysis.

Repertory grids: A system stakeholder is asked for attributes and values applicable to a set of entities. The output is in the form of an entity-attribute matrix which is amenable to automated analysis.

Brainstorming: In brainstorming, stakeholders are asked to generate as many ideas as possible, in some cases with the help of some tools. This technique has proved to be beneficial in eliciting high-level domain entities and questioning the underlying assumptions which are not explicitly manifest otherwise.

Prototyping: In prototyping, the stakeholders are asked to comment on a prototype of a physical model of a desired system. This technique is used to uncover deficiencies and flaws in the model and also to catch taken-for-granted issues.

Scenario analysis: A scenario describes the sequence of actions and events for a specific task the system is intended to accomplish. Generally, scenarios form a part of a use-case which captures a contract between the stakeholders of a system about its behaviour under various conditions (Cockburn, 2001).

Focus groups and JAD/RAD workshops: A focus group is a kind of group workshop where groups of individuals are brought together to discuss some topic of interest.

Document analysis: In this technique requirements are extracted from documents and other natural language text that govern stakeholder actions (Goldin & Berry, 1994).

Requirements derived from documents such as policies and regulations, are expressed as formal descriptions of rules which are readily amenable to analysis (Breaux *et al.*, 2006).

Model-driven techniques: This technique provides a specific model of the type of information to be gathered, and the requirements engineer uses this model to elaborate the requirements and drive the elicitation process. These include goal-based methods, such as KAOS (Dardenne *et al.*, 1993) and I* (Yu, 1997) or problem-oriented methods such as Problem Frames (Jackson, 2001).

Ethnomethodology: Contextual techniques such as ethnomethodolgy focus on gathering requirements through participant observations. Normally, a requirements engineer would spend extended periods of time observing users in their normal workplace to identify their interaction patterns. Early methods such as Soft Systems Methodology (SSM) (Checkland, 1999) relied on studying users in their natural environment. Similarly, recent user-centric approaches such contextual design (Beyer & Holtzblatt, 1998) make use of ethnographic principles to elicit requirements.

No matter which technique is adopted and used, any requirements gathered are likely suffer from various issues such as inadequacy, inconsistency, contradictions, ambiguity, noise, forward references and overspecifications (Meyer, 1985; Roman, 1985). To minimize and eliminate these problems, abstract descriptions of requirements are constructed as *conceptual models* to make them amenable to interpretation and analysis. These models can also be used further as an elicitation tool where modelling notations and partial models are used as drivers for further information gathering (Dardenne *et al.*, 1993).

2.3.2 Models

RE offers several ways to model requirements, each one varying in expressive and reasoning capabilities. Broadly they are grouped under the following categories as suggested by Nuseibeh & Easterbrook (2000):

Enterprise modelling: captures objectives of a system and describes behaviour of an organisation in which the system will operate. Enterprise modelling provides a means to understand an organisation's structure, its business rules, goals, tasks and responsibilities of all its constituent members.

Data modelling: describes what and how the information held by the system corresponds to the real world phenomena being represented. Data can be modelled using Entity-Relationship-Attribute (ERA) or object-oriented (OO) techniques. Further, data flow models can be used to represent data manipulations required within a system (Wieringa *et al.*, 2006, chp. 9).

Behaviour modelling: describes dynamic or functional behaviour of stakeholders and systems, both existing and those that are required in new system. In behaviour modelling, a model mirrors how work is currently carried out in a current physical system, and based on the analysis of this model, behaviour of a new system can be determined.

Domain modelling: provides an abstract description of the world in which an envisioned system will operate. Explicit domain models allow detailed reasoning and validation of domain assumptions. Domain models also expose how systems interact with the environment. Domain modelling also provides opportunities for requirements reuse within a domain.

Non-functional requirements (NFR) modelling: Non-functional requirements or quality requirements are desired properties of a system and include properties such as safety, security, privacy, reliability, usability, portability, testability etc. NFRs are usually represented as constraints and attributes of a system. Equally they can also be represented as detailed functions.

2.3.3 Languages for requirements

The above models are often expressed and described using formal, semi-formal or informal languages. It's important that requirements are not only modelled but also expressed in a way where they can be effectively communicated to all stakeholders. In this respect, RE provides several languages and notations from formal, semi-formal to informal languages. Again each of these languages have different levels of expressive, reasoning and communication capabilities.

One of the main benefits of using formal languages is that they offer a high degree of precision in the formulation of statements with precise rules for interpretation and analysis. For example, in Z (Spivey, 1992) or VDM (Bjrner & Jones, 1978), requirements specifications use mathematical notations to precisely describe those properties which a software system must have, without unduly constraining the way in which these properties are achieved. This kind of abstraction makes formal specifications useful because they allow questions about what the system does to be answered confidently, without the need to disentangle the information from detailed program code or from phrases in an imprecisely-worded prose description (in natural language). Further, the precise nature of formal specifications make it amenable to automated reasoning and analysis using tools (Heitmeyer *et al.*, 1996). The downside however is that they are difficult to construct and use.

In semi-formal languages such as UML (Rumbaugh *et al.*, 1999), requirements are described using graphical or visual representations such as structure diagrams (e.g. class and object diagrams), behaviour diagrams (e.g. use-case diagrams, activity diagrams, etc.) and interaction diagrams (e.g. sequence diagrams, communication diagrams etc.). In additional to the graphical representation of requirements, there are other semi-formal notations such as ‘descriptions’ (Jackson & Zave, 1993) which can be used to describe properties of a system and its environment.

Specialised languages can be used to describe particular system behaviour, for example, in the case of RELAX (Whittle *et al.*, 2009), the requirements language has constructs and operators designed specifically to capture uncertainty in self-adaptive systems. Similarly, Souza *et al.* (2011) use OCL^{TM} to formalise and express ‘awareness’ requirements that impose constraints on the run-time behavior of other requirements. In other cases, argumentation grammar has been tailored to specify conditions for security (Haley *et al.*, 2008) and privacy (Tun *et al.*, 2012).

2.3.4 Modelling methods

In RE, there are several modelling methods available, from structured to object-oriented methods, and from soft to formal methods. These methods provide different levels of precision and are amenable to different kinds of analysis.

Traditional structured analysis methods (DeMarco, 1978) used guided requirements modelling which focused on flows of data within a system. Some modelling notations such as *context diagrams* and *data-flow diagrams* are still popular and have been incorporated into current methods. With the advent of object-oriented programming, object-oriented analysis and design (OOAD) (Booch, 1999) gained popularity. In object-oriented analysis (OOA), requirements are examined and analysed from the perspective of classes and objects found in the vocabulary of the problem domain.

In KAOS (Dardenne *et al.*, 1993; van Lamsweerde, 2001), goals are captured at different levels of abstraction, showing the various objectives a system under consideration should achieve. In goal modelling, goals are not only used for eliciting and elaborating requirements but also for structuring, specifying, analyzing, negotiating, documenting, and modifying requirements.

Van Lamsweerde (2010, p.260) defines a *goal* as a prescriptive statement of intent that a system should satisfy through cooperation of its agents. Here, an *agent* is a system component that plays a specific role to satisfy a goal. Agents can be humans, devices, existing software components or new software components which are to be developed. During the elicitation process, goal modelling is particularly useful because it allows high-level organisational goals to be recursively refined leading to requirements that can then be operationalised. As goal models are specified formally, identifying and resolving conflicts in goals can assist in overcoming some of the problems associated with requirements specified in a natural language. The disadvantage of using goal-oriented RE is that goal models can become extremely difficult to use when modelling complex real-life systems.

While methods such KAOS may help to refine high-level goals to produce operationalisable requirements, they do make an assumption that the initial goals have been identified or are readily available. However, this assumption may not be true as the goals may have to be determined through other means. In order to address this, Anton (1996) proposes a goal-based requirements analysis (GBRAM) method which provides some guidance for initial identification and construction of goals from various types of artefacts such as flow charts, Entity-Relationship diagrams, process descriptions and even transcripts of interviews.

Software systems seldom operate independently, they often interact and depend on other agents in the environment to achieve their goals. This aspect of inter-component dependency is captured by some goal-oriented methods such as *i** (Yu, 1997). The i* method, specifically targets requirements elicited during the early-phase of the RE process where the aim is to model and analyse stakeholder interests and how they might be influenced by various system and environment factors. In i*, organisational actors are viewed as having intentional properties such as goals, beliefs, abilities, and commitments. Actors depend on each other for goals to be achieved, tasks to be performed, and resources to be furnished. By depending on others, an actor may be able to achieve goals that are difficult or impossible to achieve on their own. On the other hand, an actor becomes vulnerable if the depended-on actors do not deliver. In this method, actors play a strategic role as they are concerned about opportunities and vulnerabilities, and seek rearrangements of their environments that would better serve their interests.

In agent-oriented methods (Wooldridge & Ciancarini, 2001), a key abstraction is that of an agent who has specific properties such as ‘autonomy’ (agents encapsulate some state and make decisions based on this state), ‘reactivity’ (agents are able to perceive the environment they are situated in and are able to respond to environmental changes), ‘pro-activeness’ (agents exhibit goal-directed behaviour by taking the initiative), and ‘social ability’ (agents interact with other agents order to achieve their goals). Tropos (Castro *et al.*, 2001) is one such agent-oriented RE method which adopts and extends the i* notions of agent, goal, task and (social) dependency to model and analyse early and late requirements leading to architecture and detailed design. Secure Tropos (Mouratidis, 2009) is an extension of the Tropos methodology, which adds security concerns during the development process. Secure Tropos introduced security constraints, i.e. restrictions related to security issues, such as privacy, integrity and availability. Security constraints influence the analysis and design of information systems under development by restricting alternative design solutions, by resolving conflicting requirements, or by refining some of the system’s objectives. In Secure Tropos, secure goals are mainly introduced in order to achieve security constraints that are imposed on an actor or the system.

For some systems such as mobile applications, its physical context plays an important role. A distinct advantage of using Problem Frames approach (Jackson, 2001) is that it facilitates the physical context to be explicitly modelled and thus allows for validation of its domain assumptions. Unlike other approaches, Problem Frames focus on software development *problems* which take into account both the precise and imprecise aspects of context and provide means to describe them. In this approach, large problems are first structured (decomposed) into smaller commonly recognisable problems and sub-problems (similar to design patterns). After structuring, problem analysis is performed to identify any concerns and difficulties which will have to be addressed in order to solve the problem.

This section discussed various RE methods used in modelling and analysing requirements. Each of these methods have unique advantages and facilitate different types of requirements analysis. The following sections describe how some of these methods have been used in engineering mobile and privacy requirements.

2.4 Requirements Engineering for Mobility

We use the term *mobile applications (or 'apps')* to refer to software that runs on devices such as mobile phones or Smartphones. In general, mobile apps can be classified into two types - *native* and *Web-based*, depending on the programming technology adopted in their development. Native mobile apps are programmed and compiled to run on vendor specific platforms. For example, on an iPhone a native mobile app uses Objective C++, while on Android[1] mobile apps are developed in Java. Web-based mobile applications are simpler and use the standard Web protocol and technologies to provide functionality to their users.

Compared to other software, mobile apps are unique in the sense that they can access multiple devices and sensors which are on board a mobile device. Modern mobile devices[2,3] may have a *location sensor* such as GPS to determine location, a *light meter* to measure the amount of light in an environment, an *accelerometer* to measure motion,

[1] http://www.andoroid.com

[2] http://www.apple.com/uk/iphone/specs.html

[3] http://www.htc.com/uk/product/desirehd/specification.html

a *gyroscope* to sense orientation in 3D, a *proximity sensor* to detect the presence of nearby objects, a *video camera* and an *audio recorder*. In addition, screen displays of current mobile phones not only support vibrant colour and resolution but also act as a multi-touch interface enabling users to interact with their devices more easily.

While mobile devices are becoming as powerful as their desktop counterparts in terms of processing speed (CPU) and memory (RAM), they are constrained by several factors. Using the taxonomy provided by Roman (1985), here are some of the constraints that are applicable to mobile apps.

Interface constraints: Mobile devices have a small-form factor to facilitate mobility but this also restricts users' interaction with their device. The size of the screen display and keypad are particularly small when compared to traditional desktop computer systems. Therefore, understanding the user's interaction needs is considered to be another key requirement for a mobile application (Jones & Marsden, 2006).

Dependability constraints: Mobile apps are expected to be dependable. Dependability means a computer system is trustworthy and it can be relied upon for the services it delivers (De Florio & Deconinck, 2002). A dependable system is one that is able persist and to some agreed extent provide services even when it encounters some faults. The dependability property highlights how mobile applications must adjust to poor network connections and limited battery life. For example, mobile devices are carried by their human users from one location to another and when the network coverage is unavailable in some places the mobile apps will have to adapt accordingly to provide service continuity.

Operating constraints: Mobile devices have limited CPU, memory, battery-life, network bandwidth, network connectivity, etc. which impacts the way mobile applications operate. For example, mobile devices run on batteries that are recharged regularly which implies mobile applications cannot perform operations that expend the battery power quickly.

Performance constraints: These constraints refer to response time, workload, throughput, and available storage space on a mobile device, which relate to the users' productivity. For example, mobile users have short attention span and therefore mobile applications should take this into consideration.

Political constraints: Mobile apps are impacted by requirements from business and legal policies. In the context of mobile apps, these requirements may be enforced by a mobile network operator (Vodafone, O2, 3 etc.) or a platform owner (e.g Google Android, Apple iPhone etc.)

The constraints identified here are only a partial list and reflect those that cover a majority of constraints in any mobile app. In addition to these constraints, *physical security* and *privacy* can viewed as additional constraints on a mobile system. Unlike desktop systems that are stationary and can be protected in physically secure environments, mobile devices can be lost due to negligence or they can be stolen. In addition, people who are in close proximity to users may be able to read personal information on the screen display, leading to a loss of privacy.

As described above, mobile apps have to dynamically adapt to various challenges posed by the operating context. Building such applications is not easy and as it requires careful consideration of *what* users need and expect in different contexts and *how* those requirements can be translated into design and implementation. The following section explores the notion of context and context-awareness.

2.4.1 Context and context-awareness

Previously, the notion of *context* had been exploited in various *ubiquitous computing (Ubicomp)* applications such as Placeless Documents (Dourish *et al.*, 2000), Easy Living (Thomas *et al.*, 2000), Electronic Tourist Guide (Cheverst *et al.*, 2000), Conference Assistant (Dey *et al.*, 2001). Ubiquitous computing is a computing paradigm where computers are expected to be seamlessly integrated, become invisible and yet provide useful services Weiser (1991). Early ubiquitous applications researched by Weiser and others were primarily of two types, one to locate people and the other for shared drawings. The earliest forms of Ubicomp application was that of the Active Badge system pioneered by Want *et al.* (1992) at Olivetti Research Labs in England. The Active Badge system was used to locate people in an office environment. Members of staff wore badges that transmitted signals providing information about their location to a centralised location service, through a network of sensors. This was then used for several purposes such as automatic phone forwarding, locating an individual for a meeting, and watching general activity in a building (considered important for telecommuting).

Schilit *et al.* (1994) make context and *context-awareness* more explicit in their work. They state that context-aware systems have the capability to both examine and react to changes in the environment where context refers to one's location, people who are co-located (social situation) and other resource objects present in the surrounding environment. They contend that context encompasses more than just the user's location because other objects of interest are also mobile and changing. Thus, context is extended to include lighting, noise level, network connectivity, communication costs, communication bandwidth and even the social situation (e.g., whether you are with your manager or with a co-worker).

Further, Dey (2001) defines context as:

> *any information that can be used to characterise the situation of entities (i.e. whether a person, place or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves. Context is typically the location, identity and state of the people, groups and computational and physical objects.*

In the above definitions, context is proposed as a *representational problem* in which context is a form of information that can be encoded and represented in software system, is delineable, is stable and that context and activities 'within' it are separable. But Dourish (2004) enriches this view by formulating context as an *interaction problem* with the following properties:

1. Contextuality is a relational property that holds between objects or activities, i.e. contextually relevant to some particular activity.

2. The scope of contextual features is defined dynamically rather than being delineated and defined in advance.

3. Context is an occasioned property, relevant to particular settings, particular instances of action and particular parties to that action.

4. Context arises from the activity. Context isn't just 'there', but is actively produced, maintained and enacted in the course of the activity at hand.

Both these notions of context (as representation and interaction problem) are important for mobile apps, firstly because context has to be represented and reasoned upon so that it can support adaptive system behaviour. Secondly, users constantly interact with their mobile applications producing contextual changes and the system is expected to respond. This implies that mobile apps, in addition to observing changes to the operating environment, would also need to monitor the user's interactions because they indicate the current context of the user.

Similar to traditional desktop systems, mobile apps have functional requirements that describe a system's behaviour but in order to provide context-aware functionality, mobile apps monitor changes in their operating environment and adapt their behaviour accordingly. Salifu *et al.* (2007) define context-awareness as consisting of two types of system behaviour - *monitoring* and *switching (or adaptation). Monitoring requirements* define what applications must do to detect changes in their operating environment which may violate their goals/requirements and *adaptation requirements* refers to what applications must do, by adapting their behaviour, to restore the satisfaction of such goals/requirements. Further, *contextual variability* is defined as a space of variables whose different values require different application behaviours.

In RE, eliciting monitoring requirements is about establishing any *assumptions* made regarding the current state of the environment. For adaptation requirements, it is about eliciting a set of remedial *evolutions* available when mismatches develop between the assumptions and the current environment. The most challenging task however is that these assumptions can often be 'soft' requirements which are tricky to analyse and design (Fickas & Feather, 1995). Further, in mobile apps, its users' *personal requirements (goals)* must be taken into account. However, these personal goals and requirements may vary according to the time, location and context of the user *(user's context)*, leading to requirements for user specific customisation and adaptation (Sutcliffe *et al.*, 2005).

Given these complexities in first capturing context and then making sense of it to satisfy the user's needs in context-aware mobile apps, we explore how RE has approached such challenges while eliciting requirements, especially for mobile users whose location and operating environment is constantly changing.

2.4.2 Mobility requirements

Seyff *et al.* (2009b) developed a software environment called ART-SCENE to discover and document stakeholder requirements by walking through scenarios that are automatically generated from use case specifications. Their extended mobile version, called Mobile Scenario Presenter (MSP), used a mobile browser and wireless access to connect to the server-side ART-SCENE scenario system. In this system, a MSP user walked through scenarios illustrating instances of future system's behaviour while at the same time observing corresponding instances of current system's behaviour. What-if capabilities generated alternative courses for each event to enable users to follow-up and ask questions about abnormal and unusual behaviour in different contexts, thus leading to more complete requirement discovery. However, this work does not focus on privacy. In theory, this system has the potential to discover missing privacy requirements which are closely linked to the functional requirements in an existing system but asking the users to articulate privacy is a difficult task. Although this approach produces real reactions, the disadvantage is that it uses fictional scenarios which may not reflect users' reality as in an empirical study. Even in a hypothetical case where MSP was used to study mobile privacy, it may not be practical to ask users to type their privacy requirements into a mobile device as they may be in motion or constrained in using their devices. The authors suggest an improvement over this, by using audio recording of requirements, which may also be difficult for users to do depending on the sensitivity of context (for example, if they find themselves on public transport).

In another work, Sutcliffe *et al.* (2006) have proposed a requirements elicitation framework called 'PC-RE' to describe not only functions that meet people's goals but also characteristics of users and how users would like computer systems to achieve their personal goals. The framework accommodates matching requirements to individual needs: these change over time, therefore requirements evolve as people learn more about capabilities of the system they use. This is a scenario-based requirements analysis method, a development over their previous work, which used ethnographic techniques to investigate users' requirements in context. There they employed prototyping or Wizard of Oz techniques to evaluate initial designs with users and to refine their requirements. The authors highlight how personal requirements may be contextual and location sensitive;

for example, non-functional requirements such as privacy, security and information accuracy can interact with functional requirements such as information display. While the PC-RE method seems to place the emphasis on a user's goals with respect to spatial and temporal dimensions, their work does not focus on the privacy goals of end-users.

Seyff *et al.* (2008, 2009a, 2010) advocate the use of mobile devices as tools to elicit end-user's requirements. Their 'iRequire' approach uses a mobile tool to ask end-users to document their own requirements. iRequire enabled end-users to capture their needs in situ by following three elicitation steps.

(i) Capturing contextual information: end-users could document information about their environment,via different media types, such as audio, video, pictures or text in natural language. In addition to this, the mobile device provided built-in context-sensing capabilities to capture additional contextual information about the end-user.

(ii) Capturing end-user needs: end-users were required to document their needs using text-based descriptions as well as different media types such as audio recordings.

(iii) Capturing rationale and task: end-users could capture their requirement rationales by explaining why a requirement was important to them.

Although this approach supports elicitation of new user requirements in different contexts, it makes several implicit assumptions (i) users will always be proactive in providing inputs to support the elicitation process, (ii) users are always aware of the mobile privacy threats they face in different contexts, and (iii) users are able to rationalise and articulate their needs, or identify relevant contextual factors which influence them. However, these might be difficult to achieve for users.

Contextual Design (CD) (Beyer & Holtzblatt, 1998) is a user-centered design process which has been popular in industry and academia for several years. CD is a front-end product and systems design process rooted in the participatory design and user-centered design traditions. Central to CD is a field inquiry technique namely 'Contextual Inquiry' which is used for requirements gathering to understand the needs of the end-user. Contextual Inquiry uses a combination of shadowing and interviewing techniques and is based on the core premise: *'go where the customer works, observe the customer as he or she works, and talk to the customer about the work. Do that, and you can't help* *but gain a better understanding of the your customer'.* Thus, Contextual Inquiry helps to gather data by going out into users' natural environments and talking with them about their activities while they do them. This data is then consolidated and used for the derivation of requirements specifications and inventions of new products and functionalities. In order to validate requirements, new systems and concepts are tested using paper mock-ups and user interviews, again conducted in the field. CD has been adapted for eliciting requirements and designing mobile applications in a recent work where a mobile application for baseball fans was developed (Holtzblatt, 2005). The modified CD approach used a combination of interviewing, facilitated enactment, and customer co-visioning sessions to enable users involvement in the design and development process. While authors claim their CD approach provided an excellent framework for designing mobile consumer applications, their work did not focus on mobile privacy. Eliciting mobile privacy using this method will be problematic because shadowing of mobile users causes them to change their behaviour, thus invalidating the observed requirements. Moreover, privacy is a sensitive issue and often users are not be able to directly articulate their choices and decisions in an emerging context.

As discussed here, the current RE approaches and techniques help in understanding the needs of mobile users and the contextual factors that influence these needs. However, these approaches are rather limited in eliciting privacy requirements.

2.5 Requirements Engineering for Privacy

Historically, security and privacy disciplines focused only on malicious intruders and technological solutions rather than issues relating to perception, usability or organisational roles of end users (Adams, 1999b). This one-sided focus produced technical solutions that were both unusable and inappropriate. Recent approaches have understood this and sought to unpack users' perceptions of privacy and to provide guidance for system design that ensures the protection of users' personal information. However, the problem with many definitions of personal information is that they concentrate on the data itself rather than how it is perceived (Davies, 1997). Ultimately privacy, like trust, is heavily reliant on users' perception. It is not necessarily important how private or safe we are (although this is a vital component) but whether we perceive ourselves to be safe and private (Schneier, 2008). Mechanisms and policies will not address users' current concerns or potential concerns in the future, unless they are based on users' accurate perceptions of privacy. Thus, individuals' ability to control data about themselves is central to many privacy approaches. Sometimes called the 'operational privacy definition' (Bellotti, 1996), these control mechanisms provide end-users with capabilities to retain privacy via access control and feedback and in this regard Bellotti & Sellen (1993) specify four factors that affect control and feedback mechanisms, which are:

1. Capture - what kind of data is being gathered; voice, work activity, key presses etc.
2. Accessibility - who has access to the data.
3. Purpose - to what use the data is put.
4. Construction - what happens to the data (e.g., stored, manipulated out of context).

Although these high-level requirements give a general idea of what might be required to protect privacy from an end-user's perspective, software systems will also have to be compliant to local privacy and data protection laws. Generally, organisations which collect and process user data (depending on the nature of business) are likely to customise and enforce policies based on existing local laws and directives. Thus, privacy requirements are also derived from standards (guidelines & principles), privacy laws and organisational policies, which are explored in the following sections.

2.5.1 Privacy standards

A primary source for privacy requirements are standards such as OECD guidelines for privacy (OECD, 2010) and US FTC's Fair Information Practice (FIP) Principles (Federal Trade Commission, 2010). OECD guidelines specify eight principles for online privacy:

OECD guidelines

(i) *Collection Limitation Principle:* This principle states that there should be limits to collection of personal data and also data should be obtained by lawful and fair means and if possible with prior knowledge and consent of data subject.

(ii) *Data Quality Principle:* Personal data should be relevant to purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.

(iii) *Purpose Specification Principle:* This principle states that purposes for which personal data are collected should be specified at the time of data collection.

(iv) *Use Limitation Principle:* Personal data should not be disclosed, made available or otherwise used for purposes other than those previously specified and agreed with users (or authorised by law)

(v) *Security Safeguards Principle:* Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorised access, destruction, use, modification or disclosure of data.

(vi) *Openness Principle:* This principle states that there should be a general policy of openness about developments, practices and policies with respect to personal data. In other words there must be transparency in how data is used within an organisation.

(vii) *Individual Participation Principle:* This principle ensures that an individual has visibility and access to her data and if required is allowed to rectify, correct, complete or erase data held by others.

(viii) *Accountability Principle:* This principle holds a data controller as accountable for complying and enforcing the principles stated earlier.

In addition to these guidelines, FIP principles are another source of privacy requirements and a brief summary of them is presented here:

Fair Information Practice (FIP) Principles

Fair Information Practice Principles consist of five ‘core’ principles:

(i) *Notice/Awareness:* Similar to OECD's 'purpose specification principle', this principle states that consumers (or users) should be given notice of an entity's information practices before any personal information is collected from them because without notice, a consumer cannot make an informed decision as to whether or not and to what extent to disclose personal information. The notice should contain details on who is collecting the data, for what purposes the data is being collected, the potential recipients, how the data will be collected (if not obvious), consequences if the data is refused and the steps taken to ensure the data confidentiality, integrity and quality is maintained.

(ii) *Choice/Consent:* The second core principle of fair information practice is consumer (user) choice or consent. At its simplest, choice means giving consumer options as to how any personal information collected from them may be used. Specifically, choice relates to secondary uses of information i.e. uses beyond those necessary to complete a contemplated transaction. Such secondary uses can be internal, such as placing consumers on a mailing list in order to market additional products or promotions, or external, such as transfer of information to third parties. Normally, two types of choice/consent regimes have been considered: opt-in or opt-out. Opt-in regimes require affirmative steps by a consumer to allow collection and/or use of information; opt-out regimes require affirmative steps to prevent collection and/or use of such information. The distinction lies in the default rule that applies when no affirmative steps are taken by consumers.

(iii) *Access/Participation:* Access is the third core principle. It refers to an individual's ability both to access data about him or herself i.e., to view the data in an entity's files and to contest that data's accuracy and completeness. Both are essential to ensuring that data are accurate and complete.

(iv) *Integrity/Security:* The fourth principle is that data should be accurate and secure. In order to assure data integrity, data collectors can take several measures such as (a) using only reputed data sources, (b) cross-referencing data against multiple sources, (c) providing consumers access to data, and (d) destroying or anonymising untimely data. Security involves both managerial and technical measures to protect data against loss and unauthorized access, destruction, use, or disclosure. Managerial measures include internal organisational measures that limit data access and ensure that those individuals having access rights do not utilise data for unauthorised purposes. Technical security measures include use of encryption in data transmission and storage; limits on data access through use of passwords; and data storage using secure servers or computers that are inaccessible by modem.

(v) *Enforcement/Redress:* The core principles of privacy protection can be effective only if self-regulatory regimes include both mechanisms to ensure compliance (enforcement) and appropriate means of recourse by injured parties (redress).

Although OECD guidelines and FIP provide the basis for many existing software systems, there have been industry specific guidelines from self regulating bodies. For example, GSMA[1] has stipulated a set of privacy guidelines for application developers and mobile phone operators (GSMA, 2012). Similarly, the US Department of Commerce has come up with a draft proposal of 'Mobile Application Transparency' for developers of mobile applications (NTIA, 2013).

Most of these standards have evolved from privacy issues that have been reported in the past through court cases and litigation. These standards in the form of guidelines and principles were developed to provide mechanisms to mitigate these known privacy issues. We now look at how software engineers have exploited these standards and guidelines to derive privacy requirements.

The work by Langheinrich (2001) has derived privacy requirements from FIP. Their design guidelines are tailored for ubiquitous applications and have four properties: *ubiquity* (found everywhere), *invisibility* (the form factor makes it invisible), *sensing* (have capability to capture environment information) and *memory amplification* (continuously and unobtrusively recording every action, utterance and movement of users and their surroundings, feeding them into a sophisticated backend system that uses video and speech processing to allow browsing and searching through their past). The design guidelines they propose are: *notice*, *choice and consent*, *anonymity and pseudonymity*, *proximity and locality*, *adequate security and access* and *recourse*. While Langheinrich claims these guidelines are readily implementable, they do not address mobile applications specifically, although mobile applications do exhibit ubiquitous properties.

[1] http://www.gsma.com

Some researchers have modelled privacy as part of a wider modelling effort. For example, Yu & Cysneiros (2002) have modelled privacy as a non-functional requirement in i* using OECD guidelines. The i* framework was used to model the way agents interact with each other to achieve their goals. In addition to the framework, the authors have developed a catalogue (based on OECD guidelines) to guide software engineers through several alternatives for achieving privacy. Each alternative is modelled showing how it contributes to privacy and other requirements within this agent (or even within another agent). The catalogue in the i* framework models privacy as a special type of goal. They also show how one can model privacy concerns for each agent with several alternatives to operationalise them. Although this catalogue is useful, it is constrained by its organisation-centric approach with no reference to users' privacy preferences or control. Secondly, this framework suffers the same disadvantage as other approaches - it does not specifically address mobile privacy.

The PriS method (Kalloniatis *et al.*, 2007) uses eight categories of security and privacy principles to derive privacy goals which elicit and address privacy requirements in software systems. The PriS method's top-down approach using high-level principles are not only organisation-centric but also preempt the discovery of fine-grain privacy threats users might face when interacting with other system users. This is particularly difficult to know a priori in a mobile app whose operating context changes constantly. In a similar approach, the LINDDUN methodology proposed by Deng *et al.* (2011) models privacy threats to elicit privacy requirements for new software systems. They claim the method also provides guidance on selecting appropriate privacy-enhancing technologies accordingly. In LINDDUN, each letter of LINDDUN stands for a privacy threat type obtained by negating a privacy principle which in turn is derived from security properties, namely confidentiality, integrity, availability, authentication, authorization, and non-repudiation. This method suffers from the same disadvantage as the PriS method in that it is a top-down and organisation centric approach which will not able to capture privacy threats emanating from user-to-user interactions at run time.

2.5.2 Privacy laws

In addition to the FIP principles and OCED guidelines, local legislation is also a source for privacy requirements. In the UK, the Data Protection Act (DPA) (UK Government,

2013) forms the basis for data collection and its protection. The organisations collecting personal data are expected to comply with the data protection principles set out in the DPA. Unlike in the United States, in Europe, the OECD guidelines were incorporated into EU Directive (95/46/EC) (European Parliament, Council, 1995) to not only protect personal data in one member state but also when the data cross borders. A later EU Directive (2002/58/EC) (European Parliament, Council, 2002), specifically addresses location data of mobile users and how it may be processed. However, such EU directives are left for member states to implement. In the US, federal regulations enacted under Health Insurance Portability and Accountability Act (HIPAA) require members of the healthcare industry who use electronic information systems to protect the privacy of medical information. Similarly, the US has other privacy regulations that target specific industry sectors such as the Fair Credit Reporting Act (FCRA) to ensure accuracy in the maintenance and reporting of personal information by credit bureaus.

As stated earlier, regulations and guidelines require software engineers to specify, design, and implement systems that are accountable and compliant. Breaux & Anton (2008) have developed a methodology for extracting access rights and obligations from regulatory texts to ensure statement-level coverage for an entire regulation (e.g. HIPAA). The method provides guidance for software engineers to create stakeholder hierarchies, identify six types of constraints on requirements, manage cross references, maintain traceability, and resolve ambiguities. They present extensions to this methodology to acquire data elements and assign law-preserving priorities between data requirements to prevent improper information disclosures. The extended methodology provides critical assistance to engineers in navigating a very complex set of constraints and requirements as expressed in regulations. While this methodology is useful in extracting privacy requirements from existing laws and regulations (OECD, FIP and EU Directives), there are several drawbacks in using such top-down approaches. First, the existing laws and guidelines are organisation-centric and thus may not sufficiently address privacy concerns emerging from the end-user data sharing practices facilitated by novel software systems. Secondly, the ambiguity associated with the high-level requirements from laws and regulations can lead software designers to interpret it in ways that can lead to either inadequate privacy controls (Lahlou *et al.*, 2005) or over-engineering of privacy (Whitten & Tygar, 1999). Lastly, when using a top-down approach, it is difficult to know a priori if the system's implementation covers all privacy issues or not, especially in a mobile context where the operating environment can vary significantly.

2.5.3 Organisation privacy policy

The FIP principles and OECD guidelines have been used by organisations and business entities as high-level privacy requirements to formulate their own privacy policies to govern their use of customers' personal data. These privacy policies embody the privacy requirements that an organisation will implement and enforce in their data processing systems.

Breaux & Anton (2005b) have used a goal-mining technique (a process where goals are extracted from text artefacts) to analyze privacy policies. The identified goals were then used to reconstruct implicit requirements met by the privacy policies of companies in three health care industries: pharmaceutical, health insurance and online drugstores.

Massey & Anton (2008) compare and contrast two separate taxonomies created to improve the understanding of privacy and to ensure a common vocabulary between engineers and lawyers to help in establishing legal compliance. The first is the Anton-Earp requirements taxonomy, which is designed to increase software engineers' understanding of privacy related software requirements based on the pre-requirement goals extracted from an organisation's online privacy policies. The second is a legal taxonomy of privacy harms from Solove. In their comparison of both these taxonomies, they found the scope of Solove's study to be clearly broader in nature which factored in perspectives from different cultures of the world as compared to the Anton-Earp requirements taxonomy which focused on online privacy policies that were very specific to the US. This conclusion was reached because the Anton-Earp requirements taxonomy was able to map only 10 privacy harms out of the 16 categories in Solove's privacy taxonomy.

Cranor (1998) advocated a privacy policy model called the Platform for Privacy Preferences (P3P), which represented requirements using privacy policies of various online organisations. Many websites post privacy policies because they are required by law and their purpose is to inform visitors as to how the website will use any personal information being collected and what privacy choices are available. Although many websites post privacy policies, only a few visitors read them as several studies have shown that consumers find privacy policies time consuming to read and difficult to understand. In addition, readability experts have found that comprehending privacy policies typically requires college-level reading skills (Cranor, 2003). In addition, privacy policies have no standardised format, making it difficult to compare them. Consumers who do read these policies are also frustrated by the fact that they can change unexpectedly. In April 2002, the World Wide Web Consortium published the platform for privacy preferences (P3P) which specified a standard computer-readable format for website privacy policies. P3P-enabled Web browsers read policies published in P3P format and compare them with user-specified privacy settings. Thus, users can rely on their Web browsers to read and evaluate privacy policies on their behalf. Furthermore, the standardised multiple-choice format of P3P policies facilitates direct comparisons between policies and the automatic generation of standard format human-readable privacy notices. A major drawback of this framework is that it is extremely difficult to detect violations and to hold organisations accountable when they do not honour their privacy policies. For example take the case of JetBlue's privacy violation which was analysed by Anton *et al.* (2004). On 19 September 2003, JetBlue publicly acknowledged that in September 2002 it had provided travel records of five million customers to Torch Concepts, a private US Department of Defense (DoD) contractor, for an anti-terrorism study to track high-risk passengers or suspected terrorists (press release, 'JetBlue Retains Deloitte & Touche to assist the Airline in Its Analysis of Its Privacy Policy' JetBlue Airways, 22 Sept. 2002). Torch then purchased additional customer demographic information on those passengers from Acxiom[1], one of the largest data aggregation companies in the US. Using the information from JetBlue and Acxiom, Torch developed passenger profiles to identify possible terrorist suspects. This data transfer directly violated JetBlue's privacy policy ('JetBlue's Online Privacy Policy,' 24 Sept. 2003;)[2].

In addition, Earp *et al.* (2005) claim their content analytical research found many website statements that lie outside of the FIP principles. They developed seven categories of statements which make users vulnerable to potential invasions of privacy. They suggested that the FIP principles provide an incomplete basis for developing and analysing privacy-policy statements. Further, when developing their survey research, the authors

[1] www.acxiom.com

[2] www.jetblue.com/privacy.html

found that some FIP principles, surprisingly, were not valued very highly by Internet users. Based on these findings they conclude that protection statements from FIP principles may not buffer users against potential privacy vulnerabilities. This research shows that FIP principles are inadequate to deal with the privacy of website users, leave alone mobile application users.

Miyazaki *et al.* (2008) propose the use of a computer-aided Privacy Requirements Elicitation Technique (PRET) tool where organisations can fill-out a questionnaire and get a set of privacy requirements as output. This tool is similar to OECD's Privacy Policy Generator tool which makes use of a questionnaire to learn about the website's personal data practices and uses the answers to produce a preformatted draft policy statement; in this case the PRET tool outputs a set of privacy requirements. The drawback of this tool is that it does not show any validation of how effective these requirements have been in building systems which address privacy of end-users given that the laws and regulations are generally very ambiguous and their interpretations can be very subjective.

2.5.4 End-user personal privacy

In addition to gathering data and information, software systems are expected to cater for privacy perceptions of end-users (Davies, 1997). In this regard, Adams (1999a,b, 2000) has shown that most invasions of privacy are neither intentional nor malicious; rather, it was the designers who failed to anticipate *how data can be used*, *by whom*, and *how this might affect users*. Seeking to address this issue, Adams proposes a model for users' privacy perception in multimedia environments. The model helps to determine what information users regard as private, from whom, and in which context, and also depending on the privacy risks, how users might make trade-offs.

Adams model represents users as persons who transmit data about themselves directly (primary information - their work achievements, consumption habits, medical records etc.) or indirectly (secondary information - their image, voice or writings) and it identifies three major privacy factors that are key to users' perceptions of privacy:

(a) *Information Sensitivity (IS)* relates to users' perceptions of the data being transmitted. Information sensitivity is reliant on the users' judgement of the sensitivity levels of the information being broadcast. Here, the sensitivity levels are not binary (private / not private) but multi-dimensional with degrees of sensitivity.

(b) *Information Receiver (IR)* is users' perception of a person who receives and/or manipulates their data. A range of factors influence users' assessment of an information receiver and the level of trust between them, often based on their relationship and roles.

(c) *Information Usage (IU)* relates to users' perception of how and what their transmitted data is used for in a data exchange. The importance that a user attributes to perceived usage is key in privacy risk/benefit trade-offs that are made. Ultimately privacy is set within its surroundings and therefore users' perception of the context within which communications occur - specifically the technology, social groupings and national/international settings, play an important role.

Although this high-level model is useful in understanding user behaviour, it is difficult to implement this in a software system because of its vagueness. Lederer *et al.* (2002) make this point when proposing an improved model combining the Adams' (2000) and Lessig's (1999) contextual model (which include Law, Market, Norms, and Architecture at a place) to identify perceptual issues that can lead towards a technical mechanism for providing control and feedback. This is demonstrated by the prototype implementation of 'faces' in Lederer *et al.* (2003). Based on their experiences, Lederer *et al.* (2004) provide further design guidelines for designers of software systems which can help users to *understand* and perform *action* through feedback and control mechanisms. The design guidelines they propose are in the form of pitfalls that system designers should avoid, which are:

(i) *Obscuring potential information flow.* Designs should not obscure the nature and extent of a system's potential for disclosure. Users can make informed use of a system only when they understand the scope of its privacy implications.

(ii) *Obscuring actual information flow.* Designs should not conceal the actual disclosure of information through a system. Users should understand what information is being disclosed to whom.

(iii) *Emphasizing configuration over action.* Designs should not require excessive configuration to manage privacy. They should enable users to practice privacy as a natural consequence of their normal engagement with a system.

(iv) *Lacking coarse-grained control.* Designs should not forgo an obvious, top-level mechanism for halting and resuming disclosure.

(v) *Inhibiting established practice.* Designs should not inhibit users from transferring established social practice to emerging technologies.

These guidelines embody both feedback and control mechanisms which are stated at a rather high-level. Moreover, it is hard to measure whether these principles have been violated in a given system, unless it has been trialled beforehand with some real users. Take, for example, the first principle of 'obscuring potential information flow'; in many of the current mobile and social software systems the interconnections between applications are so complex that it may not be practical to make all potential information flows visible to end-users, which means trade-offs must be made upfront against existing features and implementation cost of the system. In such cases, it is difficult to verify whether the software system had fully complied to the design principle.

Altman (1975) models privacy as 'selective control of access to the self' regulated as dialectic and dynamic processes which include multiple mechanisms for optimizing behaviours. In other words, privacy is a boundary regulation process where people optimize their accessibility along a spectrum of 'openness' and 'closedness' depending on the context. To achieve this control of accessibility, users may employ verbal or paraverbal behaviours such as personal space and territoriality, and other culturally defined styles of responding. Building on this theory, Palen & Dourish (2003) propose privacy control as a continual management of boundaries between different spheres of action and degrees of disclosure within those spheres. As the context changes, these boundaries move dynamically and reflect tensions between conflicting goals; boundaries occur at points of balance and resolution. In this process, information technology performs multiple roles i.e. it is a part of a context in which a process of boundary maintenance is conducted, transformed, managed, and represented. Further, the authors propose the following three boundaries that exist under unifying genres of disclosure:

- The disclosure boundary (public and private)

- The identity boundary (self and others)
- Temporal boundaries (past, present and future)

Although privacy requirements for control and feedback have been researched, it is very likely these requirements will evolve as new types of user interfaces and technologies are made available for end-users. For example, unlike their predecessors, current mobile phones have large screens with touch interfaces which has revolutionised how mobile users interact with each other. In addition, the availability of numerous mobile applications with rich functionalities often sold for a very small fee or even given free, have opened up new avenues for data collection and manipulation. These factors when combined, increase the chances for privacy violations unless new feedback and control requirements are specifically elicited for mobile applications.

High-level goals and requirements from privacy standards (including guidelines and principles), laws and self-regulating organisation policies reveal that these are useful only if software developers can be properly guided to operationalise them. The main drawback of this approach is that they are ambiguous and are open to several interpretations. While standards, guidelines and laws evolve over time, they cannot catch-up with rapid improvements in technology-driven products such as mobile phones. Above all, high-level goals from various privacy laws and standards are organisation-centric and fail to take into account the privacy needs of end-users of novel mobile applications. On the other hand, privacy models developed through empirical means provide insight into user behaviour and facilitate understanding of the needs of end-users. However, they are difficult to design and implement in software systems because they are described at a high-level. Therefore, the challenge is to elicit privacy requirements which are not only precise but also satisfy end-users' privacy needs.

2.6 Requirements Engineering for Mobile Privacy

A news article titled *'Facebook surfing while sick costs woman job'* (Thomasson, 2009) described how a Swiss insurance worker lost her job after surfing popular social-networking Website Facebook while off sick. The woman said she could not work in front of a computer as she needed to lie in the dark but was then seen to be active on Facebook, which insurer Nationale Suisse said in a statement had destroyed its trust in the employee. The unnamed woman told the '20 Minuten' that she had been surfing Facebook in bed on her iPhone and accused her employer of spying on her by sending a mysterious friend request which allowed access to her personal online activity.

In another privacy case, a Connecticut man sued a local rental company 'Acme Rent-a-Car' after it used GPS (Global Positioning System) technology to track him and then fined him $450 for speeding three times (Lemos, 2001). Further, a recent newspaper article (Arthur, 2011) reported of a study where a group of security researchers had discovered that Apple's iPhone kept track of where users go and saved every detail of it into a secret file. The file contained the latitude and longitude of the phone's recorded coordinates along with a timestamp, meaning that anyone who could access the phone could discover details about the mobile owner's movements using a simple program. It seems for some phones, there could be almost a year's worth of data stored. One of the researchers pointing to the privacy implication of this data leakage, stated *'Apple has made it possible for almost anybody - a jealous spouse, a private detective - with access to your phone or computer to get detailed information about where you've been,'*. Although such cases are rare, they demonstrate that current mobile applications pose unique privacy challenges which need to be captured and analysed if privacy is to be addressed within mobile systems.

Privacy is a fine balancing act between what information is *monitored*, and the protections that are available against its *search*. As an enabler of both monitoring and searching, the architecture of technologies play a key role in enabling privacy (Lessig, 1999). Current architectures of mobile technologies enhance both monitoring and searching of its users' information. Mobile devices are packed with several sensors such as GPS, accelerometer, light meter, camera etc. which produce rich heterogeneous data. In addition, the availability of an Internet connection anywhere and at anytime, means search and disclosure of monitored information can be continuous. Further, the large screen displays of modern mobile devices can facilitate 'proxemic' disclosures (i.e. data disclosures caused by one's physical proximity to others) in public places (Mancini *et al.*, 2009). These architectural properties make management of *mobile privacy* far more challenging; the properties are:

(i) *Mobility of user*: A mobile device can be used anywhere including public places, where users may find themselves in close proximity to others who may be able to observe a user's interaction with their device or intercept information from their mobile screen display. For example, when users travel on public transport, strangers and fellow commuters may be able to read from the mobile screen display (Mancini *et al.*, 2009).

(ii) *Rich inferences from multiple datasets*: A user's interaction with a mobile device in different locations at different times can generate contextual data that can in turn be aggregated to infer information about their movements and actions, possibly without their knowledge. For example, a photograph taken from a user's mobile device can encode their current location co-ordinates; or data about the user's current activity can be aggregated with data about their current location, revealing what they are doing, with whom they are and where they are located. (For other type of inferences see (Eagle *et al.*, 2009; Madan *et al.*, 2010; Shilton, 2009)), and

(iii) *Dissemination of data in real-time*: A mobile device can be in continuous use, with information continuously being disseminated to allow users' movements and activities to be tracked down by others. Such contextual information can be automatically generated in a passive mode or proactively by users. For instance, 3rd parties (e.g. family, friends etc.) are able to track mobile users in real-time. On the other hand, mobile devices may also be used for tracking vulnerable users such as patients (Prociow & Crowe, 2010).

Establishing requirements that account for such complex interplay between factors described above requires a framework to systematically analyse each contextual aspect separately and also as a whole. In each contextual aspect, a framework should determine whether privacy is being violated or threatened in some form, then an analysis can be carried out to negate any emerging threats. We define mobile privacy requirements as:

> ***Mobile Privacy Requirements:*** *a set of constraints on a mobile computing application that enables appropriate flow of information depending on the user's context.*

where *flow of information* refers to information sharing practices relevant to a user's context (Sutcliffe *et al.*, 2005) and norms (Jirotka & Goguen, 1994, pp.107-139) that regulate it contribute to its *appropriateness.*

Eliciting requirements for mobile privacy is then about establishing:

(a) *Information flows:* extracted from users' information sharing practices

(b) *Context of information flows:* phenomena and properties of objects to monitor in a given environment along with its assumptions.

(c) *Privacy constraints:* to regulate information flows in a given context

When analysing mobile privacy requirements, it is likely that other requirements and constraints will be affected or in turn influence these requirements. For example, the political constraints drawn from an organisation's policy or legal document can influence how information flows are handled within a mobile app. So, RE will have to take cognizance of these and other constraints when analysing mobile privacy.

Having outlined mobile privacy, we explore the state of the art in eliciting, modelling and analysing mobile privacy using current RE techniques.

There have been several studies that have used an enthnomethodological approach to elicit privacy requirements for mobile applications. Khalil & Connelly (2006) report on an in-situ study of user privacy preferences and context information sharing patterns among different social relations. In another user study, Benisch *et al.* (2010) collected detailed location-privacy preferences and then identified best policies (or collection of rules granting access to one's location) for each subject and privacy-setting type. Tsai *et al.* (2009) evaluated users' risk and benefit perceptions related to the use of mobile technologies and privacy controls in location-sharing applications. The authors conducted an online survey of US Internet users (n = 587) to evaluate users' perceptions of their risks and benefits when using location sharing services (e.g. being stalked or finding people in an emergency).

In the PRiMMA project[1], two user-studies were conducted to elicit mobile privacy requirements by observing users in their natural environment. In the first study, Mancini

[1] http://primma.open.ac.uk

et al. (2009) used social-networking services on a mobile device to elicit privacy requirements. Using an experience sampling method, the authors were able to observe users in an non-intrusive way and gather information on how mobile users share their personal information with other users in different social-networking contexts. Based on this, deferred contextual interviews were employed to debrief and gather qualitative data on privacy behaviours of users in a socio-technical context. In another user-study, Mancini *et al.* (2011) set out to identify missing requirements and gaps in a mobile tracking application prototype. In this study, the authors were able to collect qualitative data on how mobile users made use of feedback and control mechanisms to address privacy concerns raised by location-tracking technology. Although, these empirical studies were a good source for collecting qualitative data which had mobile privacy requirements embedded in them, there were no mechanisms to systematically structure, extract and represent requirements from such data so that software engineers and designers could understand and implement them.

A review of the literature suggests that privacy requirements have been elicited mainly from legal and organisation policy documentation. However, for eliciting privacy requirements in mobile applications, mainly ethnographic methods were employed. A summary of RE techniques from the literature for mobile and privacy requirements is shown in Table 2.2 and Table 2.3.

This section explored different RE techniques that were employed for eliciting mobile privacy requirements and because of the complexity privacy introduces, ethnomethodology seems to be the most favoured approach. Privacy is dependent not only on users but also the social-technical context in which users' work is performed. Although, ethnography is known to be successful in capturing this rich socio-technical context, the qualitative data gathered from such methods tend to be very fuzzy/imprecise and difficult to analyse. However, others in the RE community recognise the importance of this qualitative data as they are a good source for deriving requirements (Jirotka & Goguen, 1994). Systematically deriving requirements from qualitative data entails several steps, most importantly it will have to be structured and made amenable for modelling and analysis.

In the literature, at least two requirements structuring approaches are available which could be potentially used for deriving requirements from empirical data. In the first

Table 2.2: RE techniques used for mobile privacy requirements

Requirements type	Observation	Introspection	Interviews (strd/unstrd)	Protocol analysis	Card sorting	Laddering	Repertory grids	Brainstorming	Prototyping	Scenario analysis	Focus groups/JAD/RAD	Document analysis	Model-driven	Ethnomethodology
Mobile Requirements	q		n,o,p,q					q	n,o	n,o	n			o,p,q
Privacy requirements														
Standards and guidelines												e,f	f,g,h	
Privacy laws												i		
Organisation privacy policy			m,l						k,m			j,k	j,l	
End-user privacy model	a,b,c		a						b,c		a		d	a,b,c
Mobile Privacy Requirements	u,t		r,s,t,u,v						u,v		v			r,s,t,u,v

Table 2.3: Reference key for Table: 2.2

References			
a	Adams (1999a,b, 2000)	l	Earp *et al.* (2005)
b	Palen & Dourish (2003)	m	Miyazaki *et al.* (2008)
c	Lederer *et al.* (2002, 2003)	n	Seyff *et al.* (2009b)
d	He & Anton (2003)	o	Sutcliffe *et al.* (2006)
e	Langheinrich (2001)	p	Seyff *et al.* (2008, 2009a, 2010)
f	Yu & Cysneiros (2002)	q	Holtzblatt (2005)
g	Kalloniatis *et al.* (2007)	r	Khalil & Connelly (2006)
h	Deng *et al.* (2011)	s	Benisch *et al.* (2010)
i	Breaux & Anton (2008)	t	Tsai *et al.* (2009)
j	Breaux & Anton (2005b)	u	Mancini *et al.* (2009)
k	Cranor (1998, 2003)	v	Mancini *et al.* (2011)

approach, Hughes *et al.* (1995) have proposed the use of three viewpoints (ecology, view and flow of work) to organise work done by users. This approach is similar to Contextual Design (CD) (Beyer & Holtzblatt, 1998) where contextual inquiry is used to apprentice with users and they suggest the use of five work models (flow model, sequence model, artifact model, cultural model and physical model). The first approach (Hughes *et al.*, 1995) is rather broad and is not particularly helpful as CD. On the other hand, CD provides clear steps and guidance to structure empirical data to derive requirements. However, the problem is that both of these approaches make an assumption that software engineers are able to apprentice or shadow users in their work. In the case of mobile apps, it is not practical to shadow or apprentice users in their natural environment. Secondly, for mobile privacy requirements, any shadowing of users will invalidate any gathered data because users are more likely to change their privacy behaviour in the presence of an ethnographer. Finally, the structuring techniques for both approaches do not specifically deal with privacy. Privacy is a complex notion, which requires expertise in being able to spot potential privacy violations from empirical data. Therefore, neither of these approaches, in their current form, are suitable for deriving mobile privacy requirements.

As mentioned earlier, in the PRiMMA project we collected data from two empirical studies which focused on mobile privacy of users. However, the data in its raw form cannot be analysed using current RE approaches to derive privacy requirements. In this regard, social sciences offers several data analysis approaches which can be broadly categorised into 'inductive analysis' and 'deductive analyis'. Inductive analyses refers to approaches that derive concepts, themes or models through interpretations of raw data whereas deductive analysis refers to analyses which tests data against prior assumptions, theories or hypotheses (Thomas, 2006). Some well known inductive approaches are thematic analysis (Clarke, 2007), grounded theory (Corbin & Strauss, 2008), phenomenology (e.g. Van Manen (1990)), discourse analysis (e.g. Potter & Wetherell (2004)), and narrative analysis (e.g. Lieblich *et al.* (1998)). The difference between these inductive approaches lies in the type of output they produce. In thematic analysis a list of most important categories and themes emerge whereas in grounded theory the emerging themes are further refined to build theories. Discourse analysis provides a descriptive account of perspectives and meanings in text and phenomenology produces coherent stories or narratives about user experiences. Since our aim is primarily to derive privacy requirements from raw data, we focus on thematic analysis to help in identifying specific themes or patterns relating to 'privacy threats and violations' reported by users. We describe the use of thematic analysis in the context of our work in Section 3.4 and 4.1.2.

2.7 Summary

Privacy requirements have been researched from different perspectives, however, only a few specifically concentrate on end-users' personal privacy. Additionally, not all privacy research take end-users' mobility and operating context into consideration. On the other hand, the RE approaches for context-aware and mobile computing take context into consideration but do not focus on privacy requirements of their end-users. Although, mobile privacy has been studied using ethnographic methods, the qualitative data gathered from such studies are hard to understand and model, and are not readily amenable to analysis. Moreover, current structuring and representation techniques for such data do not specifically address mobile privacy. Therefore, the aim of this research is to address this specific need by developing a problem structuring framework to first understand qualitative data and then model requirements from it to aid analysis. The framework will also help in representing requirements in a suitable format so that software engineers can use it to design and implement privacy enabled software systems.

3

The Privacy Facets framework

This chapter introduces and describes our analytical framework called Privacy Facets (PriF) which we will later use in a RE approach for mobile privacy. The PriF framework contains several tools to structure and analyse qualitative data and also to help in modelling and analysis of requirements. The first section provides an overview of the PriF framework and the second section describes a scenario which will be used to explain the key concepts. The third section defines concepts relating to privacy threats, privacy determinants and privacy concerns. The fourth section introduces privacy facets which provides components to be used in the structuring of qualitative data. This is followed by a section on coding heuristics which is used in identifying privacy sensitive contexts. The sixth section describes a set of data extraction rules for privacy threats. The seventh section specifies generic problem patterns found in information systems which are critical to analysing privacy related issues. The penultimate section describes extensions to an existing requirements language which is used in expressing privacy requirements. The final section summarises this chapter.

3.1 An overview of the PriF framework

There are several challenges an analyst must address in order to derive privacy requirements from qualitative data. First, the analyst should be able to identify sections of the qualitative data that might be potentially useful. This could be data which points to a user's privacy being violated or her reaction to specific privacy threats. Once such privacy sensitive sections are identified, details of the context that surrounds the user's interaction and behaviour will have to be unpacked, especially in relation to specific privacy threats and concerns. Finally, the analyst will need help in modelling information-flows pertaining to each privacy threat such that counter measures can be derived in the form of new privacy requirements. To this end, the PriF framework provides four analytical tools and an overview of each is provided below:

(i) Negative behaviour pattern (NPB) and negative emotional indicator (NEI): Qualitative data may contain not only users' experience but also their social interactions with other mobile users and actors in the environment, which may or may not be relevant to privacy. Therefore, the challenge of structuring this data relates to isolating those aspects which are relevant to extracting privacy requirements. For this, the PriF framework provides a set of user-centric heuristics *(NBPs & NEIs)* to identify privacy-sensitive contexts i.e. situations or settings involving privacy threats.

(ii) Facet questions and privacy determinants: After extracting a privacy related context from qualitative data, the social aspects of the user's interaction have to be understood. For example, the actors involved, their roles and relationships with users and the type of interactions that take place between them. To this end, the PriF framework decomposes the operating context into four privacy perspectives or facets: *information, information flows, actors* and *place* and provides a set of guiding questions called *facet questions* which elicit facet information from qualitative data. These questions help in identifying critical parameters or attributes that influence privacy in each facet.

(iii) Privacy threats and concerns: Parameters which influence privacy within a facet are likely to contribute to privacy threats, these are privacy violations that are likely to happen. When privacy threats are analysed in conjunction with requirements of an existing or a new software system, system failures are exposed as privacy concerns which will then have to be addressed in the future. The PriF framework provides a list of possible threats and concerns which can be identified from the qualitative data.

(iv) Information-flow problem patterns: Since privacy requirements are related to information-flows in a software system, the PriF framework provides problem patterns to model them. The first part of the information-flow relates to how information

is created and this aspect is captured under an *information creation problem pattern*. The second aspect, which relates to dissemination, is captured in an *information dissemination problem pattern*.

In addition to these four components, the PriF framework also includes a language to express privacy requirements. Here, an existing argumentation framework called 'Privacy Arguments' (Tun *et al.*, 2012) is extended and customised with the informal description language provided in Problem Frames (Jackson, 2001). Privacy arguments not only provides a structure but also acts as a glue binding specification of requirements. The following sections will describe each of these components in detail.

3.2 A scenario

In order to exemplify some of the key concepts introduced in the PriF framework, we use a fictitious scenario of a new (mobile) appointment booking system being developed for a local clinic called 'GetWell Surgery'. While the scenario is designed to be plausible, most importantly it models the type of realistic challenges an analyst or software engineer is likely to face when analysing privacy-related issues associated with a new software system and its end-users. The scenario is tailored to show how the different facilities provided within the PriF framework can be exploited to derive privacy requirements.

Alice and Bob are both registered with a local GetWell Surgery. Traditionally, appointments for a General Practitioner (GP) are booked via telephone or at the walk-in reception in the mornings. Patients had complained that they were unable to get through to the receptionist and make urgent bookings when the Surgery opened in the mornings. This was mainly because of the sudden influx of phone calls the Surgery had to handle as soon as it opened in the morning. In order to make the appointment booking more easier and transparent, the GetWell Surgery implemented a Mobile Appointments System (MAS) which allowed its users to make Surgery appointments via their mobile phones. Since MAS displayed all the currently available time slots for several GPs, patients could now choose their appointments and indicate their medical condition in real-time. Being a very sophisticated location-based service, the MAS didn't require its users to check-in when they arrived at the Surgery instead the system would automatically do the check-in by first determining their current location and then inferring their arrival at the Surgery. Further, the

MAS indicated this automatically inferred patient arrival status to the GP.

The software engineers who had designed the MAS were not fully proficient in engineering for the privacy needs of end-users and thus it was not a privacy-aware system. Patients were allowed to see other appointments, including their identifiable medical details. For example, when Alice booked her appointment using the MAS, another user Bob could see her appointment with a particular GP and when Bob tapped on Alice's appointment, he could even see Alice's medical condition which she had indicated in her GP appointment. In addition to this, the MAS also conflicted with the Surgery's mobile phone policy which required patients to switch-off their mobile phones when they arrived inside the Surgery building. However, when users (patients) switched-off their mobile devices to comply with the Surgery's policy, the MAS failed to identify them on arrival and check-in for their appointment.

During the initial months of using the MAS, several complaints were lodged by its users regarding privacy. Some even stopped using the MAS application, instead they reverted to using phone calls to make appointments. To address this problem, the GetWell Surgery decided to appoint a software engineer to investigate the privacy issues raised by end-users of the MAS and to report on changes required to make the system privacy-aware. The software engineer has decided to make use of the PriF framework for her analyses.

We use the above scenario in the following sections.

3.3 Privacy norms, determinants, threats and concerns

Although, there are several definitions for privacy depending on the context of use (Solove, 2006), here we adopt the version proposed by Nissenbaum (2010) where privacy is stated as *"a right to appropriate flow of personal information'*, also known as contextual integrity (CI). From a software engineering perspective, we adopt CI and restate privacy as users having the *right to information-flows*, in which *users approve their flow of personal data and information.* This appropriate information-flow within the context of a software system is referred to as a *privacy norm.*

Privacy norms can be either positive or negative. Positive privacy norms indicate the type of information-flows a software system should support whereas negative privacy norms are information-flows which should be avoided in order to protect users' privacy. Privacy norms are complex because they consist of information-flows that have subcomponents which are inter-linked with each other. Moreover, descriptions of these information-flow components are likely to be very verbose leading to ambiguity and misunderstanding. In order to address this, we formalise some of the key concepts using the notations shown below. Note that these notations are meant to be be 'light-weight' and avoid any overtly complex formalisations that are difficult to use. Thus, we formalise privacy norms as:

$$Privacy\_Norm(PN) = ALLOW[+F_i] \mid DENY[-F_i] \in W \tag{3.1}$$

In the above, F_i represents information-flows, the '+' sign indicates a positive information-flow while the '-' sign indicates a negative information flow. In a privacy norm, the 'appropriate flow of information' is achieved when positive (+) privacy norms are allowed (ALLOW) to execute while negative privacy norms are prevented from any type of execution (DENY). The '|' sign indicates that either or both conditions apply. W represents the context.

Modelling information-flows is not new. Denning (1976) describes a mathematical framework for formulating requirements for secure information-flows in software systems. Barth *et al.* (2006) provide formalisations of information-flows to support CI but they introduce additional concepts for messaging and knowledge states of agents which can increase the level of complexity. However, our information-flow model is based on CI and has some resemblance to Barth et al. It is symbolically represented as:

$$Information\_flow(F_i) = \pm\langle\ A_s \rightarrow A_r, I_a(A_u), L_p \| G_t\ \rangle$$
$$where\ A_u, G_t \neq \varnothing;\ A_s, A_u, A_r \in W \tag{3.2}$$

In the above, a user/sender A_s is sending information (or information attributes) I_a about subject A_u to receiver A_r complying with goals and purposes G_t at place L_p; where A_s, A_u and A_r correspond to roles in a given context W. Context W refers to the setting or arrangement of entities (i.e. a person, place or object) in the environment that are relevant for the flow of personal information in the software system. Context can also refer to location, identity and states of people, groups and computational and physical objects (Dey, 2001). Here, the goals govern the flow of information,

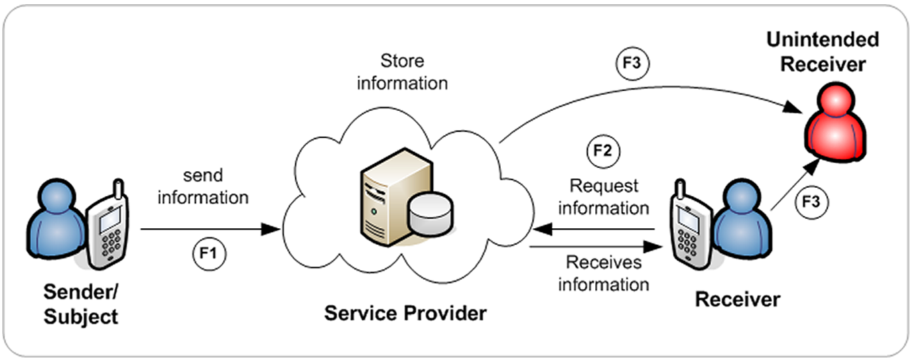


Figure 3.1: Information flows considered by PriF framework

encompassing the purpose for which the information is being transmitted.

As pointed out by Madsen *et al.* (2006), there are several types of information-flows in a software system that are relevant to addressing privacy, but this work concentrates only on those that are critical to addressing privacy in mobile applications which support peer-to-peer user interactions (i.e. personal privacy). This work does not focus on privacy issues emanating from information-flows relating to intermediate service providers or organisations. Figure 3.1 shows a generic architecture containing three information-flows - information created and sent to a service provider (F1), stored information is requested and is sent to a receiver (F2) and information sent to unintended receivers by either the service provider or the receiver (F3). The unintended receivers can also refer to actors who are co-located and in close proximity to the sender or receiver and are able to access the information without making a request to the software system.

When a privacy norm (positive or negative) is unsupported or its implementation is weakened by a software system, it can cause harm to users, leading to a *privacy violation*. As described earlier, an information-flow consists of several parameters $(I_a, \{A_s, A_u, A_r, \}, G_t, L_p)$, among them there might be one (or more) parameters which can strongly influence the privacy norm and these are called *privacy determinants*. *Privacy threats*, which map unsupported privacy norms in a software system to harms a user can suffer, can cause privacy violations when realised. Privacy threats are described as a tuple consisting of an unsupported privacy norm and its associated harm:

$$\begin{aligned} Privacy\_Threat(T) &= \langle \otimes[PN],\ H \rangle \\ & where \otimes [PN] = DENY[+F_i] \mid ALLOW[-F_i] \in W \end{aligned} \tag{3.3}$$

In the above, the unsatisfiability operator '⊗' indicates that the unsatisfiable (unsupported) privacy norm PN leads to the realisation of a set of privacy harms H. The 'inappropriate' flow of information in a privacy threat is caused when positive (+) information-flows are prevented (DENY) and/or negative information-flows are allowed (ALLOW) to execute, achieving exactly the opposite effect of privacy norms.

Privacy concerns describe the gap between current requirements model (or its implementation) and the identified privacy threats. Privacy concerns actually indicate the ways in which the privacy norms in a software system are unsatisfied (or unsupported), thus leading to privacy threats. Privacy concerns can be symbolically represented using the following notation:

$$Privacy\_Concern(PC) = \langle\ conditions\_for\ \otimes [PN],\ H\ \rangle\ where\ H \neq \varnothing \tag{3.4}$$

Privacy requirements address privacy concerns by ensuring the privacy norms of a software system are supported/satisfied. Privacy requirements achieve this by introducing suitable feedback and control facilities such that users have better control over information-flows that are linked to specific privacy threats. A semantic model of these privacy concepts is shown in Figure 3.2.

When privacy violations takes place, the most visible aspects are the injuries or *privacy harms* people suffer. Since there are at least sixteen types of privacy threats (see Table 2.1), it is only reasonable to assume there could be several types of privacy harms arising when those threats are realised. Solove (2008, p.174) broadly discusses these harms as: *individual & societal harms, financial loses & property harms, reputational harms, emotional & psychological harms, relational harms, vulnerability harms, chilling effects* and *power imbalances.* Apart from Solve, there are also others who broadly classify privacy harms. For instance, Nissenbaum (2010, p.78) categorises harms under:

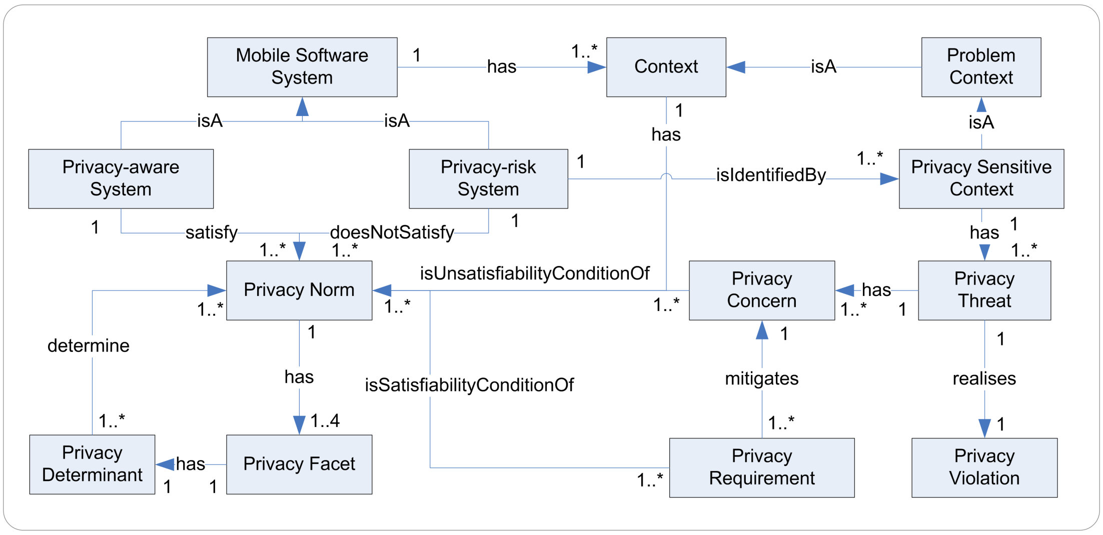


Figure 3.2: Semantic model of key concepts (with cardinality)

informational harms, informational inequality, informational injustice and *encroachment on moral autonomy.* There are two problems in using these classifications. Firstly, they are too broad to be useful for software engineers who might see very specific instances of these harms in user reports/qualitative data. Secondly, these harms may not directly relate to the harms suffered by users of mobile applications. In this regard, we incorporate the privacy harms suffered by mobile users from a previous empirical study (Mancini *et al.*, 2009). Combining these three contributions, we refine and present a set of ten privacy harms (see Table 3.1) tailored for mobile applications.

Since privacy harms are consequential effects of privacy violations, it was natural that this taxonomy is mapped to relevant privacy threats. Using a similar approach as before, we synthesized ten privacy threats applicable to mobile applications from the works of Solve, Nissenbaum and Mancini et al. as shown in Table 3.2. The main aim of producing these taxonomies was to comprehensively represent as many privacy harms and threats as possible in a tangible and recognisable form which software engineers can comprehend and identify during analysis.

Some of these threats are combined from over-lapping concepts. For instance, Solove makes a distinction between *'exposure'* and *'disclosure'*. However, we collapse them into a single threat - *'exposure'* since it represents a superset which includes the properties of *'disclosure'* (Solove, 2008, p.105). Similarly, we use a single threat *'power*

Table 3.1: A taxonomy of privacy harms

ID	Harms caused by privacy violations	Source
H1	Identity theft	[1][2]
H2	Financial & property loss	[1]
H3	Loss of reputation / trust	[1]
H4	Emotionally affected (anxiety, stress, fear etc.)	[1][3]
H5	Biased decision / discrimination	[1][2]
H6	Loss of freedom / autonomy	[1][2]
H7	Loss of anonymity	[1]
H8	Break-down of relationship	[1]
H9	Embarrassment / humiliation	[1][2]
H10	Compromise of physical safety / security	[1][3]

[1] Solove (2008, p.174), [2] Nissenbaum (2010, p.78), [3] Mancini *et al.* (2009)

imbalance' to refer to both *'increased accessibility'* from Solove and *'power imbalance'* from Nissenbaum as these relate to identical privacy harms.

The privacy harms and threats identified in these taxonomies are by no means exhaustive or complete. In other words, they are very likely to be revised and improved upon. The harms associated with the threats are only indicative, as there could be additional combinations of these harms linked to each of these privacy threats.

Privacy concepts stated so far can be explained further using the scenario described earlier (in section 3.2). In the scenario, the MAS supported a feature where Alice and Bob could put in their personal data into the system, including their medical condition when making an appointment. Here, one *privacy norm* refers to the appropriate information-flow between the Alice (or Bob) and the GP. While Alice might approve the GP to view the appointment which had her medical condition on it, she would not expect her appointment information to be seen by a stranger i.e. Bob. When the MAS allowed Bob to see this information against her wishes, it was a *privacy violation* realising a *privacy threat (T2).* The resulting embarrassment for Alice is a *privacy harm (H9).* The fact that the MAS could not stop this privacy threat from being realised is

Table 3.2: A taxonomy of privacy threats

ID	**Privacy Threat**	**Inappropriate information-flows (Privacy norms)**	**Harms to the user**	**Source**
T1	Identification	Subject's personal information is revealed	H1, H2	[1]
T2	Exposure	Personal/sensitive information received by unintended recipients	H5, H7, H8, H9, H10	[1]
T3	Surveillance	Receiver makes frequent requests for information about the subject	H4, H6, H10	[1]
T4	Aggregation	Receiver combines datasets produce a new type of information without the subject's approval	H5	[1]
T5	Misinformation	Inaccurate or insufficient level of information about the subject is transmited	H3, H4, H5	[1]
T6	Breach of trust	Receiver forwards the information to others contravening the subject's terms and conditions	H3, H4	[1]
T7	Power imbalance	Receiver uses information to control the subject	H6, H8	[1][2]
T8	Cross-contextual information flow	Information belonging to a particular context may be used in another context	H3, H5	[2]
T9	Proxemic access	Unintended receivers can access information because they are in close physical proximity to the sender.	H3, H6, H7, H9	[3]
T10	Intrusion	Information flow disturbs receiver's tranquillity	H4, H6	[1]

[1] Solove (2008, p.104), [2] Nissenbaum (2010, p.78), [3] Mancini *et al.* (2009)

a *privacy concern.* The new *privacy requirements* will state what the MAS should do to protect Alice in the future, for example, one of the privacy requirements could be to anonymize Alice's personal details so Bob cannot identify her. The information-flow in this privacy norm has Alice as sender/subject and the GP as receiver. The appointment information in the information-flow has personal data in it, for example, Alice's name and address, her date-of-birth and medical condition. Among these, there were two attributes that significantly influenced the privacy norm - the personal details of Alice and the roles of GP and Bob. If the information wasn't identifying Alice (assume her name and address was substituted with a large random number), she wouldn't mind Bob seeing her appointment details. On the other hand, if Bob was also a GP at the GetWell Surgery, she wouldn't mind her personal details being viewed by Bob. Therefore, in this case, the appointment details and the roles of the information receivers are *privacy determinants.*

3.4 Thematic codes for structuring data

Under qualitative data analysis (QDA), 'thematic analysis' is a method for identifying, analysing, and reporting patterns (themes) within data (Braun & Clarke, 2006). In thematic analysis, raw data is often structured by tagging or assigning sections of it to specific categories and sub-categories of concepts, a process referred to as *thematic coding* or just *coding.* In coding, several passages or sections of qualitative data are identified and are then linked with a name or label that correspond to a concept i.e. a *code.* All the data which refer to a similar concept or which exemplifies a single concept are often coded with the same label or code. Thus, coding is a way of indexing or categorising the data in order to establish a framework of thematic ideas (Gibbs, 2007). While some refer to thematic coding simply as 'coding' (Corbin & Strauss, 2008), others refer to it as 'inductive coding' (Thomas, 2006). (we use the term coding and tagging synonymously).

In QDA, codes are developed using either *inductive* or *deductive* approaches. In inductive or data-driven approach, thematic codes emerge from analysis of qualitative data. The deductive or concept-driven approach relies on codes developed through other means such as research literature, previous studies, topics in an interview schedule, etc. In reality neither of these methods are exclusive; analysts may go backwards and forwards using both these approaches in order to develop and refine their codes (Gibbs, 2007, p.44).

In the PriF framework, qualitative data is structured through the use of pre-defined codes (shown in Table 3.3) for two main purposes:

(a) To identify privacy-sensitive contexts, and

(b) To elicit privacy facets knowledge, which is critical for deriving specific privacy threats.

Privacy-sensitive contexts are identified when analysts spot either negative emotional indicators (code: NEI) or negative behaviour patterns (code: NBP) exhibited by the users while interacting with software systems. Once the data is coded for either of these two categories, the analyst will explore this data closely to see if the details of the privacy-sensitive context are linked to any one of the privacy facets - information, actor, information-flow and place. Both these categories (privacy-sensitive context & privacy facets) play a critical role in identifying specific privacy threats experienced by users. In the PriF framework, privacy threats are determined by matching the codes applied for privacy-sensitive contexts and privacy facets. For example, if a section of data was first coded for negative behaviour/emotion (codes: NEI or NBP) and then for information-attribute (code: I-ATTR) in privacy facets, applying certain extraction rules (described later in section 3.7), the data would indicate that users have experienced a 'misinformation' privacy violation or threat.

The categories relating to negative emotions and negative behaviour were developed through inductive analysis of qualitative data which is taken from a previous study of Mobile Facebook (MFb) users and is presented in Appendix A. Initially, there were no codes. First, raw data in the form of interview transcripts were analysed for specific episodes where users may have experienced a privacy threat or violation. It emerged that users often exhibit negative emotions in such situations. For example, users would state 'I don't feel comfortable' or 'I am worried' when they were experiencing a privacy threat/violation. We identified and labelled several of these negative emotions such as 'worry', 'dislike', 'uncomfortable', etc. These codes were captured and aggregated under the 'negative emotional indicator' in Table 3.3. If users' negative emotions could

Table 3.3: Pre-defined categories and sub-categories for QDA

	Code	Categories	Code	Sub-categories
Privacy-Sensitive Context	NEI	Negative emotional indicator	CONCRN	Concern
			DSLIKE	Dislike
			WORRY	Worry
			ANXIOS	Anxious
			UPSET	Upset
			DOUBT	Doubt
			UNHPPY	Unhappy
			UNCOMF	Uncomfortable
			OTHER	Other
	NBP	Negative behaviour patterns	DISUSE	Disuse
			WORKA	Work around
Privacy Facet (Information, Actor, Info. Flow, Place)	I-TYPE	Information type	PERSL	Personal info.
			SENSE	Sensitive info.
	I-PURP	Purpose of information		
	I-MODE	Information collection mode	AUTO	Automatic mode
			MANUAL	Manual mode
	I-ATTR	Information attribute	ACCU	Accurate
			COMP	Complete
			FRESH	Fresh
			ONTIME	OnTime
	ROLE	Role	RELTN	Relationship
			RESPB	Responsibility
	I-FLOW	Goal and purpose	SNDSUB	Sender-subject
			SNDRCV	Sender-receiver
			3PARTY	3rd party
	PLACE	Place	LOCATN	Location
			ETIQTE	Etiquette

not be mapped to any of the existing codes, the analyst can make use of the ‘other’ option from the code book. In addition to negative emotions, we also found that users exhibited two different types of negative behaviour when they perceive a risk to their privacy. They would switch-off and not use their mobile device (system) or they might use alternative means to share data instead of using the software system’s facilities. These negative behaviours were respectively labelled as ‘disuse’ and ‘workaround’ and aggregated under the category ‘negative behaviour pattern’ as shown in Table 3.3. These two categories are elaborated in section 3.6.

Unlike the above two categories, the categories relating to privacy facets were developed through a deductive approach. The theorectical privacy model proposed by Nissenbaum (2010) and others (Mancini *et al.*, 2009) heavily influenced the type of categories and sub-categories that were chosen to be under privacy facets (see Table 3.3). Here, the analysts are assisted in their coding by a set of guiding questions provided under each facet called *facet questions*. When answers are found for each of these facet questions, they are assigned codes which correspond to the question. While each of these categories and sub-categories are described in detail in Section 3.5, those categories/sub-categories pertaining to identification of privacy-sensitive contexts are described in Section 3.6. The corresponding codes are usually of the form: [CATEGORY(SUB-CATEGORY)].

Section 3.7, which is on extraction rules, describes the relationship between different categories and how satisfaction of certain rules indicate the presence of specific privacy threats. In the next chapter titled ‘Distillation’, we demonstrate how these codes are applied in conjunction with extraction rules to derive privacy threats from qualitative data.

3.5 Privacy facets

When studying privacy of mobile users, the context has to be systematically unpacked and analysed. However, this is often difficult due to users’ mobility which produces constant changes in the operating environment. Context is a complex notion having more than one dimension. It is described in several ways but generally most works can be classified under one of the two groups. The first group views context as proposed by

Schilit *et al.* (1994), where the authors define context as a representation of properties in the environment:

> *...where you are, who you are with, and what resources are nearby. Context encompasses more than just the user's location, because other things of interest are also mobile and changing. Context includes lighting, noise level, network connectivity, communication costs, communication bandwidth, and even the social situation; e.g., whether you are with your manager or with a co-worker.*

Similarly, Dey (2001) factors in users' interaction when he states:

> *Context is any information that can be used to characterise the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.*

The above notion of context is slightly different to the second group where Rogers (2006) and Dourish (2004) propose context as *'embodied interaction'* where users negotiate and evolve systems of practice and meaning in the course of their interaction with information systems and the operating environment. A study by Tamminen *et al.* (2004) shows that mobile context is about how situational and planned acts intermesh in navigation, how people construct personal and group spaces, and how temporal tensions develop and dissolve. Irrespective of the group, these definitions indicate that context consists of several real world domains whose dimensions, properties, phenomena and state must be systematically analysed if we are to build useful context-aware software systems. The additional challenge in our case is that this knowledge and understanding of users and their (mobile) operating context must be derived from qualitative data.

There are several approaches available in order to understand and make sense of the operating context. For instance, Finkelstein *et al.* (1992) suggest the use of 'viewpoints' as one way to systematically organise and structure domain knowledge from different stakeholder perspectives. The authors describe a viewpoint as a 'view' or 'perspective' which an actor maintains. The Problem Frames approach (Jackson, 2001) uses context diagrams to explicitly model the physical context and thus locate and bound software problems.

With the ubiquitous nature of networked and distributed computing, (as in mobile computing), Hughes *et al.* (1994) have proposed the use of ethnography as a method to capture the social context (collaborative work and activities). The authors also demonstrate the use of three viewpoints in (Hughes *et al.*, 1995). The viewpoints are as follows: (i) *The setting of work (ecology of work):* represents spatial distribution of workplaces in terms of its participants, the work they do and the local resources they use. The purpose of this viewpoint is to provide a sense of 'where the work takes place' (ii)*Social and organisational perspectives on work (views of work):* tries to bring out the day-to-day experience of work from the point of view of various actors within a setting, and (iii) *Work flow (Flow of work):* focuses on the sequences of work activities, information flows, and so on.

Similarly, Beyer & Holtzblatt (1998) propose a framework called Contextual Design (CD) which is also specifically designed to analyse social and physical context, where work is modelled into four types (i) *The flow model:* describes the individuals who do work, their responsibilities, their roles and groups. This model also includes work-flows, the artefacts they use or manipulate, the actions performed, the places people go to and any breakdown in communication that might take place, (ii) *The sequence model:* this model captures the intent of a sequence of actions, triggers that cause a sequence of actions, steps of action and the order connecting those steps, (iii) *The artifact model:* captures the information contained in an artefact, its parts, structure, annotations, presentations, conceptual distinctions, its usage and the problems associated with its use, (iv) *The cultural model:* deals with influencers who affect or constrain work and the extent of their influence on work, and (v) *The physical model:* work happens in physical environment, and this model focuses on places where work is done, physical structures that limit or define how spaces are used, usage and movement of people within spaces, the hardware, software, communication lines and tools made available in each space and its layout. This model also includes artifacts created and used in any work.

While the approaches of viewpoints and CD provide high-level abstractions for analysing work done in a given context, there are several reasons why it could not be used in our approach. The main reasons are: (i) they were designed for static environments (e.g. situated in a room/building) and do not factor in mobility of users - freely moving from one place to another, (ii) they do not support the privacy context - our main aim was to understand users' context with regards to protecting their privacy, (iii) they make assumptions that software engineers are able to apprentice or shadow users in their work whereas for mobile applications, it is impractical to shadow or apprentice users in their natural environment. Moreover, for mobile privacy requirements, any shadowing of users will invalidate the gathered data because users are more likely to change their privacy behaviour in the presence of an ethnographer, and finally (iv) they do not specifically deal with privacy threats. Privacy is a complex notion, which requires expertise in being able to spot potential privacy violations from user behaviour (reported in qualitative data). Thus, these approaches, in their current form, are not suitable for deriving mobile privacy requirements.

We propose the use of *facets* - a notion very similar to that of viewpoints, but in our approach each facet holds partial domain knowledge relating to the privacy norms of a context and are therefore called *Privacy Facets*. The PriF framework proposes four privacy facets, where each has unique properties and functions which must be analysed both separately and as a whole to ensure completeness and consistency. The four privacy facets are: *information*, *information-flow*, *actor* and *place*.

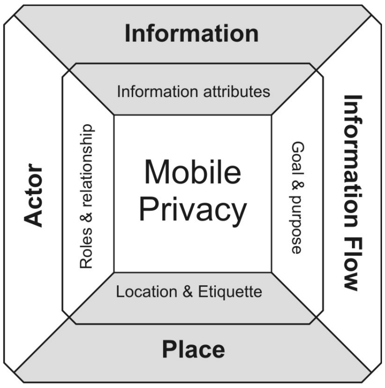


Figure 3.3: Privacy Facets for mobile privacy

As shown in Figure 3.3, each of these privacy facets have a privacy determinant - a property which influences privacy. For example, the *Information* facet has *information attributes* as a privacy determinant.

Each facet can be used to gather specific domain knowledge that affects the privacy of mobile users. The information facet elicits knowledge regarding *'what'* information is being created by a software system, while the actor facet focuses on *'who'* of the information (i.e. who is the sender, receiver and information subject), the information-flow facet identifies *'why'* the information was created about the subject and transmitted by the sender to the receiver, and finally the place facet captures *'where'* the information was created or transmitted.

3.5.1 Information

Information is the first of the four facets in a privacy norm. Software systems produce data either by themselves (e.g. log transactions) or when users interact with its functionality (e.g. take a digital photo, write an email message etc). Data produced by systems can be of several types and used for various purposes. Although, there is no clear consensus for definitions of *data* and *information*, we opt for Tenopir's definition in Zins (2007) which states that: *data* are facts that are the result of observation or measurement and *information* is meaningful data or data arranged or interpreted in a way to provide meaning.

Information has several aspects which can influence privacy, namely: information *type*, *mode* of collection, *purpose* of use and *attributes* or properties (e.g. timeliness, accuracy etc.). In this section we explore these aspects.

As such, data and information, do not cause privacy issues except when they relate to and contain data about an individual's identifiable attributes (e.g. name, date of birth, etc.). In this regard, the Information Commissioner's Office (ICO) highlights two *types* of information that we consider as being critical to protecting privacy: *personal information* and *sensitive personal information* (or *sensitive information*). Personal information relates to a living individual who can be identified from those information. Sensitive information refers to personal information pertaining to: racial or ethnic origin, political opinions, religious beliefs, memberships of organisations (e.g. trade union), physical or mental health or condition, sexual life, convictions etc. Adams (1999b) show that information sensitivity influences how information is shared, therefore privacy is primarily about controlling and moderating the use of these two types of information. In the scenario, when the MAS displays an empty weekly calendar of GetWell Surgery i.e. without any reference to either GPs or patients, it would contain information (e.g. day of the week) but this information by itself will not cause any privacy issues. In this case, the name of the GP and the patient are considered to be personal information because it refers to identifiable individuals. When the patient's medical condition is input, this information then becomes personal and sensitive, which needs to be protected.

Software systems can create information in two modes. In *automatic* mode, information is created without users' input/intervention whereas in *manual* mode, software systems create information in response to specific commands issued by users or when users interact with certain functionality of the system. Depending on the mode of information creation, different privacy threats can be uncovered. For instance, if users' information was sampled at a high frequency it can cause a surveillance effect. In the scenario, the MAS was expected to sample location of users when they arrived at the Surgery, however, if the mobile app sampled location of users too frequently, the collected data can be exploited to predict the whereabouts of patients - in effect similar to surveillance.

Another aspect of information is *purpose*. Knowing for what purpose information is being collected and used is important as it will help in checking if the information was misused in a way that is detrimental to users' privacy. Here the concern is collecting one set of data and using it for other purposes not expected or anticipated by users. Using the scenario again, say if GetWell Surgery exploited patients' details in the MAS to sell medical products to its users, this would be contrary to what had been initially agreed i.e. to provide easy and efficient means to book an appointment with a GP and nothing more.

Lastly, certain properties or quality *attributes* of information play an important role is affecting privacy of users. Attributes such as accuracy, timeliness etc. can cause users (information subject) to be misunderstood. Going back to the scenario, assume GetWell Surgery penalised all patients who came late for their appointments and one day a technical issue with the MAS causes a delay in automatic check-ins by 1 hour. When users arrive at the Surgery on-time, the MAS shows the incorrect check-in time, prompting GetWell Surgery to send a warning notice to all 'late comers'. In this scenario, users' reputation will be damaged by misinformation (privacy threat/violation).

As described above, each aspect of information contributes to privacy issues. One way of uncovering these aspects from qualitative data is by asking relevant questions. Answers to questions reveal specific knowledge which can then be tagged for identification and extraction for further analysis. Table 3.4 describes several facet questions which will assist in eliciting details on these information aspects which determine privacy and also provide coding tags for data extraction later.

Table 3.4: Questions in Information facet

<table>
<tr><td>FQ-1</td><td><i>INFORMATION TYPE</i></td></tr>
<tr><td colspan="2"><b><i>Is the information personal or sensitive?</i></b> Personal information relates to a living individual who can be identified from that information. Sensitive information refers to information pertaining to an individual that can be used to characterise them in some way (e.g., religion, ethnicity, sexual orientation, etc.) Code:[I-TYPE(PERSL | SENSE)]</td></tr>
<tr><td>FQ-2</td><td><i>INFORMATION COLLECTION MODE</i></td></tr>
<tr><td colspan="2"><b><i>Is the information collected automatically or manually?</i></b> These two modes of information creation impact the types of privacy threats that can be discovered in the software system. Code: [I-MODE(AUTO | MANUAL)]</td></tr>
<tr><td>FQ-3</td><td>INFORMATION PURPOSE</td></tr>
<tr><td colspan="2"><b><i>What is the purpose of the information or its context of use?</i></b> Knowing for what purpose the information is being collected is important, as it will help in later checking if the purpose was fulfilled or if it was used in a way detrimental to the users privacy. Code: [I-PURP]</td></tr>
<tr><td>FQ-4</td><td>INFORMATION ATTRIBUTE</td></tr>
<tr><td colspan="2"><b><i>What are the information attributes?</i></b> Quality attributes (e.g., timeliness, accuracy, completeness, etc) can influence how the information is used within the system and perceived by its users. Code: [I-ATTR(ACCU | COMP | FRESH | ONTIME)]</td></tr>
</table>

After establishing ‘what’ information is being created, we move on to the next privacy facet which deals with ‘who’ of information (i.e. actors involved in sending and receiving information and the information subject).

3.5.2 Actors

In a privacy norm involving information-flows, there could be several actors in the form of sender, receiver and information subject about whom the information is exchanged. Thus, the actors facet focuses on roles users play in a given context and their relationships with other users. A *role* is a relation between an agent and an activity; and corresponds to the behavioural aspect of the role. A *relationship* refers to relations between agents and corresponds to the social aspect of role. *Responsibility* is when one agent is responsible to another agent for something, which can be described as a possible mismatch or non-conformance relation between an actual state of affairs and a desired, expected or feasible state of affairs [adapted from Jirotka & Goguen (1994, p.87)].

Roles of actors have a significant impact on information-flows, therefore the first step is to understand the roles of each actor (i.e. role of information sender, role of information receiver, and role of information subject). Next, the relationship between the actors is established (i.e relationship between sender and receiver and information subject). Together these roles and relationships determine the level of trust, which in turn influences the sharing of sensitive information in a software system. Lastly, the actor facet must make the responsibilities associated with each role explicit, because responsibilities can affect relationships between actors, which then affects the information-flows in a given context.

For the example scenario, Alice and Bob are actors who in their specific roles are both patients visiting the GetWell Surgery. However, actors could have more than one role; at home Alice could be a partner or mother and at the same time at her office she would be an employee of a company, thus actors take on different roles depending on the context. Apart from having roles, actors can have relationships with others. In the example scenario, Alice might be Bob's partner or her colleague although both of them would have registered with the GetWell Surgery. Roles entail responsibilities which in turn impact privacy norms. For instance, if Alice and Bob were partners, when Alice is unwell, it is likely that she would expect Bob to book an appointment using the MAS on her behalf and Bob would possibly expect to see Alice's medical condition in the system. However, this would be completely different if Bob happened to be a colleague of Alice.

As explained here, the roles, relationships and responsibilities of actors play an important role in enabling privacy norms. If the knowledge pertaining to these are ambiguous, then systems implementing them will be prone to privacy violations (Adams, 1999b). Reflecting these in the actors facet, the roles, relationships and responsibilities are considered as privacy determinants as shown in Table 3.5.

Table 3.5: Questions in Actors facet

FQ05	*ROLE RELATIONSHIP*
What are the role relationships between the information sender, receiver and subject? A role is a relation between an agent and an activity and corresponds to the behavioural aspect of the role. A relationship refers the relations between agents and corresponds to the social aspect of role. Together they determine the level of trust, which influences the sharing of sensitive information. Code: [ROLE(RELTN)]	
FQ06	*ROLE RESPONSIBILITIES*
What are the responsibilities associated with each role? Responsibility is when one agent is responsible to another agent for something, and that this something can be described as a possible mismatch or nonconformance relation between an actual state of affairs and a desired, expected or feasible state of affairs [adapted from (Jirotka & Goguen, 1994, p.87-106)]. Responsibilities can affect the power relationships between actors, which in turn influence information flows in a given context. Code: [ROLE(RESPB)]	

While this facet elicits domain knowledge on the user's roles and relationships and their responsibilities, the next facet deals with 'why' information is being transmitted from senders to receivers.

3.5.3 Information-flow

The third facet of a privacy norm is information-flow. In the information-flow facet, all possible flows of information between the interacting users are examined. Each of these information flows are governed by what Nissenbaum (2010) calls *transmission principles* - informally established rules which guide the flow of information between different actors. Rather than using this ambiguous term, we break-down transmission principles as *goals* and *purposes* which constrain the flow of information between the sender and receiver and non-compliance of these constraints results in privacy violations.

As stated earlier in Section 3.3, privacy norms are appropriate information-flows in a given context and when these information-flows are not supported, it leads to privacy violations. Therefore, the aim of the information-flow facet is to elicit knowledge on the goals and purposes that determine an information-flow as being *'appropriate'* or not.

In the scenario, when Alice makes an appointment by providing her personal data and her medical condition using the MAS, the information sent to GetWell Surgery (or the MAS server) is necessary for (i) checking if Alice is registered with the Surgery, and (ii) booking of Alice's appointment with the GP. Thus, the information-flow is constrained by necessity on the part of Alice and the Surgery. Using the symbolic notion 3.2, this can be described as:

$$+\langle\ Alice \rightarrow GetWell\_Surgery, medical\_condition(Alice), *\|G_t1\ \rangle$$
$$where\ G_t1 = \{checking\_registration, booking\_appointment\}$$

However, in the scenario, Bob was able to view Alice's sensitive personal information such as her medical condition, which may cause embarrassment for Alice. Bob who is not related to Alice and had no responsibility for her, was just being curious. Thus, Bob's goal was to satisfy his curiosity, which is a negative privacy norm as formalised below:

$$-\langle\ Alice \rightarrow Bob, medical\_condition(Alice), *\|G_t2\ \rangle$$
$$where\ G_t2 = \{satisfying\_curiosity\}$$

In order to elicit knowledge for such goals and purposes, the information-flow facet provides suitable questions as shown in Table 3.6.

Table 3.6: Questions in Information-flow facet

FQ07	*INFORMATION-FLOW GOALS (SENDER-SUBJECT)*
What are goals and purposes of sender and subject in the information-flow? The goals and purposes determine the privacy expectations of the subject, for example, if the subject needs to consent before the information is sent. Coding Tags: [I-FLOW(SNDSUB)]	
FQ08	*INFORMATION-FLOW GOALS (SENDER-RECEIVER)*
What goals and purposes of sender and receiver in the information flow? These goals and purposes determine the privacy expectations of the sender and receiver. For example, the sender would expect only certain receivers and not others. Coding Tags: [I-FLOW(SNDRCV)]	
FQ09	*INFORMATION-FLOW GOALS (3RD PARTY)*
If there are 3rd-party recipients of the information, what are their goals and purposes for information use? This determines the flow of information to 3rd-parties who can misuse the information. Coding Tags: [I-FLOW(3PARTY)]	

When answering these questions, the data can be marked or tagged with the codes provided against each question.

The place facet explores the influences of location ('where') on users' privacy.

3.5.4 Place

Place refers to a unique geographic location with a material form, meaning and value (Gieryn, 2000). Geographic locations are not only identified by their names but they also have meaning and value in varying degrees. Location coordinates on a map, in addition to having a postcode and address (or latitude/longitude), may also refer to an individual's home or office. At a much more fine-grained level, location could also refer to an individual's bedroom or a chair. Thus, an individual's privacy depends on the meaning attached at every granularity of location. Mobility afforded by the mobile device adds yet another dimension to an individual's privacy because it allows users to interact with different objects that are present in the physical environment - both human agents and technologies in a place (Lessig, 1999). Therefore, the place facet contains questions that specifically elicit these details from qualitative data and are described in Table 3.7.

Apart from an individual's privacy, Westin (1967, p.42) mentions a notion of *organisation or group privacy.* When an organisation or a group of individuals wish to collectively achieve privacy they may specify and enforce rules (or policies) on all of its members and those members who do not comply with such rules may suffer exclusion. Since organsiations meet and work in specific physical environments, place is vital for achieving group privacy. Some of these rules may not be encoded in a written form but rather exist as generally accepted etiquettes. The place facet helps to uncover these norms (not to be confused with privacy norms that apply to individuals), rules and etiquettes at a given place. For example, in the scenario, GetWell Surgery had a policy where all mobile devices had to be turned off when users arrived at the Surgery, this could be because they interfere with the medical equipment or it could be a disturbance for other patients who are feeling poorly. Thus, factoring in these etiquettes at a place is important while designing software systems and Table 3.7 provides questions that help to gather knowledge regarding these.

Table 3.7: Questions in Place facet

<table>
<tr><td>FQ10</td><td><i>LOCATION</i></td></tr>
<tr><td colspan="2"><b><i>Which places are associated with the subject, sender and receiver, having an impact on users?</i></b> Used to identify the places that can be associated with a privacy-related context for the different actors. Code: [PLACE(LOCATN)]</td></tr>
<tr><td>FQ11</td><td>ETIQUETTE</td></tr>
<tr><td colspan="2"><b><i>What etiquettes or social rules apply to a place?</i></b> Used to identify the expected behaviours associated with a given place. Deviations from the expected behaviour can result in privacy threats being realised. Code: [PLACE(ETIQTE)]</td></tr>
</table>

3.6 Heuristics for privacy-sensitive contexts

Qualitative data captures users' experience in different contexts. However, there may be some specific episodes where users took action to avoid a privacy violation or had experienced one. Such episodes are referred to as *privacy-sensitive contexts* or *PS-contexts* for short. One of the main objectives in analysing qualitative data is to identify such PS-contexts and extract them for the deconstruction and analysis of privacy threats so that suitable counter measures can be developed to mitigate privacy threats.

Identifying PS-contexts in the qualitative data can be quite daunting, especially without having any guidance on what to look for. While the privacy facets helped in initially structuring data, the PriF framework also proposes the use of two categories of behaviour patterns, which can indicate the presence of a privacy threat/concern in the data. The categories are:

(i) *Negative behaviour patterns (NBPs)*

(ii) *Negative emotional indicators (NEIs)*

There are two types of negative behaviour patterns, which can indicate the privacy concern of users: *disuse pattern* and *work-around pattern.* In the *disuse pattern*, users may not want to engage with a software system because of the underlying privacy concern. For instance, users may 'switch-off' the device or choose not to use it. It may also be the case where users may ignore or choose not to respond to any prompts from the application. Similarly, in the *workaround pattern*, users may use alternative means in order to accomplish a specific task due to the perceived privacy concern. For instance, users may personally hand over a printed document instead of sending documents digitally via email due to their concern that the document may be intercepted and viewed by 3rd parties.

Users' negative behaviour patterns alone may not be sufficient to identify PS-contexts and there are other subtle cues which are equally important. Sometimes users produce negative emotional signals that can be picked-up in the qualitative data which can indicate the presence of a privacy threat/concern. *Negative emotional indicators* are a set of key words that indicate the negative emotional state of users in response to an event or action in the environment. For example, some of the key words are: *concerned,* *unhappy, worried, scared, dislike,* etc. Note there is no guarantee that these keywords will appear as they are, instead they may appear in phrases, for instance, the word 'dislike' can appear indirectly in a sentence such as 'I don't like' or 'I am not happy' for unhappy. If keywords found in the data don't match with any corresponding codes in the code book, the place holder 'other' can be used. Thus, analysts are expected to use these keywords as high level pointers and assign sections of qualitative data to the nearest category possible.

3.7 Extraction rules for privacy threats

Privacy facets described earlier are meant to unpack the context, particularly in relation to privacy norms. If privacy threats are to be uncovered, then the PS-contexts associated with them must be extracted from qualitative data. While behaviour patterns and emotional indicators help in isolating the PS-contexts, their extraction by themselves are useless because they don't map to any specific privacy threats and without this knowledge further analysis is difficult.

However, mapping user experiences to specific privacy violations and threats is probably the most difficult task of privacy analysis, for two reasons: (a) there are several types of privacy threats, and (b) more than one privacy threat can contribute to the same privacy harm (see Table 3.2 for the 10 privacy threats and their associated harms). To this end, we propose the use of data extraction rules as a way to not only extract PS-contexts but also to simultaneously associate them with specific privacy threats. These extraction rules are designed to first associate unique facet properties such as privacy determinants to negative behaviour patterns and emotional indicators, and then link these to specific privacy threats, as shown in Figure 3.4.

For example, an information attribute such as timeliness can cause a misinformation threat. When users experience this threat or violation, they would produce negative behaviour or emotional indicators which would have been tagged as either NEI or NBP in qualitative data. Further, when analysing for privacy facets, the information attribute (i.e. timeliness) would have been visible in the qualitative data which would have been

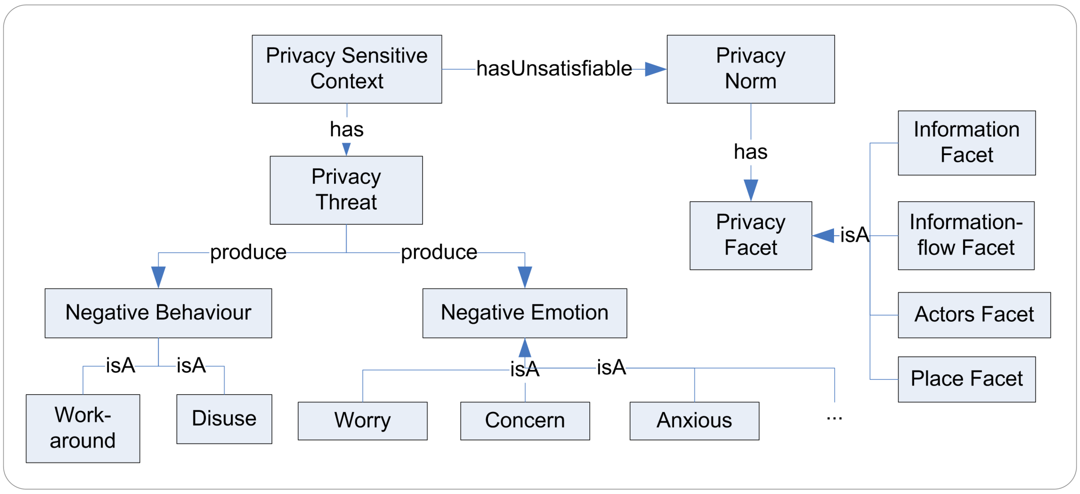


Figure 3.4: Relationship between PS-context, privacy facets and privacy threats

tagged with the code I-ATTR(ONTIME). To extract the PS-contexts satisfying the misinformation privacy threat, a rule such as the following can be used (see Table 3.8):

Misinformation rule = I-ATTR(*) AND [NBP OR NEI]

Here the rule states that all data that has been coded for information attribute (I-ATTR) and also having codes for either negative behavaiour (NBP) or negative emotions (NEI) should be extracted as a PS-context contributing to the misinformation privacy threat. The symbol asterisk (*) is used to denote all sub-attributes under I-ATTR. Table 3.8 contains data extraction rules for all the privacy threats identified earlier in Table 3.2.

So far, we have described the tools necessary to extract PS-contexts from qualitative data. The following sections describe additional analytical tools that are needed in order to analyse the PS-contexts and derive privacy requirements from them.

3.8 Problem patterns for information-flows

There are different types of problems (e.g. mathematical, mechanical, civil etc.). Software problems refer to purposes or goals that must be accomplished by a computer

Table 3.8: Data extraction rules for privacy threats

ID	Privacy Threat	Inappropriate Information-flows (Privacy norms)	Data Extraction Rule
T1	Identification	Subject's personal information is revealed	I-TYPE(PERSL) AND [NBP OR NEI]
T2	Exposure	Personal/sensitive information received by unintended recipients	I-TYPE(SENSE) AND [NBP OR NEI]
T3	Surveillance	Receiver makes frequent requests for information about the subject	I-MODE(AUTO) AND [NBP OR NEI]
T4	Aggregation	Receiver combines datasets produce a new type of information without the subject's approval	I-PURP AND [NBP OR NEI]
T5	Misinformation	Inaccurate or insufficient level of information about the subject is transmitted	I-ATTR(*) AND [NBP OR NEI]
T6	Breach of trust	Receiver forwards the information to others contravening the subject's terms and conditions	ROLE(RELTN) AND [NBP OR NEI]
T7	Power imbalance	Receiver uses information to control the subject	ROLE(RESPB) AND [NBP OR NEI]
T8	Cross-contextual information flow	Information belonging to a particular context may be used in another context	I-FLOW(*) AND [NBP OR NEI]
T9	Proxemic access	Unintended receivers can access information because they are in close physical proximity to the sender.	PLACE(LOCATN) AND [NBP OR NEI]
T10	Intrusion	Information flow disturb's receiver tranquility	PLACE(ETIQTE) AND [NBP OR NEI]

program. Since computers serve many purposes, there is a need to describe and structure software problems systematically. One well known approach from software engineering is the Problem Frames (PF) (Jackson, 2001) approach which is used to describe and model software problems. An attractive feature of PF is that it explicitly models context as real world domains (i.e. physical domains). Thus, PF was considered to be more suitable for the analysis of mobile computing applications whose physical context has to be modelled in order to address privacy. PF represents an important advantage when compared to other methods, such as KAOS (Van Lamsweerde, 2010), GBRAMS (Anton, 1996), NFR (Chung & do Prado Leite, 2009) or i* (Yu & Cysneiros, 2002), which do not naturally model physical domains.

In PF, a software problem is a requirement, R, in a real world context, W. The problem context is a collections of domains *(W = D1;...;Dn)* described in terms of their known, or indicative properties which interact through their sharing of phenomena (i.e, events, commands, states, etc.). A domain is a set of related phenomena that are usefully treated as a behavioural unit for some purpose. S is the solution specification needed to achieve requirement R. This can be described using the notation (Jackson, 2001):

$$W, S \vdash R \tag{3.5}$$

where the entailment symbol '$\vdash$' denotes that the satisfaction of W and S entails that of R. From here on, the term 'problem' and 'software problem' are used interchangeably as they have the same meaning.

In the PF approach, a *problem frame* is a type of pattern that captures and defines a commonly found class of simple subproblems. As shown in Figure 3.5, a problem frame includes a machine representing a computer that will run software to carry out necessary transformations to solve a problem.

The machine is represented as a rectangle with two vertical stripes. The rectangles with a single vertical stripe signifies a model domain, which means it did not exist prior to the problem but was created specifically to hold values of variables. The rectangles with no stripes signify physical domains in the real world. The domains are connected to the machine via solid lines representing the interfaces between the domains and the machine. In Figure 3.5, `a` and `b` describe the shared phenomena between the domains

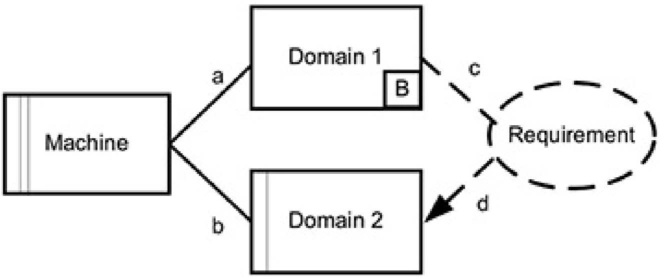


Figure 3.5: Problem frame diagram

and the machine. Dotted ovals signify requirements. The dotted lines that connect the domains to the dotted oval represent the interfaces `c` and `d` through which domain properties are visible to the world. The dotted line with the arrow-head signifies the domain we want to control and manipulate using the machine. The domains marked as 'B' signify people and are defined as biddable domains to indicate that people cannot be predicted or forced to cause (or not to cause) an event.

The PF approach catalogues five basic problem frames (and their variants) which refer to common software problems found in many software applications. However, privacy is dependent on information flows within a software system. Information flow in its simplest form consists of a *sender*, *receiver* and the *information* which is transmitted between them. As privacy relates to flow of personal/sensitive information, the *subject* whose properties the information describes should also to be taken into account. In addition, information flows have a *purpose or means/ends* to achieve, also called as *transmission principles* which play an important role to the flow of personal information from the sender to the receiver (Nissenbaum, 2010, p.127).

Putting all these together we define privacy problems as follows:

Privacy problems are concerned with building machines which will allow the appropriate flow of personal information and alternatively avoid inappropriate flow of personal information (avoid privacy violations).

Here the *appropriate* or *inappropriate* flow of personal information is a function over the four domain properties: (i) *information type* describes the nature of the information (ii) *roles of actors* such as information sender, receiver and subject; (iii) *goals and purposes* for the information-flow, and (iv) place or location of information-flow.

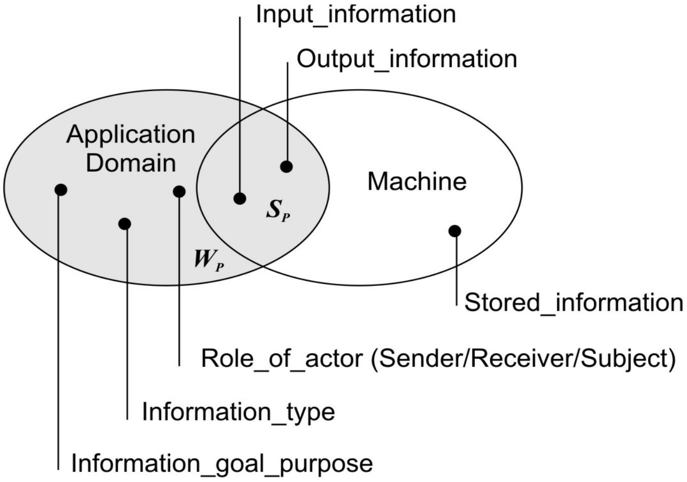


Figure 3.6: Venn diagram shown the parts of the privacy problem

At a lower level, the information-flow is concerned with data input streams and output streams. In a program, the transmission of data from the sender to the receiver is made possible through the use of commands and operations provided for the user. These do not necessarily have to be explicit commands to 'send' data, they could also include operations where the receiver has been given access to view information on the sender's system. The information flow applies to any form of information access given by the sender to the receiver. But these operations are dependent on the availability/input of data and thus, in a broad sense, information flows are primarily concerned with two phenomena - *input* and *output* of information.

A Venn diagram in Figure 3.6 shows where the privacy problem is located in the problem space and how it is bounded. It is expressed in terms of interface phenomena S_p, application domains W_p and privacy requirements R_p using the entailment relation as:

$$W_p, S_p \vdash R_p \tag{3.6}$$

where R_p *is requirements satisfying privacy norms,*
S_p *is specification to control input and output of information,*
W_p *is the observed and controlled phenomena relating to privacy determinants such as information attributes, roles of actors, goals & purposes and place*

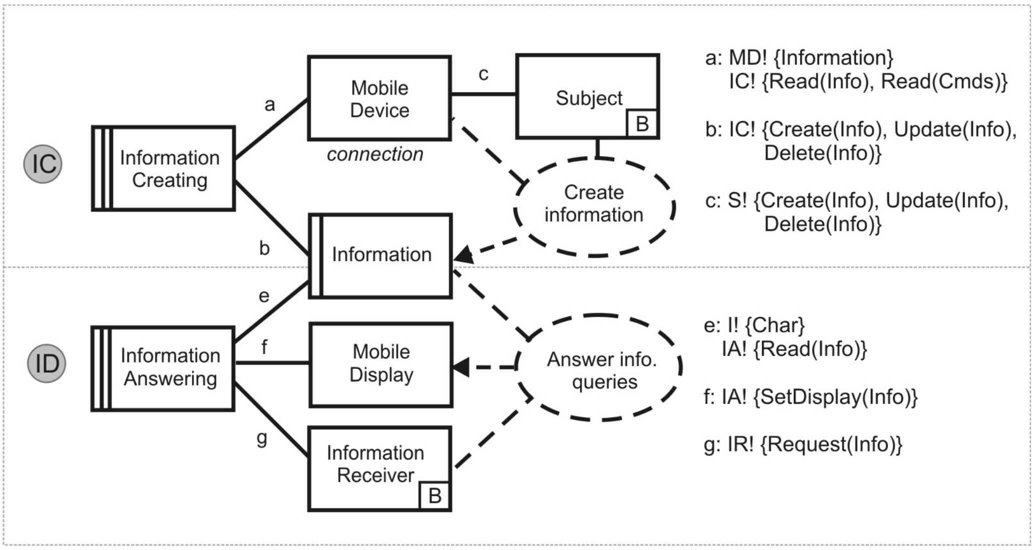


Figure 3.7: Information-flow problem frame diagram

A composite model of the information-flow problem frame is as shown in Figure 3.7. It is composed of two smaller sub-problems: *information creation (IC) problem* and *information dissemination (ID) problem.*

The IC problem deals with the building of the machine `Information Creating` whose goal is to enable its users to produce different types of `Information`. A mobile device (e.g. Smartphone) typically has several types of sensors (e.g. GPS - location sensor) and other devices (e.g. Camera) whose facilities are made available for mobile applications to use thus influencing the type of information which can be created within the software system. This in turn has an impact on the privacy of the user which will have to be addressed. Thus, it is critical to model what is being created and how it impacts the privacy of the user.

In the IC problem frame, the machine is connected to a `Mobile Device`. The `Mobile Device` is attached to the `Subject` about whom the information is being created, it could be the user herself who is able to issue commands to create, modify and delete information at interface `c` as shown in the Figure 3.7 or it could be automated in such a way that the information about the subject is sampled at regularly frequency. The machine is able to read the subject's information commands `Read(Cmds)`, via interface `a`. Depending on the type of command issued by the user (e.g. create, modify or delete), the machine issues corresponding commands to the model domain `Information` change the state as required.

While the first part information-flow problem frame relates to information creation, the second part is concerned with information dissemination - the building of the software which will enable the information receivers to query for information. In the composite problem frame shown in Figure 3.7, the problem frame labeled ID refers to information dissemination. The requirement in this problem is to answer the information requests for the information that was created in the previous problem. The `InformationReceiver` is a biddable domain (marked as 'B' in Figure 3.7) requesting information from the machine using enquiry commands. The `Information Answering` machine reads the commands received through its interface `g` and based on the subject's answering rules (or privacy policy) associated with the information, the machine is able to process the requests accordingly and issue commands to set the screen display.

While the above problem frames modelled the information-flow within a software system, its *control variants* help in constraining how the information is created and disseminated, thus addressing privacy. Although both problem frames described here are generic, they are applicable to mobile applications. Out of the five basic problem frames (Jackson, 2001), the information creation problem is a modification of the basic *workpiece problem frame* while the information dissemination extends an operator variant of an *information problem frame.*

In the PF approach, each problem has a *frame concern* which highlights a certain aspect of the problem demanding the attention of the analyst/developer. Similarly, in the PriF framework the *privacy concern* is nothing but a special type of frame concern relating to privacy which the analyst/developer must take notice of and address in the software system to make it privacy-aware. Therefore, the privacy concerns extracted from the qualitative data through the use of facets should be addressed in order to support the privacy of the user when using the software system.

3.9 Privacy arguments

One advantage of using formal or semi-formal notations in representing requirements is that they are amenable to (semi) automated analysis and reasoning. While the level of precision allowed by the language will dictate to what extent the requirements can be automatically checked for consistency and correctness, it should also be in a format suitable for the stakeholders such as business analysts, systems designers and implementers.

The PF approach adopted by the PriF framework uses a semi-formal notation to describe the domains and a combination of pseudo-code, state-charts and natural language to describe the observed and required phenomena. While this elegant and simple style in representing requirements is one of the most appealing aspects of PF, in our case we find this to be a limiting factor because it lacks the expressive power to describe privacy requirements. Further, our approach draws on evidence from qualitative data to instantiate problem contexts and these need to be referenced while representing the requirements to provide traceability. Lastly, representing the privacy requirements in a more structured way can make it more amenable for automated reasoning and model checking as well as helping in developing tool support.

Extending their earlier work on argumentation frameworks for software engineering (Yu *et al.*, 2011), Tun *et al.* (2012) have produced a language to express privacy norms as *privacy arguments* which are structured in the style proposed by Toulmin (2003). This is similar to the argument structure adopted in engineering security requirements (Haley *et al.*, 2004) where rebuttals and mitigations are recursively represented to reason about the satisfaction of security requirements. In Haley et. al's approach, security requirements are expressed as claims and are supported by grounds and warrants. While rebuttals show evidence that contradicts other arguments, mitigations describe how rebuttals can be avoided or tolerated.

Tun et al. construct a privacy argument as a *claim* that needs to be justified. The *ground* is the collection of facts that can be observed from the world domains, which supports the claim. The *warrant* is the collection of domain-specific rules that links the ground to the claim of the argument. A *rebuttal* uses one argument to rebut another, falsifying the claim of the rebutted argument. A *mitigation* uses one argument to mitigate a rebutted argument, restoring the claim of the rebutted argument. A *preference* decides on the precedence of two arguments over each other under certain optional conditions and exceptional conditions when a particular argument should not be applied. These argumentation concepts are syntactically described using the TXL grammar (Cordy, 2006) as shown below:

```
include 'OpenArgue.grm'
keys
...  mitigates rebuts preferred precedes except when depends
end keys
redefine argument
      argument [claim] {
           supported by
                [ground*]
                [evidence*]
           warranted by
                [warrant*]
           preferred by
                [preference*]
      }
end define
define evidence
      [id]
      | [stringlit]
end define
```

In order to show a link between an argument and the evidence found in the qualitative data and thereby providing traceability between the two, this grammar was extended to include a new construct called *'evidence'* which links a claim to a reference in the qualitative data. The reference is usually an identifier (id) pointing to specific quotes in the qualitative data. Here, evidence is treated as a specialisation of a ground. In addition, the grammar was extended to include constructs to support dependencies between two arguments and both these extensions are shown above in blue. The 'include' statement states that the new syntax reuses the existing syntax of arguments defined earlier (Yu *et al.*, 2011).

In the PriF framework, the above privacy arguments language is used to construct privacy norms as claims while the privacy threats and concerns rebut or falsify those claims. In order to protect privacy, additional arguments will have to be constructed to weaken or mitigate the effects of the privacy threats and concerns. A general structure for such a description of privacy requirements is shown below incorporating the symbolic notations introduced earlier, where PN_i is a privacy norm (from notations 3.1 and 3.2), PC_i is the privacy threat/concern (from notation 3.4) and PR_i is the privacy protecting requirement that mitigates the effects of PC_i to ensure the privacy norm PN_i continues to be satisfied. Privacy concerns can rebut one or more privacy norms; similarly, privacy requirements may mitigate one or more privacy concerns.

$$\begin{array}{l}
PN_i - \ ALLOW|DENY[\pm\langle\, A_s \rightarrow A_r, I_a(A_u), L_p \| G_t \,\rangle] \\
PC_i - \langle\, conditions\_for \otimes [PN_i],\ H \,\rangle\ \texttt{rebuts}\ PN_i \\
PR_i - \langle Privacy\,protecting\,requirement\rangle\ \texttt{mitigates}\ PC_i
\end{array}$$

Privacy arguments not only provide a link between problem diagrams and privacy requirements, they also provide traceability between high-level privacy requirements and low-level access control policies that should implement those requirements. Apart from providing an elegant solution to expressing privacy requirements, the other benefits of using the grammar is that it supports automated reasoning and tool support, which will be discussed in Chapter 6.

3.10 Summary

In this chapter, the PriF framework was introduced and explained. By specifying a list of privacy harms and threats, the PriF framework provided a basis for understanding the different types of privacy threats that can be found in any mobile software system. The four privacy facets were designed to structure and organise qualitative data such that privacy-sensitive contexts can be extracted and their privacy concerns identified. Using the well known approach of Problem Frames, problem patterns in the form of information-flow problem frames were introduced to explicitly model privacy problems. Further, the PriF framework adapted and extended the Privacy Arguments grammar to express complex privacy requirements to make it amenable to automated reasoning and tool support. The novelty of the PriF framework is that it was specifically designed to assist software engineers analyse end-user reports in the form of qualitative data for the purpose of engineering privacy requirements. The next chapter will demonstrate how the PriF framework will be put to use to extract and refine privacy requirements from qualitative data.

4

The Distillation approach

In this chapter, we demonstrate the application of the PriF framework to produce privacy requirements from qualitative data. In the first section, we explain the method and its evaluation criteria. The second section provides a brief overview of a novel approach we are proposing to tackle our research problem, which is followed by a detailed description of the process and the steps involved. As a conclusion, the last two sections briefly discuss and summarise our approach.

4.1 Method

Generally, in software engineering, empirical methods are applied to understand how software engineers develop and maintain complex, evolving software systems (Easterbrook *et al.*, 2008). In this respect, empirical studies have served to investigate not only the tools and processes software engineers use but also the social and cognitive processes affected. The motivation of our research is quite different as it focuses on enabling software engineers to understand end-users' privacy issues in a mobile context, where end-user privacy experiences are captured and reported as interview transcripts. Thus, the design of our approach is based on existing qualitative data analysis methods.

From the four dominant philosophical stances (Easterbrook *et al.*, 2008), we choose *'pragmatism'* because it is more aligned to our objective of wanting to engineer *'practical solutions'* to real-world problems. This of course has an impact on method selection and collection of evidence to answer our research question. Thus, our method is a mix of both *qualitative research methods* (Braun & Clarke, 2006; Corbin & Strauss, 2008; Gibbs, 2007; Thomas, 2006) and *case-studies* (Yin, 2009). Moreover case-study research is generally considered to be well suited for answering 'how' questions in social sciences. Our research is similar to Lee & Rine (2004) who have shown that case studies can be effective in designing novel requirements engineering approaches. We make use of case study principles to provide structure to our research while also focusing on data analysis techniques drawn from qualitative research methods.

First, we report on data collection, followed by a description of our method used for data analysis and finally state our evaluation criteria.

4.1.1 Data from deferred contextual interviews

Our research question (as stated in Section 1.1) is concerned with extracting privacy requirements from qualitative data, thus, the *unit of analysis* is qualitative data. Here qualitative data refers to *interview data* we collected from two empirical studies which focused on mobile users. The terms 'qualitative data' and 'interview data' are used interchangeably but they refer to one and the same.

The Experience Sampling Method (ESM) has been used in ubiquitous computing user studies to capture data about participants' feelings and behaviors in daily life situations in a non-intrusive manner and over an extended period of time (Consolvo & Walker, 2003; Hormuth, 1986). Usually, this is done by giving or delivering a set of questions to participants in a study, either on paper or electronically, automatically or manually, at regular intervals, or when specific episodes/events occur. The ESM is used when it is impractical to use direct observation methods such as 'shadowing' where an observer spends time with participants in their own environment. This is precisely the case with studying mobile privacy - any direct observation of participants results in a modification of what would otherwise be spontaneous behavior, potentially invalidating the observations.

However, gathering meaningful data through the ESM requires considerable commitment from participants, since they may have to spend a significant amount of time answering experience sampling questions, possibly in circumstances that may or may not be convenient (Miner *et al.*, 2005). This becomes even more pronounced when gathering privacy-related data about the use of software applications on mobile devices. Devices such as mobile phones are carried by people more or less everywhere and, when relevant episodes occur, people may be in transit or engaged in activities that make them unable to spend time answering questions. Even if they were able and willing to do it, their activities would be disrupted, which is likely to affect their state of mind and behavior.

Therefore, in mobile privacy studies an experience sampling questionnaire can only ask for minimal feedback that can be provided in the shortest possible time. On the other hand, in order to be useful, feedback needs to provide detailed information about specific situations and the contexts in which they occur. This detailed information can be better communicated by participants during one-to-one interviews, but these usually take place with some delay with respect to participants' experiences. Therefore, it is likely that some important contextual information will end up getting lost as participants usually remember partial details relating to their experiences. In order to gather rich and meaningful data from real-life experiences, we devised a method which combined experience sampling and semi-structured interviewing techniques specifically adapted for our studies.

Our experience sampling questions were delivered via mobile phone, which made it easy for participants to contribute their feedback. The participant could quickly answer the questions by choosing from a set of predefined multiple-choice answers. Their answers were then discussed in one or more follow-up interviews. Since participants provided feedback on multiple events in a day and debriefing interviews were conducted several days later, participants were advised to use a *memory phrase*. Memory phrase is a short text users associate with a particular event they provide feedback on. Since participants themselves choose these memory phrases, the phrase constituted a powerful trigger, which was capable of bringing participants back to the event's context (Mancini *et al.*, 2009).

Once participants reconnected to those events, they were able to provide detailed information about their experiences in debriefing interviews. The interviewer reminded participants of the memory phrase they associated with a particular event and, as participants went back to it in their mind, the interviewer could use the experience sampling questions and answers as pointers to different aspects of the participant's experience. Given the effectiveness of memory phrases in bringing participants back to a particular experience and the context in which the event took place, the interviewer could carry out what effectively constituted *deferred contextual interview* to elicit users' context information on specific events (Mancini *et al.*, 2009).

Employing these techniques, we collected interview data from several participants who took part in two user studies - the first one focused on mobile users of a social networking application (Facebook) and the second study focused on location-tracking within families. Both these studies where conducted in the context of a larger EPSRC funded project called 'Privacy Rights Management in Mobile Applications' (PRiMMA) [1]. Further details on the user studies along with the interview data are available as Appendix A and B respectively.

4.1.2 Data analysis using Distillation and Privacy Facets

Answering our main research question - *'how can we derive mobile privacy requirements from qualitative data?'* (see Section 1.1) actually entails finding answers to several sub questions, which are:

(i) *How can interview data be structured so that privacy threatening/violating contexts can be identified?*

(ii) *How to isolate the privacy threats associated with such privacy threatening/violating contexts?*

(iii) *How to discover why the current mobile system failed to protect the user?* and

(iv) *How to discover what is required in a new mobile system to protect the user's privacy?*

A novel approach we are proposing is called *requirements distillation*, a systematic process which will employ analysis models and patterns to extract and refine software requirements from qualitative data. We refer to *privacy requirements distillation* when deriving privacy requirements from qualitative data and *mobile privacy requirements*

[1] http://primma.open.ac.uk

distillation when it entails extraction and refinement of privacy requirements specifically for mobile applications. (For the sake of brevity, we will use the term *distillation* to synonymously refer to privacy requirements distillation). Distillation consists of three steps - structuring of qualitative data, information-flow modelling and privacy problem analysis.

Structuring of qualitative data
For quite some time, researchers from social sciences have been employing thematic analysis techniques to structure raw interview data. As Braun & Clarke (2006) describe it, *thematic analysis* is a method for identifying, analysing and reporting patterns or themes within qualitative data. In thematic analysis, a *theme* is said to capture something important about the data and it represents some level of *patterned* response or meaning within the dataset. Themes visible within the data are encoded using appropriate *codes* (or labels) in a process called *coding.* Coding can be thought of as a tagging process, where specific labels or codes are attached to the data according to the theme it represents.

Thematic analysis provides two approaches in which themes can be identified *(a) inductive or bottom-up* approach (Boyatzis, 1998) where codes emerge from raw qualitative data, it is also known as 'data-driven' for the same reason, and *(b) deductive or top-down* approach (Crabtree *et al.*, 1992) which makes use of a 'template of codes' derived from a theoretical framework. However, the way thematic codes are applied in distillation is different from how it is used in qualitative data analysis. Here themes and codes are pre-defined for analysts to readily apply on raw data. The analysis in distillation is limited to identifying these pre-defined patterns and applying the corresponding codes for them.

The pre-defined thematic codes used in distillation were initially developed through inductive process. Since the first challenge was to identify *privacy-sensitive contexts* (or situations where the user had experienced a privacy violation or threat), we used an inductive process to analyse interview data from our first study - 'Mobile Facebook'. We found that users displayed two types of behaviours when they perceived their privacy was being violated - they either stopped using the mobile application causing it (disuse) or used alternative means (workaround) to accomplish their information sharing tasks. We captured these two *negative behaviour patterns* as *disuse* and *workaround* patterns.

When these behaviour patterns couldn't be seen, there were other distress signals which the users gave out. They would sometimes express negative emotions (e.g. 'I don't like', 'I am not happy' etc.) when using the mobile application. We captured these *negative emotional indicators* in the PriF framework presented earlier (see Section 3.6 and Table 3.3). (We use *'PS-context'* to synonymously refer to 'privacy-sensitive context').

Once the PS-contexts are isolated, the next challenge is to pinpoint the type of privacy threat or violation it is associated with. We found this was difficult, if not impossible, to determine just by looking at the data, because it required theoretical models of privacy and privacy threats. Thus, we developed a model of privacy based on the *contextual integrity of information flows* (Nissenbaum, 2010) and linked this model to a set of privacy threats proposed by Solove (2008). This privacy model was unpacked and described as privacy facets in the PriF framework (see Section 3.5). The codes for each of the four facets i.e. information, actor, information-flow and place were derived from the underlying privacy theory. In other words, we followed a deductive process in formulating the codes. The *extraction rules* described in our framework (under Section 3.7) were based on the privacy threat descriptions provided by Solove.

Thus, the code book in the PriF framework (see Table 3.3) made use of both bottom-up and top-down techniques to develop codes. However, when applying the framework, i.e. in the distillation approach, we only make use of these pre-defined codes and there is no requirement to develop additional ones.

Information-flow modelling

Once the PS-contexts and their associated privacy threats have been categorised through coding of data, the next logical step was to analyse why the current software system failed to support users. In other words, we turn to the next step of isolating *privacy concerns* of the underlying software system. At this point the thematic analysis ends and software engineering takes over. Thus, distillation significantly differs in relation to qualitative research methods where theory building normally follows thematic analysis.

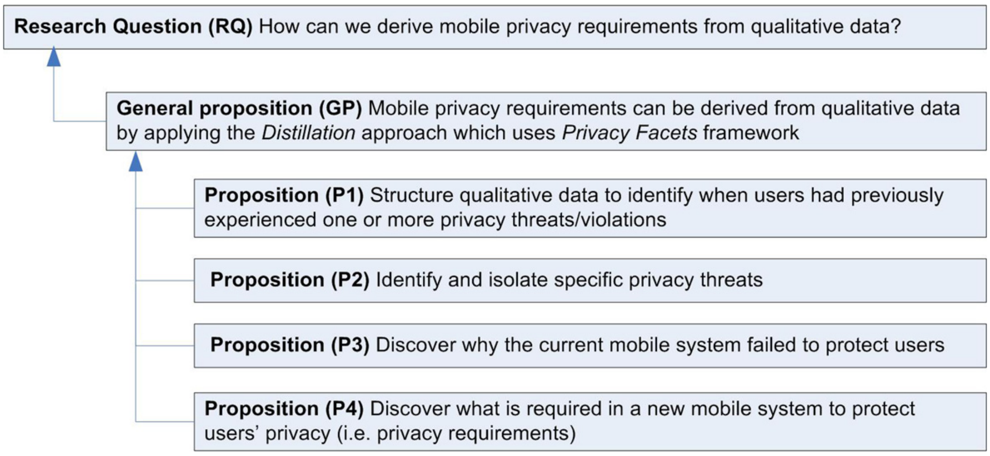
In order to analyse the privacy concerns (i.e. to analyse how and why the current system is not able to the protect privacy of users), we develop models of the software system's information-flow in a physical context. For this modelling task, we chose the Problems Frames (Jackson, 2001) approach for two reasons: firstly, it can model physical domains and secondly, it can effectively abstract recurring software engineering problems as 'problem patterns' which could be reused (similar to 'design patterns' in the solution domain). From Nissenbaum (2010), we discovered information-flow as being the single most important abstraction which underpins privacy in any software system. Information-flows have two sub components - one for *information creation* and another for *information dissemination.* The PriF framework captures both these information-flow abstractions as two problem patterns (Section 3.8).

Privacy problem analysis
The control variant of information-flow problem patterns reveal how the current system is able/unable to support the privacy requirements of end-users, leading to the identification of privacy concerns. Analysing information-flows these privacy concerns results in privacy requirements being discovered. In order to describe these privacy requirements, we extend and use an existing privacy requirements language (see Section 3.9). The distillation approach and its use of the PriF framework is explained in greater detail in the later sections of this chapter.

4.1.3 Evaluation

Here we describe the evaluation approach used for the three distillation steps described earlier. Since distillation is a mix of both qualitative analysis and software engineering approaches, for the purpose of evaluation we borrow principles from both case-study and qualitative research methods. Similar to the structure followed by Lee & Rine (2004), here our *research question* (from Section 1.1) is restated in relation to general and specific *study propositions* which this research addresses. The general proposition is stated in relation to the distillation approach and the PriF framework it uses. We derive four propositions from the general proposition as shown Figure 4.1.

As shown in Figure 4.4, each step contributes to one or more specific study propositions mentioned earlier. We build our case for validity by collecting evidence for each of these analysis steps as shown in Table 4.1. 'SE' refers to a Software Engineer performing the analysis steps in distillation.



Research Question (RQ) How can we derive mobile privacy requirements from qualitative data?

General proposition (GP) Mobile privacy requirements can be derived from qualitative data by applying the *Distillation* approach which uses *Privacy Facets* framework

Proposition (P1) Structure qualitative data to identify when users had previously experienced one or more privacy threats/violations

Proposition (P2) Identify and isolate specific privacy threats

Proposition (P3) Discover why the current mobile system failed to protect users

Proposition (P4) Discover what is required in a new mobile system to protect users' privacy (i.e. privacy requirements)

Figure 4.1: Research question and study propositions

Table 4.1: Evidence collected for distillation

Step Id.	**Distillation approach steps (evidence is underlined)**	**Supports study propositions**
Structuring of qualitative data		
1.1	SE identifies and applies codes for **privacy-sensitive context**	GP, P1
1.2	SE identifies and applies codes for privacy-determinants	GP, P1, P2
1.3	SE derives **privacy threats and concerns**	GP, P1, P2
Information-flow modelling		
2.1	SE models **information-creation** in system	GP, P3
2.2	SE models **information-dissemination** in system	GP, P3
Privacy problem analysis		
3.1	SE analyses if model addresses privacy concerns	GP, P3, P4
3.2	SE defines **privacy control and feedback requirements**	GP, P4

The distillation approach will be deemed successful when there is evidence for each of its steps (except intermediate steps 1.2 & 3.2 which have no outputs). In addition to collecting evidence, we also look other properties which might add rigour to distillation. For thematic analysis and qualitative research in general, deciding on such properties is difficult because there is no clear consensus between the positivist or constructivist groups (Yardley, 2008). Therefore, we stick to our approach of pragmatism i.e. not leaning towards either of these groups but instead concentrating on properties which can demonstrate the distillation approach as a practically viable and beneficial process.

While several evaluation strategies can be found in qualitative research (Baxter & Eyles, 1997), some properties specific to thematic analysis (Gibbs, 2007, p.38-56) overlap with those in case-study design (Yin, 2009). For instance, the meaning of 'reliability' is not exactly the same in both of these methods. For this reason and also because the distillation approach itself is unique, we formulate four criteria drawn from both qualitative research and case-study design with an engineering (pragmatic) perspective to demonstrate rigour. The criteria are:

(1) *Employs a systematic process:* The first evaluation criterion for the distillation approach is that its procedures and steps can be followed and reapplied to produce similar results (i.e. privacy requirements). Here, the emphasis is on *'doing'* the same procedure over and again; in principle it is similar to the *'reliability'* property found in both case-study design (Yin, 2009) and qualitative analysis (Golafshani, 2003). By following a clear and unambiguous process, distillation should demonstrate its *'logical flow'* and also its *'clarity of purpose'* or rationale for the steps in its process (Corbin & Strauss, 2008, p.303-306). We aim to achieve this objective by providing a clear description of distillation, the steps involved and its logical flow, and how the final product (i.e. privacy requirements) is derived using two distinctly different datasets (one dataset from mobile Facebook and another from location-tracking of families). If successful we should be able to produce privacy requirements from both of these datasets.

(2) *Applicable to multiple datasets:* Unlike the first, the second criterion focuses on the *'results'* obtained through the application of the distillation approach. Based on the principle of multiple case-study design (Yin, 2009, p.46-56), here the objective is to achieve a *replication* of results by applying distillation on datasets taken from

two studies (i.e. mobile Facebook and location-tracking in families). While it is expected that distillation will produce privacy requirements using the data from both of these studies, there is a subtle and contrasting variation between them which the approach must successfully handle. The first study on mobile Facebook concentrates mainly on the information subject/sender (victim of a privacy violation) whereas the second study focuses on multiple viewpoints i.e. information subject, sender and receiver (it includes both victims and offenders in a privacy violation). This is similar in principle to *theoretical replication* where contrasting results are obtained but for predictable reasons (Yin, 2009, p.47). In our case, the distillation approach will produce contrasting privacy requirements from different users' viewpoint.

(3) *Links to data:* The third criterion relates to linking of outputs to qualitative data (as evidence). In the distillation approach, it means clearly demonstrating how the qualitative data links to each output to finally derive privacy requirements. In other words how distillation is *'grounded in the data collected and interpreted'* (Gibbs, 2007) and provides a *'chain of evidence'* (Yin, 2009, p.105). Such data linkage, in qualitative data analysis not only provides transparency but improves the overall credibility of the process (Yardley, 2008). Since qualitative data provides the *context* for all its derived outputs (Corbin & Strauss, 2008, p.306), establishing this link is important in the distillation approach.

(4) *Informs systems design:* The last criterion refers to *applicability* or usefulness of results (Corbin & Strauss, 2008). Here the objective is to demonstrate that the privacy requirements obtained from the qualitative data are useful and they can potentially inform the design of new systems.

In subsequent sections, we demonstrate the working of the distillation approach using data from our first study i.e. mobile Facebook. In the next chapter, we use data from our second study to demonstrate how privacy requirements are derived from multiple viewpoints and discuss how these satisfy the above evaluation criteria.

4.2 An overview of 'Distillation'

To enable software engineers analyse qualitative data, the distillation approach not only provides *analytical tools* but also guidance on the *process* employed in extracting privacy requirements. First, we briefly describe the input and the output of the distillation approach, followed by an overview of its analytical tools and the process involved.

4.2.1 Input: Qualitative data and initial requirements

The distillation approach takes two inputs - *qualitative data* and *initial requirements.* The qualitative data are from empirical studies where the users' interactions with the environment are observed and qualitative interviews capture the users' reaction in different contexts. Since the focus of distillation is on privacy, it is expected that in qualitative studies, the questions will indirectly probe the users' response to privacy related incidents. In order to capture the contextual details of mobile computing users, the study may employ electronic means to sample the environment. If such data is made available, it can provide additional insight during problem modelling and analysis. However, the primary input for the distillation approach is the *interview transcript* in its raw form.

The second input for the distillation approach is a set of initial requirements of a software system. Here the initial requirements refer to the functions or features of the software system users would normally interact with to accomplish their tasks while participating in an empirical study. The initial requirements form the foundation from which the privacy requirements can be derived (as noted previously, privacy requirements cannot exist on their own). Unlike traditional software development, mobile applications often follow agile methods which mean documents on requirements may be difficult to find. In such cases, we simply observe the functionality of the software system (being investigated) and document the requirements it supports.

4.2.2 Output: Privacy requirements for mobile applications

As described earlier in Section 3.8, our privacy problem is to build a machine which will allow appropriate flow of personal information and avoid any inappropriate flow of

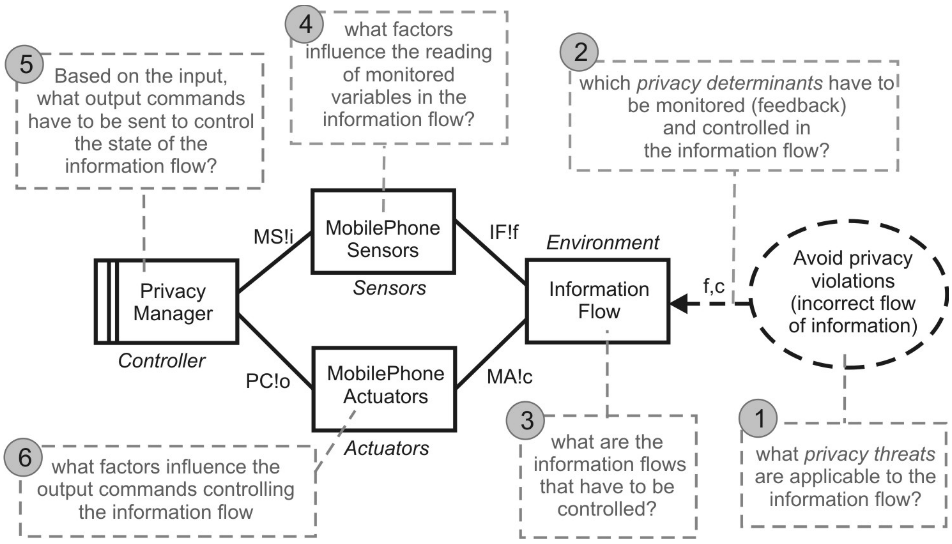

5
Based on the input, what output commands have to be sent to control the state of the information flow?
4
what factors influence the reading of monitored variables in the information flow?
2
which *privacy determinants* have to be monitored (feedback) and controlled in the information flow?
MS!i
MobilePhone Sensors
IF!f
Environment
Avoid privacy violations (incorrect flow of information)
Privacy Manager
Sensors
Information Flow
f,c
Controller
PC!o
MobilePhone Actuators
MA!c
6
what factors influence the output commands controlling the information flow
Actuators
3
what are the information flows that have to be controlled?
1
what *privacy threats* are applicable to the information flow?

Figure 4.2: Problem diagram for privacy manager (controller) of a mobile application (i=input, o=output, f=feedback, c=control)

personal information to prevent privacy violations from taking place. The behaviour of a *privacy manager* (a piece of software which controls information-flow) is dependent on a number of variables in the environment relating to privacy determinants such as *information type, roles of actors, goals & purposes* and *location & etiquette at a place.* As shown in Figure 4.2, the privacy manager is an instance of 'required behaviour' problem frame with connection domains (Jackson, 2001, p.221). In this problem frame, the privacy manager is able to receive input phenomena `i` from mobile phone sensors and translate into output phenomena `o` for actuators to control in the information-flow domain. The feedback phenomena `f` and the control phenomena `c` are concerned with monitoring and controlling variables whose changes in value affects the privacy state of the information-flow domain. To avoid privacy violations (i.e. inappropriate flow of personal information), the privacy manager must send control commands to the information-flow so that it does not enter into a state where privacy threats are realised. The privacy manager exerts continuous control over the information-flow domain by monitoring its variables (i.e. privacy determinants). For mobile applications, all sensors and actuators are assumed to be deployed on the same mobile device.

For mobile applications, modelling the sensors and actuators as connection domains works well because it takes into account the unreliability of mobile devices. Mobile phones can be switched-off, exhaust their battery power, lose their network connectivity, have no GPS signal etc., making it difficult to reliably read the values of any monitored variables. Since the actuators also face similar unreliability concerns, they too affect how the information-flow is controlled. Thus, the privacy manager should not only take into account the unreliability of both sensors and actuators, but also implement necessary measures in case one or more variable cannot be monitored and controlled to satisfy the requirements.

Consequently, the main objective of the distillation approach is to capture the privacy manager's problem context, its requirements and the real world phenomena which influence information-flows in a software system. The following are the type of questions, the distillation approach is expected to answer:

1. *What privacy threats are applicable to the information flow?*
2. *Which privacy determinants have to be monitored (to provide feedback) and controlled in the information flow?*
3. *What are the information flows that have to be controlled?*
4. *What factors influence the reading of monitored variables in the information flow?*
5. *Based on the input, what output commands have to be sent to control the state of the information flow?*
6. *What factors influence the output commands controlling the information flow?*

The distillation approach can be deemed successful only if it is able to help software engineers realise some or all of the above objectives by analysing qualitative data taken from user studies. To achieve these stated objectives, the distillation approach makes use of some tools tailored specifically for its use. The following section describes these tools and the context of their use within the distillation approach.

4.2.3 Analytical tools: Privacy Facets framework

One of the main steps in the distillation approach is the structuring and processing of qualitative data. Just as in qualitative data analysis (QDA) where several analytical tools are available to systematically organise qualitative data, the Privacy Facets (PriF) framework provides a set of analytical tools for the distillation approach. The PriF framework has a set of *facet questions* to stimulate the classification of data, this is very similar to the *'guiding questions'* in QDA (Corbin & Strauss, 2008, p.72) which are used to reveal categories, dimensions and properties. However, in the PriF framework, the guiding questions are pre-defined and the analyst only has to go through each of them while analysing the data. The PriF framework also provides guidance by describing relevant privacy threats and privacy determinants under each facet to help analysts pinpoint the type of concerns a software system needs to address.

In addition to the facet questions, the PriF framework provides two types of *negative behaviour patterns (NBPs)* that reveal the negative actions and interactions of users in a study. This is an extension of two tools in QDA namely - *'waving the red flag'* and *'looking for the negative case'* (Corbin & Strauss, 2008, p.80,84). Users' negative behaviour can indicate a potential privacy issue experienced in a specific context and these negative behaviour patterns help identifying such contexts for further analysis. Sometimes the users may not explicitly state their negative experiences and in such cases the negative behaviour patterns are inadequate to capture the relevant contexts in the data. To overcome this disadvantage, the PriF framework provides additional tools such as the use of *negative emotional indicators (NEIs)*, which are keywords that point to the negative emotional state of users in a given context. Again, the use of keywords indicating negative emotions is based on a strategy used in QDA namely - *'looking at emotions that are expressed and the situations that aroused them'* (Corbin & Strauss, 2008, p.82). For both these analytical tools, the PriF framework provides codes which can be assigned to the data and thus automate the extraction of PS-contexts for further analysis (see Section 3.6).

The next part is to model privacy norms from PS-contexts. For this purpose the PriF framework provides two problem patterns based on the Problem Frames approach (Jackson, 2001). Privacy norms contain information-flows. Since, information flows are made up of two parts - information creation and information dissemination, the PriF framework provides the *information creation problem frame* and *information dissemination problem frame* respectively. Thus, the PriF framework provides the necessary models and patterns to equip analysts to not only structure the qualitative data, but also extract and model problems from it which can be analysed further.

4.2.4 The process: Distillation approach

As a starting point, the distillation approach relies on the software system which implemented the initial requirements. This is the same software system the user would have used during the empirical study and whose experiences are captured in the interview transcripts. The interview transcripts (qualitative data) are one of the two inputs to the distillation process as shown in Figure 4.3. The initial (functional) requirements supported by the software system forms the second input to distillation.

The distillation process consists of three main phases, namely:

(1) *Structuring of qualitative data*

(2) *Information-flow modeling*

(3) *Privacy problem analysis*

In the first phase, the qualitative data from an empirical study is structured using coding techniques. Coding is similar to what QDA analysts use but here the difference is that the PriF framework provides pre-defined codes tailored for the identification of PS-contexts. Once the PS-contexts are isolated, additional codes from each privacy facet help in deriving the relevant privacy threats and concerns.

In the second phase, using the initial requirements of the software system (used in the study) as the basis, problem models of privacy norms with its information-flows are developed by instantiating the problem patterns in the PriF framework. These problem models not only capture how the information is created but also how it is disseminated to other users.

During the third phase which relates to privacy problem analysis, privacy threats and concerns are analysed in conjunction with the information-flow models to identify any gaps, leading to the discovery of *privacy requirements.*

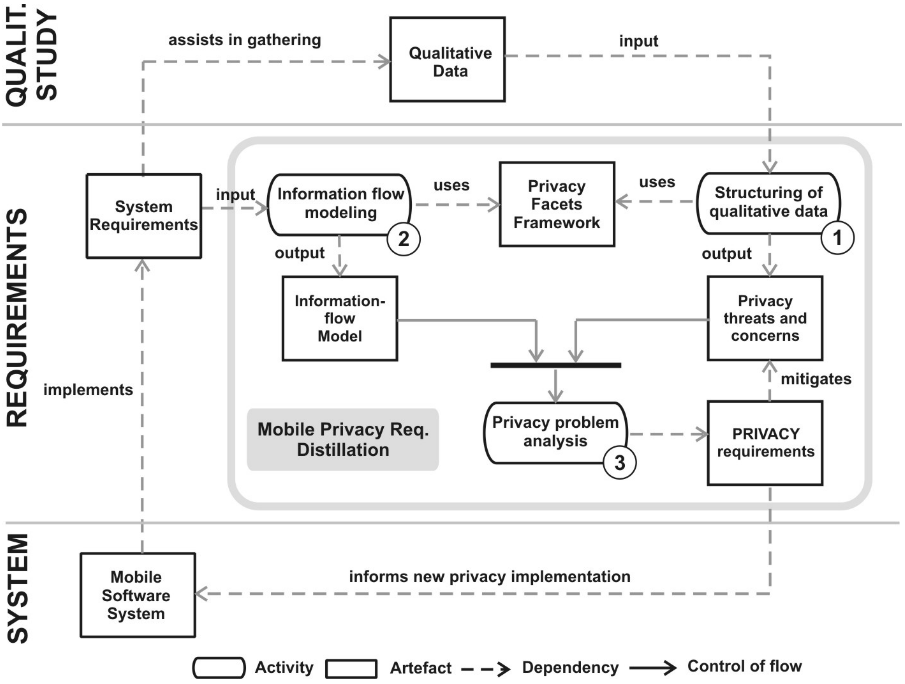


Figure 4.3: An overview of Distillation process

The distillation approach is designed to be a sequential process where one phase follows the next. Therefore, the structuring of qualitative data precedes information-flow modelling and privacy problem analysis.

If the distillation approach was successful, the resulting set of privacy requirements will inform the design and development of new software systems that are privacy-aware. The next sections provide detailed descriptions on the internal workings of the distillation approach.

4.3 Phase 1: Structuring of qualitative data

Qualitative data contains *noise* such as inputs from users which may not be relevant or useful for analysis, therefore the first task is to isolate the important sections from the rest. For this type of structuring, QDA make use of coding techniques where analysts tag data with special labels called *codes* to organise the data into discrete categories. Initially, the coding process starts with no codes (i.e. labels for the concepts) but as and when analysts discover them in the data, they accumulate codes that semantically align with the emerging concepts in the data. These codes will be later reused to tag other sections of the data when it semantically matches to the code, and in cases where the data does not match any of the existing codes, the analysts will have to assign new ones. Thus, the data is categorised and structured according to its properties and dimensions. Although the distillation approach makes extensive use of this technique, there is a fundamental difference in the way it is executed. The distillation approach does not expect analysts to find new codes, instead it uses the list of pre-defined codes from the PriF framework (under privacy facets, NEIs and NBPs) which collectively represent a coding model.

The end result of data structuring using codes of NEIs and NBPs is that PS-contexts are identified, this is shown as the first step 1.1 in Figure 4.4. PS-contexts are then subjected to facet questions provided in the four privacy facets as shown in step 1.2. of Figure 4.4. This also entails tagging of PS-contexts with appropriate codes under each facet question when answers are found for privacy determinants.

In step 1.3 of Figure 4.4, privacy threats are identified by applying a set of extraction rules. Once the privacy threats have been successfully identified, they are then analysed for privacy concerns and requirements.

The following subsections explain in detail the coding techniques applied under each step and also describe how the facilities provided within the PriF framework assist in coding. In order to exemplify and improve confidence in the distillation approach, sample qualitative data taken from a previously concluded study on Mobile Facebook (MFb) [1] is used.

4.3.1 Coding for PS-context

The MFb study observed how the participants used the Facebook (Fb) application on their mobile device as they carried out their daily routines and the privacy issues that

[1] In this study, mobile users of Facebook (http://www.facebook.com) application were electronically shadowed and their responses to privacy issues were captured through a detailed qualitative interview. The method used in the study is described elsewhere (Mancini *et al.*, 2009) and all transcripts from this study are available in Appendix A.

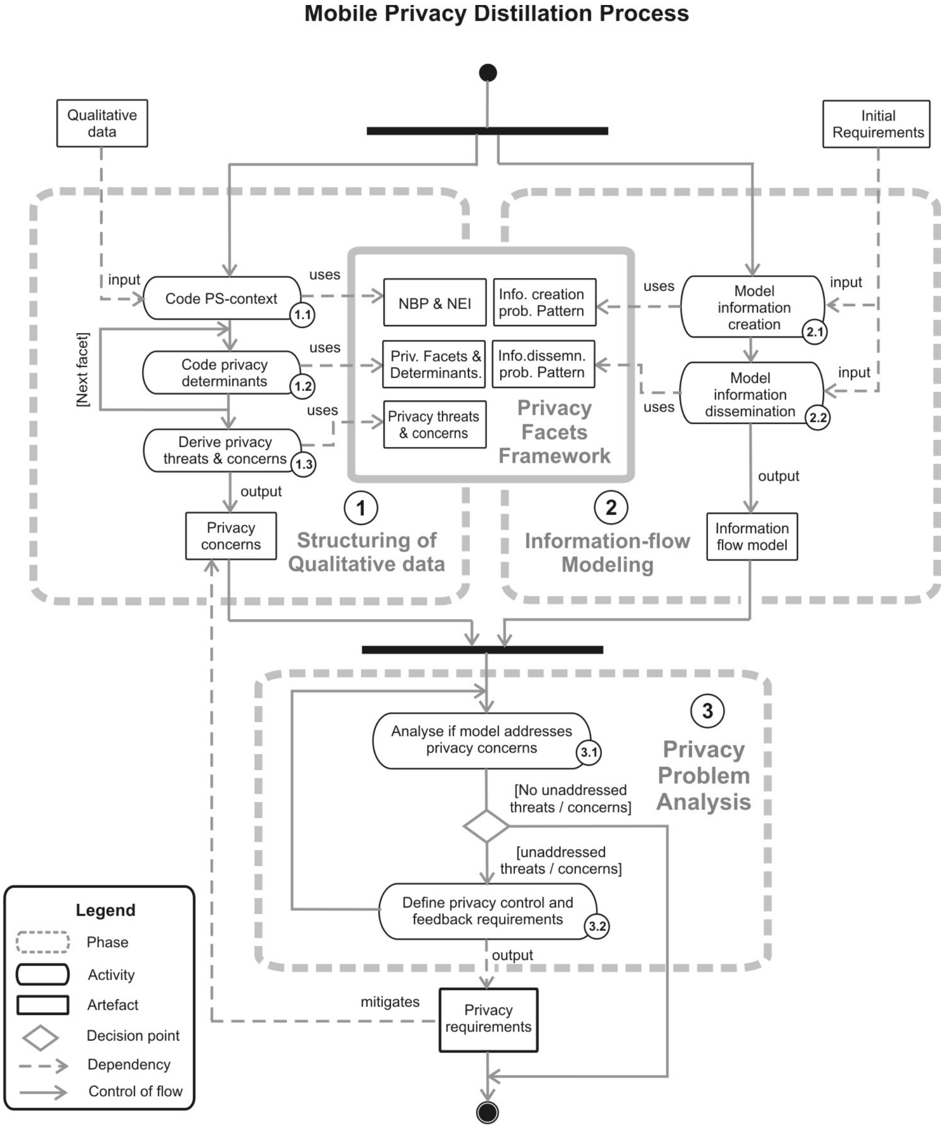

Mobile Privacy Distillation Process
Qualitative data
Initial Requirements
input
uses
Code PS-context
1.1
NBP & NEI
Info. creation prob. Pattern
uses
Model information creation
2.1
input
[Next facet]
uses
Code privacy determinants
1.2
Priv. Facets & Determinants.
Info.dissemn. prob. Pattern
Model information dissemination
2.2
input
uses
uses
Derive privacy threats & concerns
1.3
Privacy threats & concerns
Privacy Facets Framework
output
output
Privacy concerns
1
Structuring of Qualitative data
2
Information-flow Modeling
Information flow model
3
Privacy Problem Analysis
Analyse if model addresses privacy concerns
3.1
[No unaddressed threats / concerns]
[unaddressed threats / concerns]
Define privacy control and feedback requirements
3.2
output
mitigates
Privacy requirements
Legend
Phase
Activity
Artefact
Decision point
Dependency
Control of flow

Figure 4.4: Distillation process in detail

arose in different contexts. The Fb application supported several functions such as updating of status message, viewing of friends' status messages, uploading of photo, tagging of photo, writing on Fb wall etc. However, to limit the scope and better illustrate the distillation approach, the examples used here concentrate on the participant's use of a single feature - *updating of status message*.

During the first pass, when the users' negative behaviours and the negative emotions were spotted, they were tagged using the pre-defined codes described in Table 3.3. The entire paragraph or surrounding text containing the negative behaviour or emotion is tagged and this segment of data is then referred to as a privacy-sensitive or PS-context.

In the first example taken from the MFb study, a user states:

> **[P1-01(A3.6)]** *Usually I tend to be specific on a certain topic which I'm ok that people know...I'm not happy people knowing about my relationship...or my personal problems, my working problems...I usually don't put these* **[NEI(UNHPPY), NBP(DISUSE)]**

The above identified data uses our own labeling convention [1].

The user indicates his reluctance in using Fb on his mobile phone when travelling on a bus or train, possibly concerned that his input of status messages might be visible to other passengers seated next to him. Here the code **NEI-UNHPPY** corresponds to 'unhappy' emotional state under the category - negative emotional indicator **NEI** in Table 3.3. Similarly, in the same paragraph, the code **NBP-DISUSE** indicates the 'disuse' behaviour of the user where he describes *'I usually don't put these [sensitive information]'* to indicate he does not use the Fb status message facility to input sensitive personal information, this is a negative behaviour pattern **NBP**.

As seen in the above example, negative behaviours may arise due to some negative emotional state of the user. In such cases, it is inevitable that codes for negative

[1] The interview transcript is labeled using an identifier of the form
$[< LOCAL\_ID > -(< APPENDIX\_REF > . < PARAGRAPH\_NO >)]$
For example, in [P1-01(A3.6)], P1-01 is the local id which indicates it belongs to the participant named 'P1', A1 refers to Appendix A and section 1, and 6 refers to the paragraph under this section. The referencing system provides traceability to the original transcripts from the study.

behaviour and negative emotional states apply to the same context, both indicating privacy issues in the context.

> **[P1-02(A3.16)]** *If I am out with friends I don't take my phone out, I don't do facebook ...yes, ok, if I am with my sister I keep to read emails, but no I don't use facebook and I tend not to use the mobile...because I am busy with other stuff, talking with them, socialising...facebook tends to fill the gaps...if I am with a person I concentrate with that person* **[NBP(DISUSE)]**
>
> **[P2-03(A4.25)]** *...things like buses and trains I don't feel so comfortable..., because I don't know...lots of people I don't know...if they for example read some of the posts I have done...they don't know the people that they are aimed at or the back story...they'd probably come across quite differently and they would not understand them, it would look a little weird..[they would get] the wrong sort of almost the wrong first impression* **[NEI(UNCOMF)]**
>
> **[P3-04(A5.42)]** *anything I feel is private to myself I keep it to myself. I have a lot of good friends so if I want to share it I am happy to share it with all my friends. If there was something private, that is more close to me, like a girl that I liked and I wanted to share it with a friend I would do that in person rather than on Facebook* **[NBP(WORKA)]**

The PS-context P1-02 is coded for a disuse pattern because it shows the user is unwilling to use his mobile phone when he is in the company of his friends. Similarly, the PS-context P2-03 is coded for its negative emotional indicator because the user mentions being uncomfortable when posting a message in a bus or a train. Finally, P3-04 is coded for the workaround pattern because the user indicates he would meet in person and share information rather than on Fb.

As the previous examples show, the surrounding contextual data is vital as it contains additional details of a privacy episode. Therefore, when coding, this surrounding data should be included. It may be difficult to prescribe the exact boundary of a PS-context in advance because only the analyst is in a position to judge which of the objects in the environment are contributing to the users' privacy and which of them are not. In other words, analysts will have to make a subjective judgement on coding the surrounding contextual information and thus define the boundary for a PS-context.

4.3.2 Coding privacy determinants

In the previous step, when the qualitative data were coded for NBPs and NEIs, this actually raised a 'red flag' on those data segments which could potentially have privacy threats or concerns. However, it may not be very clear from the onset, which of the ten privacy threats described in the PriF framework apply to this data. The privacy determinants are designed just for this purpose, they assist in pin-pointing privacy threats associated with each PS-context.

The strategy is to identify the causal factors for users' negative behaviours and negative emotional states observed in the PS-contexts. To achieve this, the PS-context is iteratively probed using the facet questions in the PriF framework to discover privacy determinants. When answers are found, the data is tagged with appropriate codes (from Table 3.3) for privacy determinants.

The first question (FQ1 in Table 3.4) focuses on the *information types* - here the aim is to identify whether the information created by the user is personal or sensitive. In this case of PS-context P3-04, the 'information' which is being referred to is the 'status update' message input by the user in the MFb application. The status messages can be any text but here the user specifically points to the message being 'private'. Further, the user indicates that the status messages could sometimes be sensitive - relating to his personal life (*'girl that I liked'*). Thus, based on the term 'private' and also because the information is sensitive, the problem context is coded for **sensitive information [I-TYPE(SENSE)]** as shown below.

> **[P3-04(A5.42)]** *'anything I feel is private to myself I keep it to myself. I have a lot of good friends so if I want to share it I am happy to share it with all my friends. If there was something private, that is more close to me, like a girl that I liked and I wanted to share it with a friend I would do that in person rather than on Facebook...'* **[NBP(WORKA), I-TYPE(SENSE)]**

The next question in the Information facet (FQ2, FQ3 and FQ4 in Table 3.4), elicits details on 'purpose', 'mode of creation' and 'information attributes' respectively. In the above PS-context, the purpose (FQ2) of the status message is rather implicit and is not visible here. The next question assess the method of information creation (FQ3), again there is nothing in the PS-context to indicate this is a causal factor. Further, there are no specific information attributes (FQ4) visibly connected in the PS-context.

In the next iteration, the PS-context is analysed using the Actors facet where the first question (FQ5 in Table 3.4) elicits the roles and relationships of information subjects, senders and receivers. The PS-context P3-04 contains the phrase *'all my friends'* and *'a friend'* which indicates the roles of information receivers. It is implicit, the *user* was the information sender. Here, several receivers are related to the sender by a *'friend'* relationship. Since the roles and relationships are visible, the PS-context is coded for **Role and Relationship [ROLE-RELTN]** as shown below.

> [**P3-04(A5.42)**] *'anything I feel is private to myself I keep it to myself. I have a lot of good friends so if I want to share it I am happy to share it with all my friends. If there was something private, that is more close to me, like a girl that I liked and I wanted to share it with a friend I would do that in person rather than on Facebook...'* [**NBP(WORKA), I-TYPE(SENSE), ROLE(RELTN)**]

The next Facet question (FQ6) refers to the responsibilities of the roles. Since there are none visible in the PS-context P3-04, we move on to the next facet which is Information-flow.

The Information-flow facet helps with identifying the information-flow's goals and purposes, which may be influencing the user's interaction. Since the PS-context contains nothing to match the first question in the facet (FQ7) we move on to the next question (FQ8).

FQ8 is interesting because some goals of the user (sender) and his friends (receivers) can be inferred from the PS-context. The user wants to reveal sensitive information (*'the girl I like'*) only to certain friends but not all of them and this is the main reason why the user is reluctant to use the MFb application. He is unable to safely send sensitive status messages only to select friends[1]. The user had an expectation that other than a select few, no one else in his friends list should be able to intercept and read his sensitive messages. He indicated this by wanting to keep the messages 'private'. This expectation of the sender is coded as **I-FLOW(SNDRCV)**] and is shown below.

[1] At the time when the study was carried out, the Facebook application revealed all information, sensitive or not, to all the friends in the Facebook network.

[P3-04(A5.42)] *'anything I feel is private to myself I keep it to myself. I have a lot of good friends so if I want to share it I am happy to share it with all my friends. If there was something private, that is more close to me, like a girl that I liked and I wanted to share it with a friend I would do that in person rather than on Facebook...'* **[NBP(WORKA), I-TYPE(SENSE), ROLE(RELTN), I-FLOW(SNDRCV)]**

The third question (FQ9) refers to 3rd party recipients, which does not apply to this PS-context. Similarly, the questions in the Place facet which deal with the location and the etiquettes at a place (FQ10 and FQ11) are not relevant because there is no reference to any specific location and its etiquette influencing the users' behaviour.

This iterative analysis using facets can be applied to any PS-context to code their privacy determinants. As another example, consider the PS-context P2-03, where the user is feeling uncomfortable and is concerned about the sensitive information being visible to strangers co-located in a bus or train. In this case, the user stops using the mobile phone to ensure his status messages are not being read by those travelling with him on the same bus or train. Since in this PS-context a specific location (i.e. inside a bus/train) is impacting the user, it is coded with **LOCATN** as shown below.

[P2-03(A4.25)] *...things like buses and trains I don't feel so comfortable ..., because I don't know...lots of people I don't know...if they for example read some of the posts I have done...they don't know the people that they are aimed at or the back story...they'd probably come across quite differently and they would not understand them, it would look a little weird..[they would get] the wrong sort of almost the wrong first impression* **[NEI(UNCOMF), PLACE(LOCATN)]**

Once the privacy determinants are tagged, the data is ready to be processed for privacy threats and concerns.

4.3.3 Deriving privacy threats and concerns

The extraction rules for each privacy threat are provided as a combination of NEI/NBP plus a privacy determinant as shown in Table 3.8. These extraction rules when applied, produce a list of privacy threats linked to PS-contexts. For example, after running the extraction rules, P3-04 was shown to linked to three privacy threats - *(T2) Exposure*, *(T6) Breach of trust* and *(T8) Cross-contextual information-flow*. This is because the codes assigned to the PS-context satisfy the extraction rules of specific privacy threats (see Table 3.8). For the first privacy threat (T2), the extraction rule is:

(T2) Exposure: *I-TYPE(SENSE) AND [NBP OR NEI]*

Linking PS-context P3-04 to exposure threat implies that personal/sensitive information can be received by unintended recipients. However, this appears to be quite accurate given that Facebook did allow all members in the user's friend list to view his sensitive status message, resulting in the exposure threat being realised. Conversely, it also implies that the software system did not satisfy a privacy norm. This privacy norm relates to the user wanting to send sensitive messages only to a select group of 'close friends'. This privacy norm is represented using the privacy arguments language (see Section 3.9) as shown below:

```
argument:  Fb_Close_Friends_Norm
 PN1 $ALLOW[+\langle User \rightarrow close\_friend, sensitive\_StatusMessage(User*), *\|$
$providing\_update\rangle]\{$
 supported by
      E1 User has friends and close-friends
      (P3-04:  ‘...all my friends...  a friend’)
      E2 User creates sensitive status message
      E3 User wants sensitive status message to be seen only
      by close-friends
      (P3-04:  ‘...If there was something private, that
      is more close to me, like a girl that I liked
      and I wanted to share it with a friend...’)
      F4 Close-friends want to see the sensitive status message
 warranted by
      R5 User inputs sensitive status message on Fb field
      R6 When a close-friend taps the Fb icon on his mobile device,
      the application opens with sensitive status message displayed
 }
```

Here PN1 is a positive privacy norm denoted by the '+' sign, the '$User \rightarrow close\_friend$' indicates the direction of information-flow i.e. from the user to a close-friend. The sensitive status message pertain to '$User*$' - the status message is about the user and other unidentified persons (e.g. the user's girlfriend). Unidentified or unknown entities are denoted using a wild card character '*'. In the next part of the norm the wild card character '∗' is used because location can take on any value and this privacy norm is not constrained by any specific location. The last part indicates the goal or purpose for this privacy norm, which is '*providing_update*'. In PN1, 'E' is the evidence found in the qualitative data, 'F' is a known fact and 'R' represents an existing requirement in the software system. The supporting qualitative data for evidence E2 and E3 are combined and shown under E3 because it is sometimes difficult to split sentences without losing the context.

The next step is to derive privacy concerns (unsatisfiability conditions) for PN1 by asking questions such as *'Why is the software system not able to satisfy the privacy norm PN1?'* Such questions are aimed at helping analysts locate the gaps in the current software system. Thus for PN1, the unsatisfiability conditions could be that the software system treats (a) sensitive and non-sensitive status messages alike, and (b) friends and close-friends alike[1]. These two gaps can again be described as privacy concerns (PCs) using the notation in Section 3.9 as shown below:

```
argument:  Fb_Exposure_Concern
 PC2.1 ⟨Status messages are treated as not sensitive by the system,
 (H9 : exposure − causes embarrassment/humiliation to User)⟩ rebuts PN1 {
 supported by
      F7 User is unable to classify a status message
      as being sensitive or non-sensitive
      F8 The system is unable to differentiate between
      sensitive and non-sensitive status message
 }
```

In the above, the convention is to label each privacy concern according to the threat it is derived from, thus 'PC2' represents a privacy concern derived from 'T2' exposure threat. Similarly, PC6 would correspond to threat T6. The second part of the concern, starting with 'H' is the harm suffered by the user.

[1] This may not be applicable to the current version of Facebook which provides a facility for users to specify separate lists or groups for different types of friends e.g. close-friends, family members etc.

```
argument:  Fb_Exposure_Concern
 PC2.2 ⟨Close_friends and friends are treated alike by the system,
 (H9 : exposure − causes embarrassment/humiliation to User)⟩ rebuts PN1 {
 supported by
      F9 User is unable to classify a friend
      as being friend or close-friend
      F10 The system is unable to differentiate between
      friends and close-friends
 }
```

These two concerns of the existing software software will have to be addressed during the privacy problem analysis phase. We now look at the second privacy threat (T6) that was uncovered. The codes of P3-04 show that it satisfies the the extraction rule:

(T6) Breach of Trust: *Role-Relationship(ROLE-RELTN) AND Work-around pattern (NBP-WORKA)*

When receivers re-send some personal or sensitive information to others, without the sender's approval or consent, then the sender experiences a breach of trust. Although privacy threats are not explicit and cannot be substantiated with direct evidence in the qualitative data, it is helpful to check for potential realisations of such threats. Therefore, when the PS-context does not explicitly state that the user's friends are able to forward confidential status messages to others, we infer this by observing the current functionality of the MFb application. The MFb application allows the user's friend to forward or re-send sensitive status messages to other friends who may or may not be related to the user. This is made possible with the MFb application having a button labeled 'like' which when pressed indicates that a receiver 'liked' a message that was sent. However, this 'liked' status message will be visible to a wider group of friends ('friend-of-a-friend'). This has the same effect of the receiver re-sending the status message to their group friends who may or may not be friends of the user (who created the status message). This concern can be described as shown below:

```
argument:  Fb_Breach_Of_Trust_Concern
 PC6 ⟨User's close_friends can forward sensitive status messages to others,
 (H3 : friends lose User's trust)⟩ rebuts PN1 {
 supported by
      F11 User's close-friends can have other friends
      who are not related to the main user
 warranted by
      R12 If the close-friend tags the sensitive status message
      by pressing the 'like' button, then all his friends are
      able to see the main User's sensitive status message
 }
```

The last privacy threat (T8) relates to the flow of information from one context to another. The PS-context P3-04 contains the codes which satisfy the extraction rule:

(T8) Cross-contextual information-flow: *Expectation of Sender-Receiver [I-FLOW(SNDRCV)] AND Work-around pattern (NBP-WORKA)*

This threat is realised when information belonging to one context is transmitted to another without the user's knowledge and against their expectation. Here, the user is willing to share status messages but not when it is 'private' (sensitive). The privacy concern relates to the software system allowing sensitive status messages (about a person the user 'likes') to cross-over to non-sensitive domain where 'all friends' have access. This of course is against the user's expectation.

The third privacy concern can be described using the privacy argument construct as shown below.

```
argument:  Fb_Cross_Contextual_Info_flow_Concern
 PC8 ⟨Sensitive status message is visible
 to all friends, (H3 : User's loses reputation)⟩ rebuts PN1 {
 supported by
      F12 Friends should access only non-sensitive status messages
      F10 Close-friends should access both sensitive and
      non-sensitive status messages
 }
```

Earlier, the privacy norm PN1 was extracted from the PS-context P3-04 because only then can privacy concerns be represented. Since privacy concerns are always linked to privacy norms, privacy concerns PC2.1, PC2.2, PC6 and PC8 are linked to PN1. To derive privacy requirements, each of these privacy concerns will have to be mitigated to ensure that privacy norm PN1 relating to 'close-friend' is supported in the software system. These privacy concerns will form the basis for information-flow modelling and privacy problem analysis in the following sections.

4.4 Phase 2: Information-flow modelling

In this phase, we model information-flows that are relevant to the privacy concerns that were derived earlier. Information-flow is modelled as two parts: *information creation (IC)* and *information dissemination (ID)*. The first part concentrates on how users create information while the second, focuses on how the information is sent and received.

In order to model an information-flow, first its constituent domains will have to be identified. For instance, we need to establish who the information senders, receivers and subjects are. Further, we need to model what information was created in the system and how was it created and transmitted to other users.

For the information-flow in PN1, the user was both the information sender and subject. Friends (including close friends) of the user were information receivers. Status message was the information that was sent from the user to the friends. The problem frame for this information-flow is modelled as shown in Figure 4.5.

Here a mobile phone acts as connection domain between the user and the location creating machine. The user is a biddable domain representing a human operator. To create a status message, the user issues a command `Create(SM)` at interface `c` which is executed by the location creating machine. This machine sends equivalent commands to the model domain called status message where it is stored. Further, updates and deletions to the status messages are performed by the user issuing commands `Update(SM)` and `Delete(SM)` respectively. However, MFb had no facility for the user to specify if a status message was sensitive or not.

In information dissemination, the emphasis is on how the information reaches its recipients. Therefore, this second part of the information-flow model deals with the viewing or receiving of information. Normally, users are able to view information when they

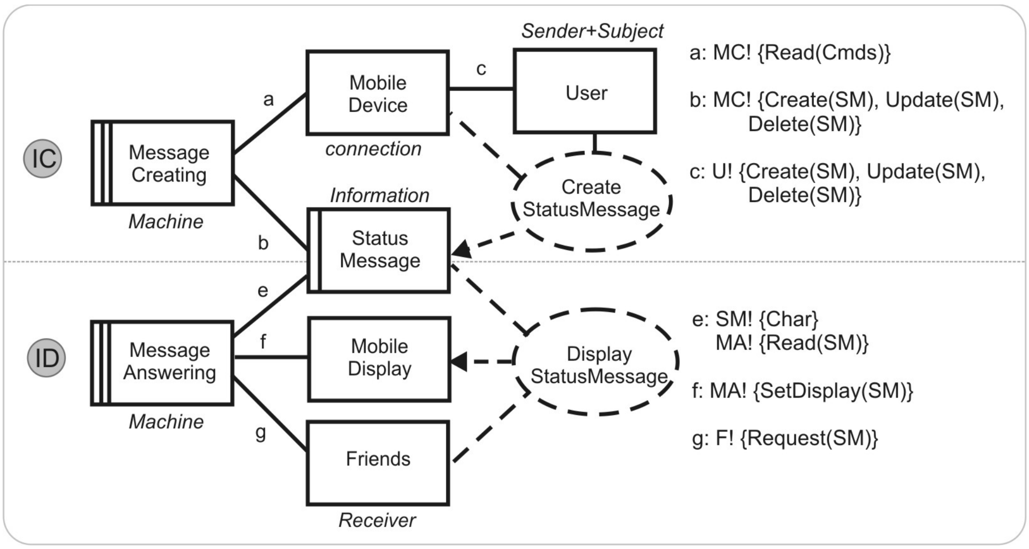


Figure 4.5: Problem frame: Information-flow in PN1

query the software system. However, in some cases this functionality is made default, as in the case of the MFb application where it is designed in such a way that the users' status message is visible to all their friends by default (i.e. immediately after they log into the system).

In the MFb application, friends are other system users who are related to the main the user. From PS-context P3-04, we can infer that there were two types of friends - ‘close friend’ and ’friend’. A ‘close friend’ being a specialisation of ‘friend’, enjoys a higher level of trust than normal friends. We can define both these relationships in an informal way as shown below.

Friend(f, u) $\approx$ *role: a person f is a known to user u*
Trusted(cf, u) $\approx$ *state: a person cf is trusted by user u*
CloseFriend(cf, u) $\leftrightarrow$ *Friend(cf, u)* $\wedge$ *Trusted(cf, u)*

In the MFb application, by default all users are able to view the status message of other users (i.e. friends) when they log-in. This is because the system automatically makes a request issuing a command `Request(SM)` at interface `g` and the message answering machine responds to this query by reading the information from the status message at interface `e` and setting the mobile display accordingly.

Although the information creation and dissemination models establish the current functionality relating to sharing of status messages in the MFb application, its control variant is required to analyse the gaps highlighted by the privacy concerns identified earlier. Thus, the next logical step is to model the *control variant* of this information-flow where the analysis focuses on (i) exerting control over the types of status messages being created, and (ii) controlling of information receivers' mobile display when requests are made for status messages.

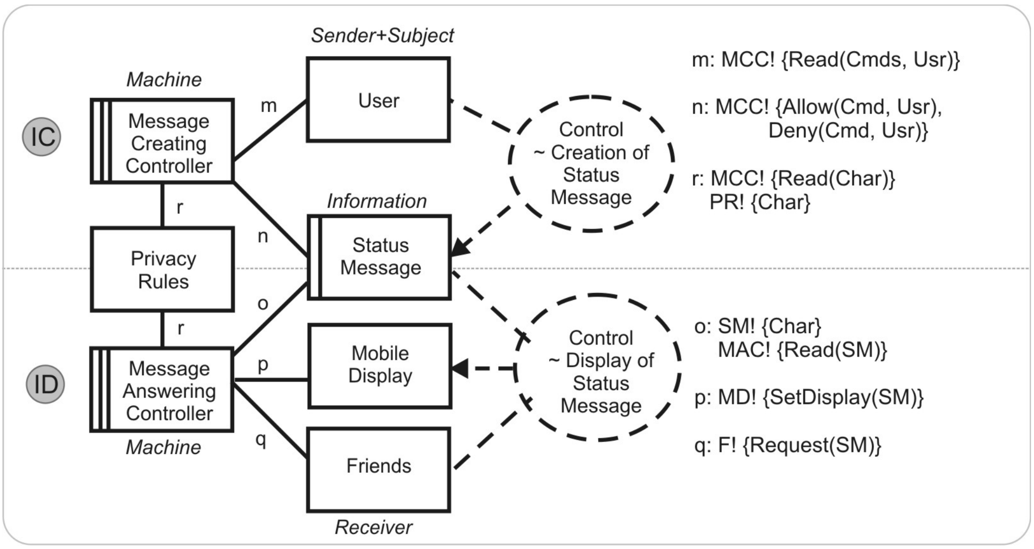


Figure 4.6: Problem frame: Control variant of information-flow (PN1)

The control variant in Figure 4.6, shows a model domain called privacy rules which contains rules to not only control the creating and editing of status messages but also for controlling information queries. The message creating controller checks the privacy rules to determine if the user is *allowed* to issue commands to create, modify, delete or forward status messages. Similarly, when the friends make a request for status messages, the message answering controller checks the privacy rules to determine if the request should be answered or not and also setting the mobile screen display accordingly. One of privacy rules which enables all friends of the user to view his status message could be of the form:

```
$\langle User \rangle$ HAS $\langle R\_friends \rangle$
$\langle User \rangle$.StatusMessage VIEWED_BY $\langle R\_friends \rangle$
..
$\langle R\_friend \rangle$ HAS $\langle O\_friends \rangle$
IF $\langle R\_friend \rangle$ LIKES $\langle User \rangle$.StatusMessage THEN
$\langle User \rangle$.StatusMessage VIEWED_BY $\langle O\_friends \rangle$
```

The above is described using our own language[1]. Here, 'R_friends' denotes all the user's friends who are receivers of status message and 'O_friends' represents others who are associated with 'R_friend' (i.e. friend-of-a-friend). Given that this is the policy used in the MFb application, the controller machines executing this policy cannot control sensitive status messages from being read by all the user's friends and their friends, leading to the privacy threats being realised.

4.5 Phase 3: Privacy problems analysis

This final phase aims to analyse the information-flow models along with the previously identified privacy concerns so that new privacy requirements are uncovered. These new privacy requirements are needed to mitigate the effects of privacy threats/concerns.

In the first privacy concern PC2.1 which relates to exposure privacy threat, the software system is unable to distinguish between sensitive and non-sensitive information. This is one of the reasons why sensitive information is visible to all friends. To address this concern, the existing message creating machine can be extended to allow users to also input an additional attribute called 'sensitivity' when creating a status message. Further, a new machine could be built that will automatically check if the information is sensitive or not, alternatively the machine can prompt the user to make this decision. As a sub-problem of the controller discussed earlier, a sensitivity detecting machine could be modelled as shown in Figure 4.7, where the machine monitors the status message being created.

[1] We extend the Structured English language (DeMarco, 1978) by incorporating new keywords namely HAS, LIKES and VIEWED_BY to reflect the properties found in the MFb application. Further we use the $< CLASS/OBJECT > . < PROPERTY >$ notation usually found in the object oriented languages to denote a class or an object and it's properties.

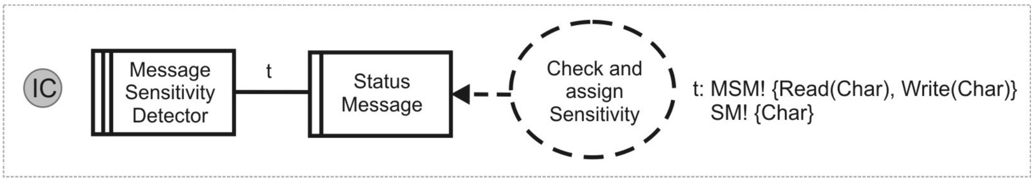

IC
Message Sensitivity Detector
t
Status Message
Check and assign Sensitivity
t: MSM! {Read(Char), Write(Char)}
SM! {Char}

Figure 4.7: Problem frame: status message sensitivity detecting

It detects if the status message is sensitive or not and then classifies it accordingly. Determining information sensitivity may not be trivial but one potential approach could be to check for words in the sentences that relate to real entities such as people, place or thing. Another strategy could be to look for key words related to emotion such as 'like' or 'love'. If the machine is unable to distinguish, it can be programmed to prompt to the user to make the decision, possibly by means of a message on the screen display. By being able to determine information sensitivity, the software system will be able to ensure that sensitive information is not visible to any unintended recipients. For instance, in this case, if the user selects a group of friends/recipients who should not be viewing the sensitive information, then the software system can immediately alert the user of potential threats. In this way, we can mitigate the effect of exposure where sensitive information cannot be accidentally leaked to a wider unintended audience. The following informal designations are used in when describing privacy requirements.

StatusMessage(m) ≈ *entity: m is the status information input by the user, describing his/her current activity*
CreateMessage(ce) ≈ *event: ce is an event in which status message is created by the user*
ViewMessage(ve) ≈ *event: ve is an event in which the user's status message is viewed by his/her friend*
StatusMessageView(m,f) ≈ *state: holds if friend f is able to see the status message m of the user in event ve*
SensitiveMessage(m) ≈ *state: a message m is said to be sensitive*
CheckingSensitivity(sme,m) ≈ *event: sme is an event in which a message m is classified as being sensitive or non-sensitive*
CreateGroup(ge,N) ≈ *event: ge is an event where a group of recipients is created with a name or label N*
AssignGroupMember(ae, $\langle g, f\rangle$***)*** ≈ *event: ae is an event in which a friend f is assigned to a group g*
∀ ***r*** · ***CloseFriendOfUser(f,u)*** ↔
(Friend(f,r) ∨ ***CloseFriend(f,u))*** ∧ ***CloseFriend(f,u)***

In order to mitigate the exposure concern, the requirement detailing the checking of information sensitivity is captured in the following argument construct.

```
argument:  Fb_Information_Sensitivity_Detecting
 PR1 Status message sensitivity can be determined mitigates PC2.1 {
 warranted by
      Cr1 System prompts user to input message sensitivity
      Cr2 If no user input in Cr1, system automatically detects
      sensitivity of status message by:
      CheckingSensitivity(StatusMsg)
      Fr3 If Cr1 is indeterminate, system prompts the user
      for input.
 }
```

Here the privacy requirements start with letter ‘Cr’ to represent control requirements and we use ‘Fr’ to denote feedback requirements.

Detecting sensitivity of information is just one part of the problem, the other relates to information receivers. In PS-context P3-03, the user wished to share sensitive status messages only with close friends. Here the concern was that the software system did not facilitate the creation of such specialised groups. Therefore, the next privacy requirement is about allowing the user to create a recipient group called ‘close friends’ with whom sensitive status messages can be shared.

```
argument:  Fb_Close_Friends_Group
 PR2 User can create a group of close friends mitigates PC2.2 {
 supported by
      E1
 warranted by
      Cr3 User issues a command to make close-friends group
      CreateGroup(‘close-friends’)
      Cr4 User assigns members of the group:
      AssignGroupMember(f,close-friends)
 }
```

On their own, the privacy requirements PR1 and PR2 may not be sufficient, another requirement regarding query answering must be defined such that only those who are members of the close-friends group may be allowed to see status messages marked as being sensitive. This is done by a additional requirement in PR3 as shown below. Incidentally, this requirement also mitigates the privacy concern of cross-contextual information (PC8) where status messages crossed over from a sensitive to a non-sensitive context.

```
argument:  Fb_Close_Friends_Viewing
 PR3 Only close friends can see sensitive status messages
 depends on (PR1, PR2) mitigates PC8 {
      Cr5 Sensitive status message is visible only to close friends:
      IF SensitiveMessage(m) ∧ CloseFriend(cf,u)
        THEN StatusMessageView(m,cf)
 }
```

We use the ‘depends on’ clause to indicate PR3 has a dependency on other requirements such as PR1 - the system’s ability to determine if the status message was sensitive or not and PR2 - the user’s ability to create a group (list) of close friends.

Now that only close friends of the user can view sensitive status messages, it still leaves one issue open - what will happen if a close friend revealed the sensitive status message to all her friends who may not be close friends of the main user. This is the privacy concern raised in PC6 - breach of trust. The software system can address this concern by doing a simple check to determine if the friends or close friends of the receivers are also the close friends of the user and thus reveal the status messages only to them, as shown below:

```
argument:  Fb_Receivers_Friends_Checking
 PR4 (preceeds PR3) mitigates PC6 {
      when CloseFriendOfUser(Receivers_friend,User)==true
 }
```

With this condition in place, even if the receivers click on the ‘like’ button, the status message will not ‘forwarded’ to other friends, instead it will now be visible only to those who are ‘close friends’ of the user, thus mitigating the privacy concern in PC6.

As shown here, a privacy requirement can mitigate one or more privacy concerns and similarly a privacy concern can rebut one or more privacy norms of a software system.

The third phase of distillation showed the working of only three examples of privacy concerns but it is possible to carry out similar analyses of additional concerns derived from the qualitative data.

4.6 Discussion

The PriS method (Kalloniatis *et al.*, 2007) uses eight categories of security and privacy principles to derive privacy requirements but these high-level principles are organisation-centric and do not cover fine-grain privacy threats a user might face when interacting with other system users. In a similar approach, Deng *et al.* (2011) have produced the STRIDE threat taxonomy which they use to identify security threat types. STRIDE is an acronym for Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, and Elevation of privilege. These threats were obtained by negating the main security properties, namely confidentiality, integrity, availability, authentication, authorization, and non-repudiation. Further, they introduce the LINDUUN methodology to elicit privacy requirements and select privacy-enhancing technologies. Unlike distillation which follows a bottom-up approach to identify real and very specific privacy threats related to a software system, the top-down approach of LINDUUN suffers a disadvantage in that it is difficult to identify a priori, all potential privacy threats that are applicable to a software system. Moreover, privacy requirements from top-down approaches such as LINDUUN can lead to over-engineered solutions. On the other hand, distillation produces privacy requirements that focus on existing privacy threats in a software system.

In Semantic Parameterization (SP) (Breaux & Anton, 2005a) (Breaux & Anton, 2005b) (Breaux *et al.*, 2006) (Breaux & Anton, 2008) privacy requirements are extracted from legal documents which are often opaque and ambiguous for software engineers who are not familiar with the lawyers' vocabulary. The methodology requires software engineers to analyse each statement from a regulation text and identify the statement as a definition, right, obligation, or constraint and then restate into what they call as restricted natural language statements (RNLS). When the RNLS are mapped into semantic models they become amenable to formal analysis. This method bears resemblance to our distillation approach, where both draw upon QDA techniques such as extraction patterns and keywords. The difference lies in the input and output of both these approaches. Semantic Parameterization uses legal text as its input which is structured in way that is less appropriate for RE but it is definitely not as imprecise and noise-filled as qualitative data which is used in the distillation approach. Secondly, Semantic Parameterization produces a set of privacy requirements which are organisation-centric, as compared to the distillation approach where user-centric privacy requirements for mobile users are derived.

4.7 Summary

In this chapter, we introduced and illustrated the use of our distillation approach. The distillation approach adapted and used thematic coding to structure raw data and to make it amenable for requirements extraction. A set of pre-defined codes specifically designed for mobile privacy was used along with behaviour patterns and emotional cues to detect and tag privacy threatening or privacy violating contexts in the data. In the next step, privacy threats, concerns and their contextual details were extracted using a combination of facet questions and data extraction rules provided by the PriF framework. Privacy norms and concerns relating to the privacy threats were then derived and represented using the privacy arguments language. Later, the information-flows linked to the privacy norms were modelled using problem patterns adapted from Problem Frames approach. These helped in understanding the application domains and its related phenomena. Finally, we showed how the privacy concerns and the information-flow model, when analysed together, help in specifying privacy requirements which mitigate the effect of the privacy concerns and uphold the privacy norms.

In this chapter, we showed how the novel distillation approach is able to successfully structure, model and analyse qualitative data for the purposes of privacy requirements engineering. This was achieved through the integration of qualitative research (Braun & Clarke, 2006; Corbin & Strauss, 2008; Gibbs, 2007; Thomas, 2006), and software engineering approaches (Jackson, 2001; Tun *et al.*, 2012). In the next chapter, we demonstrate distillation using a different qualitative dataset obtained from a study of mobile location tracking software.

5

Distilling Privacy Requirements From Multiple Viewpoints

In the previous chapter we showed the workings of the distillation approach using qualitative data from a study involving the MFb application. However, the aim here is to validate this approach by applying it on a different dataset. The objective is to prove that the distillation approach can be consistently applied, without any modifications, on other datasets and yet produce new privacy requirements. Here the qualitative data is from an empirical study which focused on location-tracking of mobile users within a family. Unlike the first study, this data contained privacy experiences from multiple perspectives - both 'victims' and 'offenders' and our aim is to use the distillation approach to capture privacy requirements from both these viewpoints.

As previously described, distillation consists of three phases. In the first phase, qualitative data is structured using the PriF framework. The second phase concentrates on modelling information-flows and in the third phase, the problem analysis of information-flow models produce privacy requirements. The following sections focus on collecting evidence for each of these analysis phases to validate the distillation approach (see Table 4.1).

The qualitative data from the location-tracking study consisted of a number of transcribed interview transcripts. The participants of the study were members of a family and were not related to the researchers. Further details on the study are available in Appendix B - Location Tracking Study. For the validation, we use two transcripts from the interview. The first participant is named as ‘CW’ (for short) and is one of the family members who took part in location-tracking. The interview data captures CW recollecting her experiences with using the location-tracking application (LT-App) on her mobile device to locate her other family members, especially her boyfriend ‘RN’.

5.1 Distillation Phase 1: Structuring of qualitative data

The first phase of distillation deals with organisation of qualitative data in order to identify relevant episodes relating to privacy threats in specific contexts i.e. privacy-sensitive contexts (PS-contexts). These PS-contexts were extracted using coding technique. In order to code PS-contexts, the interview transcripts were scanned for instances where the user would have experienced negative emotions such as fear, unhappiness etc. when using the LT-App. Alternatively, users could have exhibited negative behaviours when accomplishing a task. The negative behaviour could be where the user stops using the LT-App (disuse pattern) or accomplishes work through other means (workaround pattern). The ‘application’ or LT-App refers to the Smartphone with the software application which was evaluated for privacy.

5.1.1 Coding for PS-context

As shown previously, in order to identify PS-contexts, we look for negative behaviours (NBPs) and negative emotions (NEIs) in users’ experience. While there could be several instance of of NEIs and NBPs, our focus is only those instances which are linked to specific functionality of the software system used in the study.

During the first pass of the interview script (below), the user CW indicated that she was worried about being unable to know her boyfriend’s current location. To indicate the negative emotion exhibited by the user, this part of the data was appropriately coded with NEI(WORRY).

[CW-01(B3.18)]*'...his phone was off, actually then I was like it is 2 o'clock in the morning, I don't know where you are and I can't get hold of you, where are you and who are you with...I thought he had turned his tracker off. I thought he didn't want me to see where he was'***[NEI(WORRY)]**

CW's response is again highlighted in the following data, where the worried user who is not able to determine the current location of her boyfriend. CW explains her concern for her boyfriend whose mobile phone battery had previously discharged, requiring the LT-App to be reinstalled. After restarting his device, CW's boyfriend failed to start the LT-App leaving her with no information on his whereabouts, resulting in her worrying, this is coded as shown below.

[CW-02(B3.24)] *'Just the fact that it was off, yeah even if it was work and it was off or at home and it was off, you know, I would still think ...well, why don't you want me to see it [location] and that was one time where I got really upset it...And actually what had happened, I think on his phone because he never charged it [mobile phone] properly so when it dies he had to reinstall it [restart tracking application] so that is why it was off. The tracker was off but his phone was still on so I thought maybe his battery has just died and his phone was off but it was ringing so the fact that he had turned off [tracking application] I really [felt] like a bit cheated. Why wouldn't you want [me] to know where you are?'* **[NEI(WORRY)]**

In any PS-context, the LT-App users can take on one of two roles: *tracker* and *trackee*. A user whose location is being sampled and stored is a trackee whereas a user who makes enquiries for another user's (trackee's) location is said to be a tracker.

PS-contexts CW-01 and CW-02 above are both linked to each other and it would suffice to use the second one (CW-02) as it contains additional details about the tracker and trackee.

CW worked on the same road as her boyfriend, but inaccuracies in location made it difficult for her to determine the exact location of her boyfriend. This caused CW to be upset because she inferred her boyfriend was not at work and was elsewhere and did not want to talk with her. This can be seen in the next PS-context identified.

[CW-03(B3.28)] *'...at Westfield the [network] reception is not great so it doesn't pinpoint where exactly where [boyfriend's] work was. Sometimes it was just around [a round circle on the map]. I remember first times I saw it when [he] wasn't at work and I thought he was at his friends house, and I thought that's fine, I don't mind but you said you couldn't phone me because you were at work...Well you are not speaking to me because you are at work but you are not at work, you are down the road and I can see where you are. Then I realised he was [still] at work so then I felt a bit bad about that...Then I realised it [tracking application] didn't always pinpoint the exact spot all the time'* **[NEI(UPSET)]**

In the above, it is not always easy to clearly assign codes as both PS-contexts indicate more than one negative emotion. For instance, the PS-contexts CW-01 and CW-02 could also be tagged for NEI(WORRY) or NEI(ANXIOUS) but what is important is that it is tagged with at least of them - the codes in distillation are designed to cope with such variations.

While analysing the transcript of JW (sibling of CW), we come across an episode where JW switched-off her mobile phone in order to meet and chat with her close friend at a certain restaurant. Since this matches the disuse pattern in the PriF framework where users might stop using their mobile phones for privacy reasons, the following data is coded for DISUSE.

[JW-04(B4.195,197)] *'I went for drinks with one of my best friends from [another country] who just moved here and we are very close and we have this, we talk about personal stuff so it's got a long history so there's always lots to catch up on and she's had quite a roller coaster the last few weeks. We hadn't seen each other so we met up and it's always like a very close personal things, silly girly gossip whispering you know all this stuff and I quite liked knowing that it was really private from everybody. Everyone knew I was having some drinks because I tell everyone and she tells everybody, but while we were chatting it was just us and I quite liked that because that's how we open up to each other...You are conscious that other people, out of love, out of interest, or positive things but they are kind of aware [of location]. Which is fine there is nothing inherently wrong about it but sometimes you do just want to close the curtains.' [The user had switched-off the mobile phone]* **[NPB(DISUSE)]**

Once NEIs and NBPs are coded, the next step involves coding for privacy determinants - the properties which influence privacy.

5.1.2 Coding privacy determinants

While NEIs and NBPs point to specific privacy issues the user might have experienced, at the onset it may not be clear whether the PS-context is associated with a specific privacy threat or not. For this purpose, the Facet questions from the PriF framework are used to sift through potential privacy determinants that may be associated or contributing to the NEIs or NBPs of users.

In PS-contexts CW-01, CW-02, and CW-03, it was found that the user was either worried or upset when checking the location of her boyfriend. While this is an indication that there could be potential privacy threats, it may not be clear which of those privacy threats apply in this case. By identifying the privacy determinants contributing to the user's emotional states such as 'worry' and 'upset', we can isolate the specific privacy threat and its associated concerns.

As described previously, the Facet questions help to explore the PS-context in much greater detail, especially to identify the type of information, the actors involved, the information-flow produced and the place. Although each of these facets elicit information about the PS-context, specific factors, namely the privacy determinants, contribute directly to the NEI or NBP. Therefore, the aim here is to identify these privacy determinants and code for them.

In this case, when the Facet questions are applied to PS-contexts - namely CW-01 and CW-02, the information facet stands out because the user's worry relates to the quality of location information. When the location of CW's boyfriend is *not fresh*, she initially assumes that the LT-App might have been switched off - *'I thought he had turned his tracker off'*. One possible explanation could be that the user may have seen that the location of her boyfriend hasn't changed for several minutes or hours. However, had the LT-App been working normally, the location would have changed/refreshed every 10 minutes. Therefore, the user correctly assumes the LT-App had entered into an error condition when the location is not updated. Using the codes provided under each Facet (see Table 3.3), the PS-context can be coded appropriately. Since in this PS-context, the user's worry [NEI(WORRY)] is directly related to the information attribute (I-ATTR) such as freshness of location (FRESH), we tag the PS-context as shown below.

> **[CW-01(B3.18)]** *'...his phone was off, actually then I was like it is 2 o'clock in the morning, I don't know where you are and I can't get hold of you, where are you and who are you with...I thought he had turned his tracker off. I thought he didn't want me to see where he was'***[NEI(WORRY), I-ATTR(FRESH)]**

PS-context CW-02 is related to CW-01 as it follows from the previous PS-context where the user is concerned about not being able to determine the location of her boyfriend, the difference here is that she explains the possible reasons why the location data was not fresh. We tag this PS-context with the same code because it relates to freshness of location data.

> **[CW-02(B3.24)]** *'Just the fact that it was off, yeah even if it was work and it was off or at home and it was off, you know, I would still think ...well, why don't you want me to see it [location] and that was one time where I got really upset it...And actually what had happened, I think on his phone because he never charged it [mobile phone] properly so when it dies he had to reinstall it [restart tracking application] so that is why it was off. The tracker was off but his phone was still on so I thought maybe his battery has just died and his phone was off but it was ringing so the fact that he had turned off [tracking application] I really like a bit cheated. Why wouldn't you want [me] to know where you are?'* **[NEI(WORRY),I-ATTR(FRESH)]**

The LT-App is configured to update the location of CW's boyfriend every 10 minutes (see Appendix B for more details on the functionality of the location-tracking system) and there could be several reasons why it did not update according to the expected frequency. Some potential reasons could be (i) the mobile device was switched-off (ii) the LT-App was switched-off, and (iii) the trackee changed his settings to avoid being tracked.

In the third PS-context CW-03, the user CW is again concerned with the location of her boyfriend, however, here the cause is slightly different from the first two. Here the issue relates to the accuracy of location data. CW expected her boyfriend to be at work but the LT-App displayed a location that was vague and inaccurate, causing her to misunderstand her boyfriend's location. Since inaccuracy of location data played a significant role, we can tag this PS-context for accuracy (ACCU) under information attributes (I-ATTR) of the Information facet.

[CW-03(B3.28)] *‘...at Westfield the [network] reception is not great so it doesn’t pinpoint where exactly where [boyfriend’s] work was. Sometimes it was just around [a round circle on the map]. I remember first times I saw it when [he] wasn’t at work and I thought he was at his friends house, and I thought that’s fine, I don’t mind but you said you couldn’t phone me because you were at work...Well you are not speaking to me because you are at work but you are not at work, you are down the road and I can see where you are. Then I realised he was [still] at work so then I felt a bit bad about that...Then I realised it [tracking application] didn’t always pinpoint the exact spot all the time’* **[NEI(UPSET), I-ATTR(ACCU)]**

It may not always be the case that users are able to readily identify reasons for data quality issues (e.g. freshness, inaccuracy etc.) they experience while using location-tracking applications (as shown here). Normally, it is the analyst’s task to examine the potential causes and to look for evidence in qualitative data or alternatively examine the application’s system design to determine the causes. However, here it was rather straightforward because the user was able to give definitive reasons as to what had gone wrong with the LT-App when it produced inaccurate or missing location data.

When iterating through the Facet questions for JW-04, the PS-context had indicated that the user (trackee) switched-off her mobile device so that she is not disturbed when in the company of her best friend. While the place was not very obvious from the context, it appeared the user *‘went for drinks’* with her best friend and the place could have been a restaurant or pub. The switching-off of her mobile phone by JW [NBP(DISUSE)] in the presence of her best friend shows that the privacy determinant here is a social rule or etiquette. Thus, the following is coded for PLACE(ETIQTE) as shown below.

[JW-04(B4.195,197)] *'I went for drinks with one of my best friends from [another country] who just moved here and we are very close and we have this, we talk about personal stuff so it's got a long history so there's always lots to catch up on and she's had quite a roller coaster the last few weeks. We hadn't seen each other so we met up and it's always like a very close personal things, silly girly gossip whispering you know all this stuff and I quite liked knowing that it was really private from everybody. Everyone knew I was having some drinks because I tell everyone and she tells everybody, but while we were chatting it was just us and I quite liked that because that's how we open up to each other…You are conscious that other people, out of love, out of interest, or positive things but they are kind of aware [of location]. Which is fine there is nothing inherently wrong about it but sometimes you do just want to close the curtains.' [The user had switched-off the mobile phone]* **[NPB(DISUSE), PLACE(ETIQTE)]**

Now that all four PS-contexts have been coded for their privacy determinants, the next step is to derive privacy concerns from it.

5.1.3 Deriving privacy threats and concerns

In order to identify the type of privacy threat associated with a PS-context we run rules provided in Table 3.8. In this case, upon processing these rules, the PS-contexts CW-01, CW-02 and CW-03 are all shown to be linked to the 'misinformation' privacy threat since they satisfy rules associated with it as shown below:

(T5) Misinformation: *Information attributes (I-ATTR) AND*
(any sub nodes of NEI OR NBP)

This privacy threat relates to information attributes contributing to misinformation about users. In CW-01 and CW-02, the user CW was worried (NEI-WORRY) about the freshness [I-ATTR(FRESH)] of her boyfriend's location and therefore these PS-contexts satisfy the conditions for the above privacy threat.

CW-03 also satisfies the conditions for this privacy threat because CW was upset [NEI(UPSET)] with her boyfriend when the LT-App displayed his inaccurate [I-ATTR(ACCU)] location.

For JW-04, the processing of rules show it satisfying the conditions for the 'intrusion' privacy threat whose extraction rule is shown below:

(T10) Intrusion: *Place - Etiquette (ETIQTE) AND (any sub nodes of NEI OR NBP)*

This privacy threat relates to violation of privacy etiquette at a place. With JW-04 being associated with 'intrusion' privacy threat, it now confirms why JW would have wanted to switch-off her mobile phone - any ringing or even tracking of her would have intruded on her solitude and her need to be left alone with her friend. This again confirms how the distillation approach is able to successfully link PS-contexts with specific privacy threats and also account for users' privacy related negative behaviours.

After identifying privacy threats associated with PS-contexts, the next step is to derive privacy concerns from them. Privacy concerns as we know, highlight the ways in which a software system is unable to prevent privacy threats from being realised. Since privacy concerns do not exist on their own but only in relation to existing privacy norms, these norms have to be described before privacy concerns.

As described earlier, in the distillation process, privacy norms (either positive or negative) state what information-flows are legitimate or allowed within specific contexts. In this case, the PS-contexts CW-01, CW-02 and CW-03 are associated with the user querying for her boyfriend's location, this privacy norm is described as follows:

```
argument:  LT_Boyfriend_Location_Norm
 PN1 ALLOW[+⟨ Boyfriend → User, location(Boyfriend), ∗‖
providing_movement_update ⟩]{
 supported by
      F1 User has boyfriend
      F2 User needs to know her boyfriend's location
 warranted by
      R1 The boyfriend's location is automatically sampled
      every 10 minutes
      R2 The location is stored in a database on a remote server
      R3 User taps on her boyfriend's icon to make a location query
      R4 When user taps boyfriend's icon, the location is
      retrieved from the database and displayed on the screen
      using a map
 }
```

The norm PN1 indicates that in a family context, a user should be able to view her boyfriend's location. Since, the PS-contexts (CW-01,CW-02,CW-03) don't indicate the number of times or how often this should be possible, we can assume that there are no restrictions with regards to the tracker or trackee.

The evidence from PS-contexts CW-01 and CW-02 (‘I thought he had turned his tracker off’) suggested that turning the LT-App off can lead to misinformation i.e. CW’s boyfriend is hiding his location. Since the concern here is that the LT-App is unable to prevent this misinformation threat from being realised, we capture this concern as follows:

```
argument:  LT_Misinformation_Concern_1
 PC5.1 ⟨Unavailability of timely location updates,
 (H3 : causes loss of reputation for Boyfriend)⟩ rebuts PN1{
 supported by
      E3 The mobile device can be switched-off
      (CW-01:  ‘...his phone was off’)
      E4 LT-App can be switched-off
      (CW-02:  ‘...the tracker was off but his phone was still on’)
      E5 The mobile device is switched-off because of
      an discharged battery
      E6 LT-App may fail to start or get into an error mode
      (CW-02:  ‘...and actually what had happened,
      I think on his phone because he never charged
      it [mobile phone] properly so when it dies he
      had to reinstall it [restart tracking application]
      so that is why it was off’)
 }
```

While the above concern was derived from PS-contexts CW-01 and CW-02, it did not include CW-03 because in this case the source of misinformation was slightly different to the one described above. In CW-03, misinformation was caused by location inaccuracy rather than unavailability of location data. In CW-03, the software system produced the trackee’s inaccurate location causing the tracker to misunderstand. This concern is captured below:

```
argument: LT_Misinformation_Concern_2
 PC5.2 $\langle$ $Inaccurate\,location\,sampling, (H3: causes\,loss\,of\,reputation\,for$
 $Boyfriend)\rangle$ rebuts PN1{
 supported by
      E7 The location created by LT-App is inaccurate
      (CW-03: '...at Westfield the [network] reception
      is not great so it doesn't pinpoint where exactly
      where [boyfriend's] work was')
 }
```

As seen above, both these privacy concerns rebut the privacy norm stated in PN1. The privacy concerns are labelled as PC5.1 and PC5.2 to show that they are instances of concerns derived from T5-Misinformation concern. Protecting the privacy of the trackee, in this case CW's boyfriend, will now depend on successfully deriving privacy requirements which mitigate the effects of both these privacy concerns.

For JW-04, the privacy norm is slightly different to PN1. Here, the privacy norm relates to the user (trackee) allowing her family members (trackers) to view her location and day-to-day movements but not when she is with her close friend in a private meeting place. This is a negative privacy norm, an information-flow that should not be supported in the LT-App, is captured as shown below:

```
argument: LT_Private_Meeting_Norm
 PN2 $DENY[-\langle User \rightarrow FamilyMember, location(User),$
$(place = pub|restaurant)||meeting\_friend\,\rangle]${
 supported by
      E8 UserJW wants to switch-off LT-App
      (JW-04: '...You are conscious that other people,
      out of love, out of interest, or positive things but they are
      kind of aware[of location]...but sometimes you do just want to
      close the curtains [wants to be left alone]')
 }
```

Previously, the *T10-Intrusion* privacy threat was associated with PS-context JW-04. Exploring the context details, we discovered that the user went for a drink with her best friend and therefore switched-off the mobile device. Even though the user was able to manually switch-off the mobile device, it was evident the software system had no facility to automatically switch off when it detected the user was in the company of her close-friend in a restaurant or pub. Instead, the LT-App allowed everyone to locate her at a private meeting which is both a violation of her freedom/autonomy (harm H6) and anonymity (harm H7) - this intrusion privacy concern is captured below.

```
argument:  LT_Intrusion_Concern
 PC10 $\langle\ System\ allows\ determining\ of\ trackee'slocation\ when$
 $place = pub|restaurant,\ (H6, H7 : disturbs\ private\ meeting)\rangle$ rebuts PN2{
 supported by
      F3 User has family members
      F4 Family members want to know the user's location
 warranted by
      R5 The user's location is automatically sampled
      every 10 minutes
      R6 The location is stored in a database on a remote server
      R7 Family member taps on user's icon to make a location query
      R8 When family member taps user's icon, location is retrieved
      from the database and displayed on the screen using a map
 }
```

The following section will demonstrate how new privacy requirements can be derived to actually address the privacy concerns identified here.

5.2 Distillation Phase 2: Information-flow modeling

In the second phase, information-flows in PN1 and PN2 are modelled with respect to their constituents such as information senders, receivers and subjects. Later, we model control variants of PN1 and PN2 to establish its current privacy controls (if any) and use them to analyse privacy concerns.

As described previously, information-flows consist of two parts - information creation and dissemination, both of these are modelled for PN1 as shown using a problem diagram in Figure 5.1.

In LT-App, a trackee's location information consists of geographical coordinates (longitude and latitude) which is determined using a global positioning system (GPS) on board the mobile device. The LT-App is by default configured to sample the location

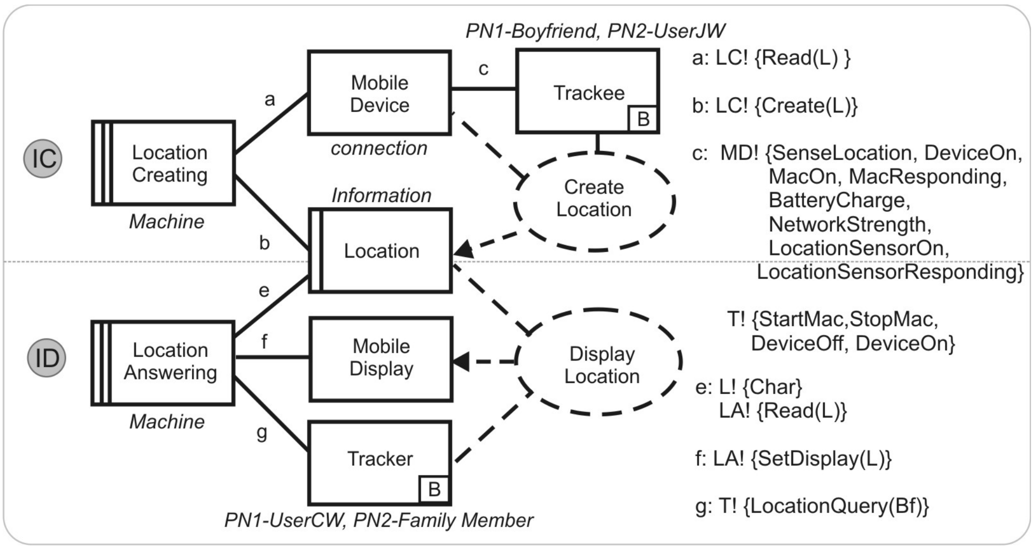

PN1-Boyfriend, PN2-UserJW
IC
Location Creating
Machine
a
b
c
Mobile Device
connection
Trackee
B
Create Location
Information
Location
ID
Location Answering
Machine
e
f
g
Mobile Display
Display Location
Tracker
B
PN1-UserCW, PN2-Family Member
a: LC! {Read(L) }
b: LC! {Create(L)}
c: MD! {SenseLocation, DeviceOn, MacOn, MacResponding, BatteryCharge, NetworkStrength, LocationSensorOn, LocationSensorResponding}
T! {StartMac,StopMac, DeviceOff, DeviceOn}
e: L! {Char}
LA! {Read(L)}
f: LA! {SetDisplay(L)}
g: T! {LocationQuery(Bf)}

Figure 5.1: Problem frame diagram: information-flow in PN1 and PN2

of the trackee every 10 minutes. This is because a higher sampling rate would drain the battery of the mobile device and impact the system's performance. Once the location of the trackee is determined, location information can be made available to trackers who may request it. PN1 is an instance of this information-flow and it can be modelled using a problem frame diagram as shown in Figure 5.1.

In PN1, UserCW is a tracker and her boyfriend takes on the role of a trackee. A mobile device periodically invokes the command `SenseLocation` on the embedded location sensor (GPS) to capture location coordinates. Once the location has been captured/sensed, the location creating machine then reads the location data through the command `Read(L)` at interface `a`. In PN2, a family member becomes the tracker and the user is the trackee. The following descriptions are relevant for location creation.

Trackee(te) ≈ *role: trackee te is a person whose current geographic location is stored in the system*
Tracker(tr,te) ≈ *role: tracker tr is a person who makes a request/enquiry for the location of trackee te*
Location(x) ≈ *entity: x is an entity which contains the location coordinates at a given place*
SenseLocation(s) ≈ *event: s is an event in which a device's location is determined*
LocationSensed(d,x,s) ≈ *state: holds if location x for device d was produced by location sensing event s*

The specification for the location sensing regime can be stated as:

For every 10 minutes
LocationSensed(d,x,s) holds true

When UserCW makes a request for her boyfriend's location using `LocationQuery(Bf)`, the location answering machine reads the location and sets the mobile display accordingly. The information query can be described as:

LocationQuery(q,tr,te) ≈ *event: q is an event where a tracker tr makes a request/enquiry for the location of trackee te*

In Figure 5.2, a control variant of the above information-flow, models existing controls for both location creation and dissemination.

By default, the LT-App provides trackees with some privacy controls in the form of 'filters' to be applied over location information. These are:

- *access control of location:* who should be able to see trackees' current location
- *granularity of location:* trackees can set at what level of granularity trackers can see their location (e.g. display city instead of street or vice versa).

In Figure 5.2, trackees can exploit both these controls by issuing a `SetFilter(F)` command at interface `s`, where F stands for a set of filtering rules. For instance, if a trackee wants to deny access to a tracker, this can be achieved through the use of a filtering rule as shown below.

Syntax: ALLOW/DENY (USER ⟨type of access⟩ DATA);

For location information, this can be instantiated as:

DENY(User ⟨ALL-VIEW on⟩ Location);

or set a rule to reduce the granularity of the location:

ALLOW(User ⟨CITY-VIEW on⟩ Location);

In addition to controlling access and granularity of location, trackees can also issue commands to stop and start the LT-App (`StopMac / StartMac`), thus, controlling the creation of location information. When the LT-App is switched-on, it implies that the location creation machine is switched-on. This state of the machine is stored in phenomena `MacOn` where its value is set to true. There could be other situations where the machine may be switched-on but it may not respond to `LocationSense` commands. When this happens, the machine can be considered to be in an unresponsive state and the phenomena `MacResponding` is set to false.

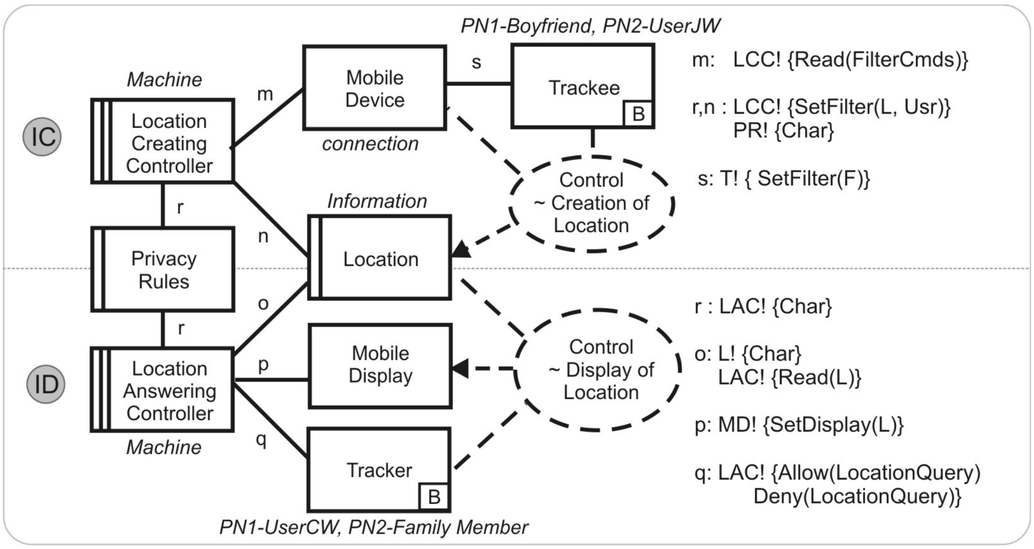

PN1-Boyfriend, PN2-UserJW
Machine
IC
Location Creating Controller
m
Mobile Device
connection
s
Trackee
B
Control ~ Creation of Location
Information
Location
r
n
Privacy Rules
r
o
ID
Location Answering Controller
Machine
p
Mobile Display
q
Control ~ Display of Location
Tracker
B
PN1-UserCW, PN2-Family Member
m: LCC! {Read(FilterCmds)}
r,n : LCC! {SetFilter(L, Usr)}
PR! {Char}
s: T! { SetFilter(F)}
r : LAC! {Char}
o: L! {Char}
LAC! {Read(L)}
p: MD! {SetDisplay(L)}
q: LAC! {Allow(LocationQuery)
Deny(LocationQuery)}

Figure 5.2: Problem frame diagram: control variant of information-flow - PN1

Mobile devices can be switched-off or switched-on by their users. Since the location creating machine is deployed and run within mobile devices, any 'switch-off' can seriously impact the collection of location data. The phenomena `DeviceOn` is set to true only if the mobile device is running.

A location sensor (GPS) inside the mobile device can also impact location creation. When the location sensor is switched-off or it enters into an unresponsive state, the location creating machine may not be able to create users' location data. These phenomena can be described informally as shown below.

MacOn(mo) $\approx$ *state: mo holds true if location creation machine is started and running*
MacResponding(mr) $\approx$ *state: mr holds true if location creation machine is running (MacOn=true) and successfully executes commands*
DeviceOn(do,d) $\approx$ *state: do holds true for mobile device d if d is started and it is running*
LocationSensorOn(so,d) $\approx$ *state: so holds true for mobile device d if the location sensor is started and is running*
LocationSensorResponding(sr,d) $\leftrightarrow$ *LocationSensorOn(d)* $\wedge$ *LocationSensed(d)*

As described earlier, trackees are able to control access to their location by specifying a set of filtering rule. These filtering rules are actually privacy rules. When a tracker issues a location request in the form of a `LocationQuery` for a specific trackee, the location answering controller uses the filtering/privacy rules to first check if the requester is allowed to access the trackee's location information. Secondly, the machine checks if the location information is at an appropriate level of granularity. Only when the request satisfies both these conditions, will trackee's location be set on the requester's mobile display.

So far, we have described the privacy norms PN1 and PN2 in a generic manner. Now we will instantiate the roles of trackers and trackees in PN1 and PN2. In PN1, the tracker was CW (UserCW) and the trackee was her boyfriend. Similarly, for PN2, we can instantiate the information-flow model by having the trackee as JW (UserJW) and her family members as trackers.

There were two main reasons for modelling information-flows in PN1 and PN2. First, it showed how location information is created within the LT-App and how the tracker made queries to access and view it. Secondly, a control variant of the models highlighted the LT-App's existing privacy controls and how it was used by the trackees to control their privacy. In the next section, we analyse these information-flow models against privacy concerns that were identified in the first section.

5.3 Distillation Phase 3: Privacy problems analysis

In this phase, the aim is to analyse privacy concerns against information-flow models which were developed earlier. By uncovering reasons as to why and how existing system requirements had failed, new privacy requirements can be discovered to address the privacy of users.

Earlier, we had modelled the information-flow in PN1 along with its control variant which established what location information is being shared between UserCW (tracker) and her boyfriend (trackee). The relevant PS-contexts showed that there were at least two instances where the T5-Misinformation privacy threat could have been realised. These two privacy concerns are analysed here.

PC5.1 Lack of location updates leads to misinformation

In the first privacy concern, UserCW boyfriend's reputation was at risk when his location information was not available to UserCW. Figures 5.1 and 5.2 show that UserCW is connected to the location creating machine via a mobile device. In other words, the mobile device is a connection domain. Generally, connection domains can be affected by various environmental factors which can impact users and their location creating machines. Thus, when the mobile device experienced power supply issues (e.g. uncharged battery) or it was switched-off by UserCW's boyfriend, it affected the machine which was creating his location data. Although the user had reported two instances where the mobile device affected location creation, there could be additional phenomena within the mobile device which may be interfering with the working of the machine and indirectly contributing to the misinformation privacy concern. It is important that these additional contributing factors are analysed in order to address this privacy concern.

So, analysing the mobile device (i.e. connection domain) and the location creating machine, it appears, there could be several phenomena which impact creation of location data and they are:

- Location Creating
  (a) User can switch-off the machine using `StopMac` command
  (b) Machine can enter into an error mode (does not respond)

- Mobile Device
  - (c) Location sensor may not respond to `SenseLocation` command
  - (d) Mobile device may be out of network coverage range
  - (e) Mobile network may be deactivated (e.g. credits exhausted)
  - (f) User can switch-off mobile device using `StopMac` command
  - (g) Battery has no charge.

In the above, some are legitimate functional requirements provided for users. For instance, in (a) and (f), users (trackees) are allowed to switch off their machines/mobile devices. There could be various reasons why users will want to do this (e.g. when in a meeting or to conserve battery power, etc.). In such unavoidable situations, software systems should allow users to switch-off their device while simultaneously providing measures to overcome the misinformation concern. One possibility could be to provide feedback to the requester (tracker) indicating the current state of the machine/mobile device. For example, requesters (trackers) can be informed about users' (trackees') machine/device being switched-off. By providing feedback in situations where users'(trackees') actions cannot be avoided, the risk of trackers misinterpreting trackee's location will be greatly reduced because of the additional contextual information being shown to them.

Therefore, the privacy requirement here will be to monitor various phenomena belonging to both the mobile device and the location creating machine such that corrective actions are initiated when necessary. This privacy requirement can be stated as:

```
argument:  LT_Achieve_Location_Data
 PR1 Achieve creation of location data mitigates PC5.1{
 warranted by
      Cr1.  MacOn=true, else inform tracker
      Cr2.  MacResponding=true, else alert user or
           automatically restart machine
      Cr3.  LocationSensorResponding=true, else alert user or
           automatically restart the location sensor
      Cr4.  NetworkStrength ≥ 20%, else
           inform tracker about network coverage
      Cr5.  DeviceOn=true, else tracker is informed
      Cr6.  BatteryCharge ≥ 10%, else it is recharged
 }
```

In control requirements Cr4 and Cr6, assumptions have been made with regards to the minimum level of network strength ($\geq 20\%$) and battery charge ($\geq 10\%$) required to enable the machine to successfully create location data for the trackee. Some control requirements should be treated as sub problems and be analysed separately. For instance, sub problem Cr6 could be addressed in several ways. One approach could be to develop a special type of battery that will automatically recharge itself through non-standard sources such as body movements or vibrations. Similarly, sub problem Cr2 may require the development of a new machine to monitor and reset the location creating machine and location sensors when they become unresponsive as shown in problem diagram in Figure 5.3.

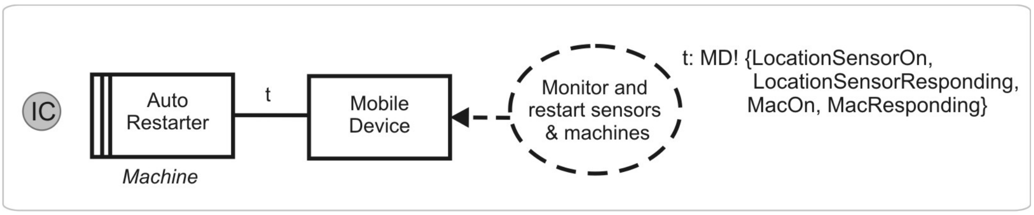

IC
Auto Restarter
Machine
t
Mobile Device
Monitor and restart sensors & machines
t: MD! {LocationSensorOn, LocationSensorResponding, MacOn, MacResponding}

Figure 5.3: Problem frame: sensor and machine auto restarting

PC5.2 Inaccurate location causes misinformation

A second factor contributing to the misinformation concern is the *inaccuracy* of location. As described in PC5.2, the accuracy of location has a direct impact on trackee's reputation, therefore, the privacy requirement is to ensure the trackee's location is accurate in order to protect him from reputational harm.

The first task is to define what can be considered as accurate and what is feasible using current mobile technologies. In PS-context CW-03, the user was shown her boyfriend's location as a circular region on a map and because the circle's radius was large it contributed to the location appearing vague. The tracker however expected to see the precise location of her boyfriend ([CW-05(B3.28)] *'didn't always pinpoint the exact spot all the time'*). In LT-App, the accuracy of location is measured by the radius of the circle drawn around a point on the map. Current mobile devices depend on several techniques to accurately pinpoint the location of users, in addition to GPS, they could also make use of cell towers and Wifi points to triangulate and determine the position. However, all these technologies may not produce optimum data and thus a measure such as ‘level of confidence’ is used to indicate the probability of correctly determining a user’s location. If accuracy is to be defined for the LT-App, then it has to be defined at least using both these measures - location radius and the level of confidence.

Location accuracy is dependent on a number of phenomena and states in a mobile device, namely (i) location sensor should be switched on (ii) location sensor’s GPS should be able to receive satellites signals, and (iii) availability of network coverage at a given place. The descriptions for these states are shown below:

LocationAccurate(d, x) $\approx$ *state: holds for mobile device d at location x if location radius <= 10 meters and confidence >= 80%*

GPSLinkToSatellite(d,x) $\approx$ *state: holds for mobile device d when its GPS has unobstructed line of sight to four or more GPS satellites at a given location x*

NetworkCoverage(d, x) $\approx$ *state: holds for mobile device d if it can access a network service at a given location x*

LocationSensorOn(d) $\approx$ *state: holds for mobile device d if it is switched-on and the GPS is functioning*

$\forall$ ***d, x*** $\cdot$ ***LocationAccurate(d, x)*** $\leftrightarrow$ ***GPSLinkToSatellite(d,x)*** $\wedge$ ***NetworkCoverage(d, x)*** $\wedge$ ***LocationSensorOn(d)***

On careful observation, the above statements point to location (and its accuracy) being determined for the mobile device and not the person who uses it. For instance, consider what would happen if trackees left their device at home or lost it while commuting, the location information shown to trackers would be that of the mobile device and not the trackee. Trackers who may not be aware of this, are most likely to accept this trackee’s current location, thus continuing to contribute to the threat of misinformation. To avoid this situation, earlier designations will have to be amended to ensure the trackee is always in possession of the mobile device when the location is sensed/determined:

MobileDeviceWithUser(d, p) ≈ *state: holds for mobile device d if it is attached to or in contact with the person p*

The description for determining accurate location can now be restated as:

$$\forall\ \boldsymbol{p, x} \cdot \boldsymbol{\mathit{LocationAccurate}(p)} \leftrightarrow \exists\ \boldsymbol{d} \cdot \boldsymbol{\mathit{GPSLinkToSatellite}(d,x)} \wedge \boldsymbol{\mathit{NetworkCoverage}(d, x)} \wedge \boldsymbol{\mathit{LocationSensorOn}(d)} \wedge \boldsymbol{\mathit{MobileDeviceWithUser}(d, p)}$$

The above descriptions ensure that the location sensed is always that of the user and not of the mobile device. Thus, the privacy requirement here is to achieve accurate location in order to avoid the misinformation privacy concern. This can be stated as:

```
argument:  LT_Achieve_Accurate_Location
 PR2 Achieve accuracy of location mitigates PC5.2{
 warranted by
      Cr7.  Location is accurately sensed for trackee:
      LocationAccurate(trackee) holds true
 }
```

Some aspects of this privacy requirement are specific to the use of location sensor technologies. When these technologies change and mobile devices are upgraded, additional analysis may have to be performed on to identify points of failure in achieving the same level of location accuracy.

PC10 Determining trackee's location when at pub/restaurant with best friend causes intrusion

The privacy requirement here is to ensure the trackee's location is not sensed and collected when she is in the company of her best friend. Achieving this will prevent the trackee from being disturbed because the location of the pub where she is having the drinks will not be visible to the rest of her family. The current software system cannot automatically stop sensing the location of the trackee, it has to be done manually where the trackee issues commands such as `StopMac` (at interface `c`) to stop the location creation machine from determining the location (see Figure 5.1). If the trackee failed to issue this command, then the intrusion privacy threat will be realised.

If the location creating machine is to automatically stop sensing and collecting the trackee's location in certain situations like the one above, it will have to do two things (i) detect the place (i.e. restaurant or pub) and (ii) detect if the trackee's close friend is co-located. The detection of close friend can be described as:

CloseFriendWith(p,cf) ≈ *role: person p is a close friend of person cf*
CoLocated(p1,p2) ≈ *state: holds if the sensed location of person p1 is the same as p2*

The privacy requirement can be described as shown below.

```
argument:  LT_Auto_Location_switch-off
 PR3 Automatically switch-off location creating
 when trackee (te) is in pub/restaurant with close friend (cf)
mitigates PC10{
 warranted by
      Cr8.  SET MacOn == FALSE ↔ CloseFriendWith(te,cf) ∧
      CoLocated(te,cf) ∧ location = pub/restaurant
 }
```

The above privacy requirements indicate two sub problems (i) checking if the trackee is at a pub or restaurant, and (ii) checking if the trackee is co-located with her best friend. Both these sub problems will have to be analysed separately and addressed.

Composition concerns

Distillation actively supports problem decomposition in the form of information creation and dissemination as two sub problems of an information-flow. But after deriving privacy requirements under each sub problem, they are recomposed into a composite problem to check the requirements for two types of concerns (i) *consistency* - in what respects are the requirements consistent and is it possible to satisfy them, and (ii) *precedence* - if they are not consistent, which requirement takes precedence.

In some cases, the newly derived privacy requirements are inconsistent because they conflict with existing functional requirements or other privacy requirements of the software system, in such cases it is important to resolve the conflicts through precedence rules or prompting the user to make the choice over its execution. In distillation, preferences over two requirements can be defined using *'preferred by'* and *'precedes'* constructs within the Privacy Arguments language (Tun *et al.*, 2012).

As described earlier, PR3 will not allow for location data to be created if the machine detects the trackee is with her close friend and is in a pub/restaurant, but this cannot be satisfied because the goal of PR1 is to achieve creation of location data. There are two possible scenarios:

(i) PR1 take precedence over PR3

In this case, the location data will be created to prevent any misinformation privacy threat from being realised but it will allow intrusion. One way to limit intrusion would be to specify additional conditions such as when the location query is made by the partner of the user, as shown below.

```
PR1 preceeds PR3
     when (tracker == Boyfriend)
```

(i) PR3 take precedence over PR1

Here the intrusion privacy threat may be prevented because PR3 is preferred over PR1 but it does not protect the trackee from misinformation threat. Again, there could be situations where the trackee will prefer to have this requirement satisfied over PR1, for example where the trackers are her family members, as shown below.

```
PR3 preceeds PR1
     when (tracker != partner && tracker == Family member)
```

The above described privacy requirements are inconsistent with one another but this is not the only type of inconsistency that might occur. Privacy requirements could also be inconsistent with other existing functional/non-functional requirements. For example, *PR2 - achieve accuracy of location* makes sense only if the user (trackee) had defined a rule such that the requester (tracker) is allowed to view the location information at the same level of granularity as it was created. In other words, if the location was accurately captured at street-level (say with an accuracy of 10 metres), it will be useful in preventing misinformation privacy threat only if the tracker's viewing setting was such that he/she can see the location at street-level and not at city-level granularity.

In this case, the user can be prompted to show the inconsistency in the privacy rule he/she has defined for the tracker(requester).

The derived privacy requirements can also create new dependencies with other functional or non-functional requirements of the software system. For example, it is quite implicit that the privacy requirement *PR1 - achieve creation of location data* depends on an existing requirement where the user is expected to switch-on the LT-App. If the trackee deliberately wants to switch-off the LT-App in order to hide his/her location, which is a legitimate privacy requirement, then this should take precedence over PR1 and the user should be allowed to switch-off the LT-App. In this case, PR1 will not be satisfied until LT-App is switched-on again.

5.4 Discussion

In this chapter, the distillation approach was validated using qualitative data taken from a location-tracking study. The qualitative data was first structured using the coding technique from the PriF framework to isolate three PS-contexts. The PS-contexts yielded two privacy concerns relating to the Misinformation privacy threat (T5). The privacy concerns related to availability of location data (PC5.1) and location accuracy (PC5.2). These were addressed by two newly derived privacy requirements: achieve location availability (PR1) and achieve location accuracy (PR2).

The choice of PS-contexts (CW-01, CW-02 and CW-03) were not predetermined, instead we choose them because they happened to be the first ones that appeared in the interview data. Although, we found several PS-contexts which could potentially indicate the presence of other types of privacy threats and concerns, we selected only a few to limit the scope of our analysis. The analysis so far indicates that the distillation approach is robust enough to be applied to all PS-contexts that may be identified in the qualitative data. However, this is not to preempt other possible approaches which might be able to uncover additional privacy threats than the ones we have identified using our approach. What we are claiming is that distillation is one of the many approaches that could be employed to uncover privacy threats within a software system through the use of end-user reports and interview data of empirical studies.

Not all privacy requirements which were derived here are complete in themselves, in fact some of them are subproblems. Subproblems require further analysis, leading to additional privacy requirements being discovered. For example, Cr7 in PR1 can be analysed as a subproblem where one approach could be to develop a special type of battery that will automatically recharge itself through vibrations produced by users' activities (e.g. walking). A rather simplistic approach would be to warn users to recharge their mobile phones when a minimum threshold is reached. However, such battery recharge warnings already exist in current Smartphones. Perhaps the requirement here is to remind users at the opportune time to act on those warnings such as in the car or at home where they are likely to have facilities to recharge the battery.

The distillation approach also showed how privacy requirements in subproblems can be checked and addressed for inconsistencies (when composed into a composite problem) through the use of 'preferences' provided within the language constructs. This was shown with privacy requirements taken from two information-flows which were combined with their information creation and dissemination subproblems.

Before we conclude this chapter, we briefly discuss some of the threats to validity.

5.4.1 Threats to validity

(i) Thematic codes were not assessed for reliability: Similar to other inductive approaches, thematic coding in the distillation approach is a subjective process because the results are based on the software engineer's interpretation of raw data. This implies that data analysis and selection (identification of PS-contexts) may be biased. Inductive approaches prescribe the use of an assessment process where an initial coder produces a set of codes and additional analysts may be asked to apply these codes to the same raw data. The variations between the initial coder and subsequent ones are statistically measured to prove the reliability of codes (Fereday & Muir-Cochrane, 2008; Thomas, 2006). However, such assessments were not carried out for thematic coding in the distillation approach. Although, it might be possible to train a group of software engineers to use our approach and measure their level of agreements in coding raw data and identification of PS-contexts and privacy threats. Further, inter-coder agreements can be improved if software engineers can be encouraged to discuss and agree

with each other's interpretations of the raw data, similar to code cross-checking (Gibbs, 2007, p.90-105). While such inter-coder assessments can improve the confidence and reliability of our approach, this is considered to be future work.

(ii) Limitations on generalisability: The qualitative data from MFb and LT-App studies had three dimensions which were common to both, namely (a) mobility of users (b) peer-to-peer information sharing of users, and (c) privacy issues. Although distillation and more specifically the PriF framework had been designed to analyse these three dimensions, it is possible that distillation cannot cover scenarios if the dimension of privacy is not included in the data. The approach is successful only because the underlying privacy norms which produce negative behaviour patterns (NBPs) and emotions (NEIs) in users are captured in the qualitative data. In other words, distillation critically relies on NEIs and NPBs as handles within the qualitative data to analyse privacy requirements and without these markers it will be difficult to apply this approach. Therefore, distillation cannot be generalised for other types of qualitative data which do not specifically focus on privacy.

(iii) Usefulness of the privacy requirements is untested: The application of distillation shown in this chapter demonstrates how the approach can successfully help software engineers derive privacy requirements that address end-users' privacy concerns. While the derived requirements could be used to improve the design of privacy functionality of the software that was studied, it was not possible to validate this by modifying the software and testing it with the users again. Using distillation in an iterative software development project, where the effectiveness of the derived requirements can be evaluated empirically remains an area for future work.

5.5 Summary

In the previous chapter we demonstrated the workings of distillation using data from a mobile social networking (Facebook) study and here we validated the same approach using data taken from a distinctly different location-tracking study. Unlike the first dataset, distillation used the second dataset to derive privacy requirements from contrasting but multiple viewpoints. We believe this satisfies the first two evaluation

Table 5.1: Summary of outputs from using distillation on MFb and LT-App datasets (ITs=Interview transcripts, PS-Cs=Privacy-sensitive contexts, PTs=Privacy threats, PCs=Privacy concerns, PRs=Privacy requirements)

Dataset	**ITs**	**ITs used in Dist.**	**PS-Cs**	**PTs**	**PCs**	**PRs.**
MFb	4	3	3	3	4	4
LT-App	11	3	4	2	3	3

criteria - employ a systematic process and theoretically replicate privacy requirements (results) on multiple datasets.

The distillation approach, with the help of thematic coding in the PriF framework identified and isolated PS-contexts in interview data which indicated potential privacy threats experienced by users. Using these PS-contexts, privacy threats were uncovered which highlighted privacy concerns in mobile applications leading to the discovery of new privacy requirements. In this process, explicit references (links) to parts of the qualitative data were maintained to provide traceability to the source and improve transparency. This satisfies the third evaluation criterion - link to data.

The distillation approach used information-flow patterns to model a current mobile system, highlighted its privacy deficiencies through privacy problem analysis and then helped in uncovering new privacy requirements. By modelling privacy threats, concerns and privacy requirements using Privacy Arguments, we constructed arguments which informed the design of new mobile software systems. This satisfied the last criterion for evaluation.

Table 5.1 provides a quantitative summary of the outputs from distillation.

Having completed the validation of distillation, in the next chapter we briefly explore some of the tooling options that are available for use in distillation.

6

Tool options for Distillation

Human-centric tasks can be error prone and the requirements engineering process is no exception. Errors in requirements are often caused by misunderstandings and misinterpretations on the part of analysts. Any failure in minimising or addressing such errors during the early phases of software development cycle can only increase the cost of fixing it later. Further, analysts' levels of expertise may vary from one to another and only a few specialise in one or more methods. Thus, achieving consistency in the application of a particular method is directly dependent on the type of guidance and support which is made available for the use of a method. This is one of the factors that has influenced the development of software tools for analysts. The software tools are designed to play the role of an 'intelligent assistant' catering to analysts' of varying degrees of expertise, reducing the level of manual work and increasing the consistency of output (Kramer *et al.*, 1988). In addition to these benefits, there are other areas where software tools play an important role, such as requirements traceability, validation and verification.

In this chapter, tool support for the distillation approach is considered where the main requirements are based on the approach's three phases, namely:

(1) Structuring of qualitative data,

(2) Information-flow modelling, and

(3) Privacy problem analysis.

Since these three phases are based on two well known and proven methods namely - Qualitative Data Analysis (QDA) and the Problem Frames (PF) approach, a software tool tailored for distillation will need to support both of these and also provide ways to link and integrate them. The following are a set of minimum requirements this tool will need to support in order to be useful.

(i) *Coding*: users should be able to (a) define a set codes for each concept in the PriF framework and then (b) apply it on interview transcripts that semantically match each concept.

(ii) *Data extraction*: users should be able to (a) define rules for privacy threats, and (b) extract PS-contexts, i.e. qualitative data segments that satisfy a privacy threat rule.

(iii) *Problem modelling*: users should be able to model information-flows of the software system being analysed.

(iv) *Problem analysis*: represent privacy threats/concerns and privacy requirements which mitigate them. The tool should also detect any requirements inconsistencies so that they can be addressed.

(v) *Requirements traceability*: the tool should provide seamless traceability from PS-context in the data to privacy requirements.

In view of the above requirements, in this chapter we explore software tool options. Since there are already software tools that support QDA and PF methods, these tools are quite likely to be the best candidates to support the distillation approach. In the first instance, QDA software tools are explored which is followed by software tools that support the PF method.

6.1 Tools for structuring qualitative data

As described in the previous chapters, structuring qualitative data involves coding of data. Here analysts should be able to not only import transcribed data files into the software tool but also readily apply pre-defined codes from the PriF framework. The distillation approach uses extraction rules to derive privacy threats, so the software tool should provide a facility to define these extraction rules. Upon execution of these extraction rules, data associated with privacy threats should be extracted. In summary, software tools should facilitate the following to support structuring of qualitative data:

- importing of transcribed data files
- define codes for privacy facets, NEIs and NBPs.
- define rules to extract data relating to privacy threats.

There are several QDA tools[1] that can potentially support analysts in structuring of qualitative data but choosing a right one is a challenge. Some like Barry (1998) have provided guidance in selecting tools for analyses that are similar to the distillation approach. While Barry compares two software tools - Atlas.ti and Nudist, others such as Welsh (2002) have demonstrated the use of NVivo.

Among the QDA tools that are available - NVivo, MaxQDA and Atlas.ti are considered to be popular and well suited for QDA (Clare, 2012). However, these three tools are proprietary and expensive (even for educational use). For the purpose of distillation NVivo was selected mainly because it was found to be the preferred QDA tool among qualitative researchers.

Our intent is not to describe all of the functionalities provided by NVivo or to demonstrate its use in performing QDA. Rather, the focus is on those aspects of the tool which can be exploited to facilitate the data structuring phase in the distillation process. (For more details on the features/functionalities of NVivo and its use in qualitative analysis, refer to the user guide available for download [2]).

Structuring of qualitative data is mainly about coding transcripts. NVivo allows transcripts to be directly imported using import commands from the File menu. Once a transcript file is imported, its data is visible as a *Source*.

One of the main advantages of using NVivo is that the codes from the PriF framework can be pre-defined, stored and re-used by other projects and analysts. The pre-defined codes in NVivo serve as a template, making it easier for analysts to readily apply the codes on transcripts. In NVivo, each *node* represents a code or a concept. Therefore,

[1] http://www.eval.org/Resources/QDA.asp

[2] http://download.qsrinternational.com/Document/NVivo10/NVivo10-Getting-Started-Guide.pdf

all the codes relating to NEIs, NBPs and privacy facets (Table 3.3) can be defined as a hierarchy of nodes in NVivo as shown in Figure 6.1.

Using these readily available nodes (or codes) analysts can go about tagging data accordingly. One feature of NVivo (or any QDA tool) is that it provides feedback on the nodes (codes) associated with data. At any given time, analysts can view the different codes linked to the data because they are displayed in different colours (thus making it easier to distinguish between the type of nodes) as shown in the Figure 6.2

The structuring of qualitative data does not stop at coding. Coding merely helps to identify and assign sections of transcribed data belonging to a particular concept or category. After the codes are applied to data, it is relatively straightforward to extract PS-contexts depending on the combination of nodes for each privacy threat described in Table 3.8. For instance, in the previous chapter, one PS-context was extracted because it satisfied the extraction rule for 'T2 Exposure'.

In NVivo, extraction rules are defined as a combination of codes. A general syntax for extraction of PS-contexts associated with a specific threat is shown below:

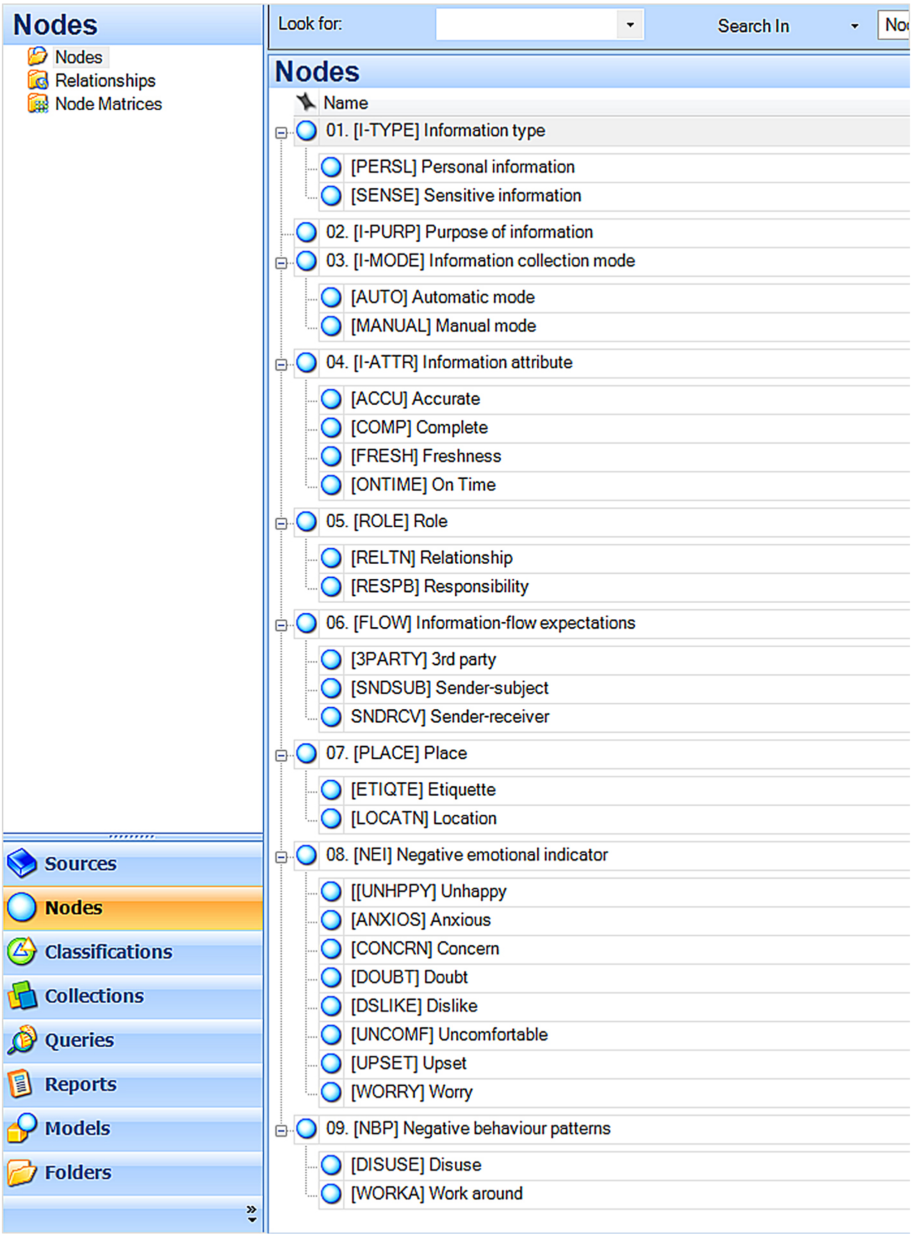
> Content Coded at $\langle NODE \rangle$ NEAR (Overlapping) content Coded at any of these nodes: { $\langle NODE \rangle$ | $\langle NODE\_LIST \rangle$ }

As noted previously, in NVivo codes for concepts are specified as 'nodes'. The above syntax states that PS-contexts will be identified by two sets of nodes (codes), the first node refers to the four privacy facets and the other node refers to NEI/NBPs.

Using the above syntax, the previously described 'T2 Exposure' threat, can be defined as a combination of codes SENSE and NEI/NPB as shown below:

T2 Exposure

```
Content Coded at [SENSE] Sensitive information
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

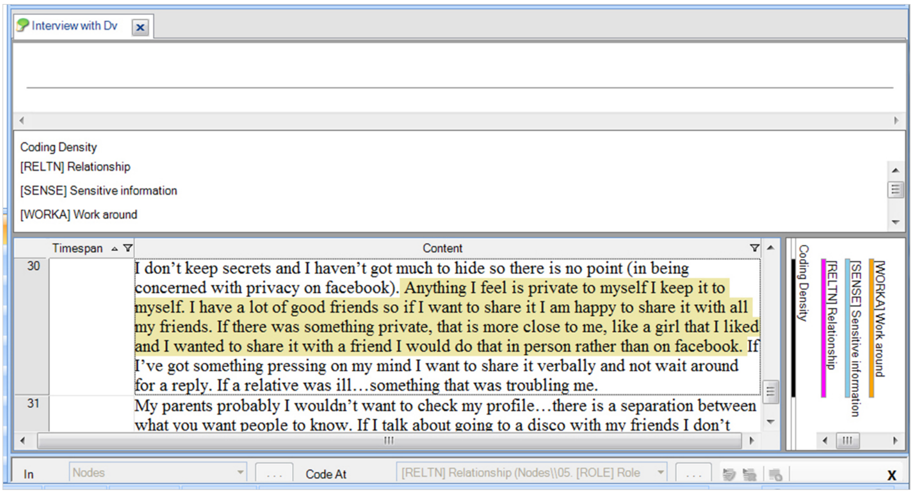


Nodes

Name

- 01. [I-TYPE] Information type
  - [PERSL] Personal information
  - [SENSE] Sensitive information
- 02. [I-PURP] Purpose of information
- 03. [I-MODE] Information collection mode
  - [AUTO] Automatic mode
  - [MANUAL] Manual mode
- 04. [I-ATTR] Information attribute
  - [ACCU] Accurate
  - [COMP] Complete
  - [FRESH] Freshness
  - [ONTIME] On Time
- 05. [ROLE] Role
  - [RELTN] Relationship
  - [RESPB] Responsibility
- 06. [FLOW] Information-flow expectations
  - [3PARTY] 3rd party
  - [SNDSUB] Sender-subject
  - SNDRCV] Sender-receiver
- 07. [PLACE] Place
  - [ETIQTE] Etiquette
  - [LOCATN] Location
- 08. [NEI] Negative emotional indicator
  - [[UNHPPY] Unhappy
  - [ANXIOS] Anxious
  - [CONCRN] Concern
  - [DOUBT] Doubt
  - [DSLIKE] Dislike
  - [UNCOMF] Uncomfortable
  - [UPSET] Upset
  - [WORRY] Worry
- 09. [NBP] Negative behaviour patterns
  - [DISUSE] Disuse
  - [WORKA] Work around

Figure 6.1: Pre-defined nodes in NVivo



	Timespan	Content
30		I don't keep secrets and I haven't got much to hide so there is no point (in being concerned with privacy on facebook). Anything I feel is private to myself I keep it to myself. I have a lot of good friends so if I want to share it I am happy to share it with all my friends. If there was something private, that is more close to me, like a girl that I liked and I wanted to share it with a friend I would do that in person rather than on facebook. If I've got something pressing on my mind I want to share it verbally and not wait around for a reply. If a relative was ill…something that was troubling me.
31		My parents probably I wouldn't want to check my profile…there is a separation between what you want people to know. If I talk about going to a disco with my friends I don't

Figure 6.2: Coding using the pre-defined codes

Thus, the ten extraction rules defined in Table 3.8 are specified as 'queries' in NVivo. These extraction rules or queries are a part of standardised template proposed within the tooling infrastructure provided for analysts using the distillation approach. The ten queries are listed below.

T1 Identification

```
Content Coded at [PERSL] Personal information
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

T2 Exposure

```
Content Coded at [SENSE] Sensitive information
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

T3 Surveillance

```
Content Coded at [AUTO] Automatic mode
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

T4 Aggregation

```
Content Coded at [I-PURP] Purpose of information
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

—T5 Misinformation

```
Content Coded at [I-ATTR)] Information attributes
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

T6 Breach of trust

```
Content Coded at [RELTN] Relationship
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

T7 Power imbalance

```
Content Coded at [RESPB] Responsibility
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

T8 Cross-contextual information flow

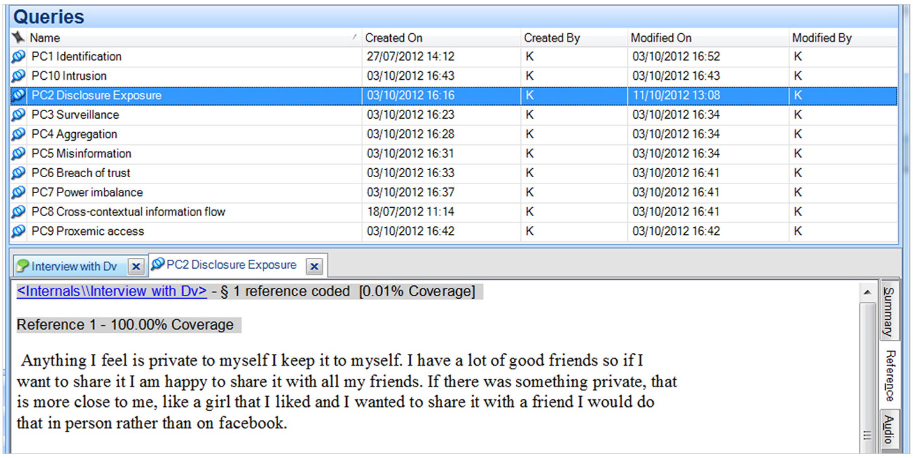
```
Content Coded at any of these nodes{
 [3PARTY] 3rd party, [SNDSUB] Sender-subject, SNDRCV] Sender-receiver}
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

T9 Proxemic access

```
Content Coded at  [LOCATN] Location
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```

T10 Intrusion

```
Content Coded at [ETIQTE] Etiquette at place
NEAR (Overlapping) content Coded at any of these nodes:{
 [DISUSE] Disuse, [WORKA] Work around, [ANXIOS] Anxious,
 [CONCRN] Concern, [DSLIKE] Dislike, [DOUBT] Doubt,
 [UPSET] Upset, [WORRY] Worry, [UNHPPY] Unhappy,
 [UNCOMF] Uncomfortable }
```



Queries

Name	Created On	Created By	Modified On	Modified By
PC1 Identification	27/07/2012 14:12	K	03/10/2012 16:52	K
PC10 Intrusion	03/10/2012 16:43	K	03/10/2012 16:43	K
PC2 Disclosure Exposure	03/10/2012 16:16	K	11/10/2012 13:08	K
PC3 Surveillance	03/10/2012 16:23	K	03/10/2012 16:34	K
PC4 Aggregation	03/10/2012 16:28	K	03/10/2012 16:34	K
PC5 Misinformation	03/10/2012 16:31	K	03/10/2012 16:34	K
PC6 Breach of trust	03/10/2012 16:33	K	03/10/2012 16:41	K
PC7 Power imbalance	03/10/2012 16:37	K	03/10/2012 16:41	K
PC8 Cross-contextual information flow	18/07/2012 11:14	K	03/10/2012 16:41	K
PC9 Proxemic access	03/10/2012 16:42	K	03/10/2012 16:42	K

<Internals\\Interview with Dv> - § 1 reference coded [0.01% Coverage]

Reference 1 - 100.00% Coverage

Anything I feel is private to myself I keep it to myself. I have a lot of good friends so if I want to share it I am happy to share it with all my friends. If there was something private, that is more close to me, like a girl that I liked and I wanted to share it with a friend I would do that in person rather than on facebook.

Figure 6.3: PS-context extracted from the qualitative data after executing the query

In NVivo, executing the `Exposure` query will result in PS-contexts associated with this threat as shown in Figure 6.3,

Once extracted, the PS-context can be analysed further as described in the distillation approach. This section showed how the facilities of nodes and queries in NVivo can be exploited by analysts to extract PS-contexts from qualitative data which, when done manually, might be considered to be cumbersome. In the next section, tool support for information flow modelling and problem analysis is explored.

6.2 Tools for modelling and analysis of privacy problems

In the distillation approach, the information-flow modelling and problem analysis phases are based on the Problem Frames (PF) method which not only uses graphical notations but also informal descriptions when modelling application and real world domains. One of the advantages of using PF is that it supports abstraction of common problems into patterns which can be reused. The distillation approach exploits this feature to provide its own problem patterns or problem frames to model information-flows in software systems (these are described in chapter 3).

PF is well respected within the RE community but because of its steep learning curve, its use in the industry is somewhat limited, which explains the lack of tool support for this method. Generally, PF notations are simple, so analysts can model with some of the popular general purpose applications such as Microsoft Word, Visio, etc. While these tools have basic facilities for vector graphics, they lack the semantics to perform automated model checking.

We found atleast two tools that claimed to specifically support PF modelling and they are: UML4PF [1] (Cote *et al.*, 2011) and OpenArgue [2] (Yu *et al.*, 2011).

UML4PF is a plugin for Eclipse [3] and it consists of a UML [4] profile with formal validation conditions expressed in OCL [5]. The OCL validation conditions make it possible to perform semantic checks on developed diagrams which could refer to only problem frame diagrams or between other types such as sequence diagrams. UML4PF represents domains as classes and interfaces as associations and it has a model generator to automatically create these interfaces based on the connections made between domain diagrams. While UML4PF offered automated model checking capabilities for PF, it did not particularly address our need to specify privacy arguments in conjunction with problem frames. Moreover, the source code for this tool is unavailable for us to modify and use.

OpenArgue which is also based on Eclipse, provides facilities to model PF diagrams and also supports incremental arguments described in propositional logic. This tool can perform syntax checking, visualizing, formalizing, and reasoning over these incremental arguments. OpenArgue integrates Decreasoner which is an off-the-shelf reasoning tool that translates propositional formulae into problems for SAT-solvers. The integrated tool supports logical deduction to check whether an argument is valid and performs model checking to obtain counterexamples to arguments. Using these, rebuttals and mitigations are generated and visualized. Although, OpenArgue was considered as a potential candidate to support privacy problem analysis in the distillation approach, it wasn't used mainly because in its current form it does not support the language extensions we require. This is considered to be future work.

[1] http://www.uml4pf.org

[2] http://sead1.open.ac.uk/openpf/

[3] http://www.eclipse.org

[4] http://www.uml.org

[5] http://www.omg.org/spec/OCL/2.0/

6.3 Summary

Outcomes of qualitative research are associated with an element of subjectivity and opaqueness. By providing tool support and automating some aspects of QDA, we can increase transparency in the process and thereby increase the level of confidence in its outputs. To this end, we showed how NVivo can be used in structuring qualitative data to extract PS-contexts and derive privacy threats, especially by creating a coding model and defining extraction rules for the ten privacy threats which were previously described in Chapter 3. Once defined, these coding model and extraction rules can be reused in other projects, thus saving time and effort when the distillation approach is applied later.

The tool support for modelling and analysis of privacy problems is still in a nascent stage. As discussed, UML4PF is not an open-source tool and it was designed with a focus on integrating PF diagrams with other design artefacts such as sequence diagrams. The one other tool we examined, called OpenArgue, looks promising as it supports PF and the privacy arguments grammar but it wasn't used because it is yet to be tailored for exploitation and use. To support the distillation approach, OpenArgue's grammar will have to be extended to include some of the new constructs we have proposed. Since, code changes and redevelopment will require time and resources, this can be considered to be future work.

7

Conclusion & future work

Eliciting mobile privacy requirements is challenging, largely due to the fact that mobile privacy issues are so dependent on the physical and socio-cultural context of users. This means that only data that captures the nuances of these contextual factors and variations can adequately inform the development of privacy requirements for privacy-aware mobile applications. The case studies in this thesis used data from an empirical field study. The advantage of such data is that it provides the necessary contextual information reflecting the reactions of real users to real experiences. However, its richness means that this data is not amenable to the traditional software engineering processes of requirements elicitation and specification. Moreover, qualitative data in the form of interview transcripts contain answers to questions which are often related to exploring end-users' experiences with regards to their actions and interactions in the operating context. While end-users' responses provide insights into their rationale and emotional state, they are often associated with other objective properties and dimensions such as information sensitivity, users' roles and relationships with other agents etc. which will have to be systematically analysed. In addition, the users' privacy requirements are made manifest indirectly as privacy wishes, expectations and needs in response to perceived privacy threats. Unlike researchers who can employ several QDA techniques (Braun & Clarke, 2006; Corbin & Strauss, 2008), software engineers have very little to choose from when it comes to analysing qualitative data. The distillation approach proposed here allows software engineers to take advantage of the richness of qualitative data by systematically refining it into a form that enables it to be used for the design of mobile applications that meet end-users' privacy concerns and needs. Our distillation process used PriF, a novel privacy facets framework to structure qualitative data and to derive privacy concerns. It then used problem patterns to identify information and its flow within a software system. Finally, it combined the outputs of these two steps to derive precise privacy requirements. We believe the distillation approach and the PriF framework, as demonstrated in the validation, were successful in achieving the research objectives and our work makes the following contributions.

7.1 Contributions

In this section we briefly summarise our research contributions and their potential impact.

Privacy threats for mobile applications
Analysis of privacy threats is made difficult by the fact that there are several definitions for privacy. Any holistic approach to analyses will require a comprehensive catalogue that covers as many privacy threats as possible for the software system currently being analysed. Unlike some who have derived privacy threats by negating privacy principles (Kalloniatis *et al.*, 2007) or security properties (Deng *et al.*, 2011), the PriF framework provides a taxonomy of privacy threats which were based on well known works such as Nissenbaum (2010) and Solove (2008). The latter produced a list of privacy threats based on real-life court cases and litigations. Morever, the privacy threats adopted in the PriF framework and distillation were specially tailored for mobile applications.

The PriF framework
Qualitative data is always difficult to work with because it is not only fuzzy (imprecise) but also contains noise that is not useful for analyses. Unlike ethnographers and social scientists, software engineers are not specialised in analysing such data. In other words, software engineers may find it hard to structure and analyse qualitative data and derive privacy requirements from it. The PriF framework was designed specifically with this requirement from software engineers in mind. The PriF framework provided a set of cues in the form of *behaviour patterns* and *emotional indicators* of mobile users to identify privacy-sensitive contexts (PS-context). Using a set of *Facet questions*,

analysts are able to elicit and explore the PS-contexts with respect to four facets - *information, actors, information-flows* and *places*. The framework also provided guidance on identifying and coding of *privacy determinant* properties which contribute to privacy threats in the PS-contexts. By running pre-defined *extraction rules*, privacy threats were identified and associated *privacy concerns* modelled. The framework also provided *problem patterns for information-flows* to analyse privacy concerns and derive new *privacy requirements*. To provide traceability and enable automated model checking, privacy requirements were expressed as *privacy arguments* by extending the previous work done by Tun *et al.* (2012).

The distillation approach
The three phases in the distillation approach provided a clear separation of concerns. In the first phase - *qualitative data was structured* to identify, extract and specify privacy concerns. Privacy concerns point to how software systems failed to address privacy threats of end-users. In the second phase, *information-flows were modelled* from initial requirements of the software system being investigated. In the third and final phase, as part of *privacy problem analysis*, privacy concerns were mapped to information-flow models to identify any gaps, leading to the discovery of new privacy requirements. In our bottom-up approach, privacy threats and privacy concerns were based on real-life experiences of end-users. The resulting privacy requirements are therefore very specific to the software system being investigated and reported on. This is unlike top-down approaches such as PriS (Kalloniatis *et al.*, 2007) and LINNDUN (Deng *et al.*, 2011) where high-level privacy goals and requirements are based on privacy principles which are rather vague and open to subjective interpretations. This may lead to faulty implementations where the privacy threats may not be fully identified.

Privacy arguments
In the distillation approach, the main challenge was to express concepts such as privacy threats, privacy concerns and privacy requirements in a coherent manner. Further, these concepts had to be integrated with information-flow models with their semi-formal domain descriptions found in the Problem Frames method. Additionally, there was a need for these models to provide traceability by linking them to qualitative data. These multiple challenges were addressed in the distillation approach by adopting and extending the Privacy Arguments language (Tun *et al.*, 2012). Here privacy norms

were described as claims and privacy concerns as counter claims. Further, privacy requirements were described as counter claims which mitigated privacy concerns and supported the satisfaction of privacy norms. The constructs in privacy arguments were extended to include a specialised ground called *'evidence'* which assisted in referencing evidence from qualitative data. In addition to providing links and traceability between the requirement models and qualitative data, the integrated approach of distillation aimed to preserve the simplicity of problem analysis by using only light-weight notations. This makes it relatively easier for both technical and non-technical stakeholders to discuss and agree upon.

Tool support
In order to facilitate automation of tasks in distillation, we explored the possibility of reusing and adapting existing off-the-shelf tools for: (i) structuring of qualitative data (ii) requirements modelling in problem frames, and (iii) expressing the resulting requirements in the privacy arguments language. Using facilities within NVivo, all codes relating to the concepts in the PriF framework were converted to 'nodes' and extraction rules for privacy threats were defined as 'queries'. The interview transcripts, in form of data files, were imported into NVivo and were later coded for their behaviour patterns/emotional indicators, facet attributes and privacy determinants. When the pre-defined queries (extraction rules) were executed, the privacy-sensitive contexts with their associated privacy threats were automatically identified. In the second part, OpenArgue (Yu *et al.*, 2011), which supports privacy arguments was considered for use in distillation. Here, our contribution only has been to state the requirements for a generic software tool which can support the distillation approach and also describe how existing tools can be exploited to achieve some level of automation.

7.2 Future work

In this section, we briefly state the options to potentially extend our work.

(i) Making the distillation approach extensible
The distillation approach was first described using qualitative data from an empirical study involving mobile Facebook users. Further the approach was validated using another dataset taken from an empirical study of location tracking users. In both

of these, the users' mobility and privacy were key dimensions which the distillation approach exploited to derive privacy requirements. It would be useful to explore the extensibility of the distillation approach by first applying it on qualitative data that does not specifically focus on mobility. This could potentially lead to the derivation of privacy requirements for desktop systems. Secondly, instead of using qualitative data from user studies, it might be worth exploring other types of end-user reports (e.g. from other sources such as online complaints and feedback) as input for the distillation approach. These two experiments could highlight the ways in which the distillation approach can be made more extensible and generalisable, thus broadening its usefulness to the software engineering community.

(ii) Automated model checking and reasoning for privacy requirements
Managing consistency of privacy requirements is a major concern, especially when requirements evolve. Further research is required to detect and resolve conflicting privacy requirements which are modelled and represented using our approach. This will require some form of semantic encoding to be designed and implemented to support reasoning in an automated model checker.

(iii) Monitoring and adaptation of privacy requirements
Privacy requirements are subject to constant change which must be monitored and adapted in software systems that implement them. For example, when a privacy determinant such as location (or roles and relationships of actors) changes, it has an impact on the privacy norms currently supported by a software system. It is therefore critical that privacy determinants under each privacy facet be actively monitored and suitable adaptations performed to maintain the privacy of end-users. Further investigations are required to, firstly, determine which of these privacy determinants are to be monitored and how to monitor them. Secondly, when these privacy determinants change, what are their impact on the current system, and what adaptations are required to maintain the privacy of end-users.

(iv) Encoding arguments as machine-readable privacy policies
The PriF framework provided formal language constructs to encode privacy norms which software systems are expected to support. If these formal descriptions were to be translated into machine-readable privacy policies, they could be used directly in

software systems to enforce privacy protection of end-users, thus making the outputs of distillation even more useful.

7.3 Conclusion

Previously software engineers had difficulty in eliciting privacy requirements for mobile applications because such requirements were influenced by the frequently changing operating context. On the other hand, qualitative methods which collect rich data on the users' experience were not compatible with existing software engineering approaches. Thus, the aim of this research was to develop a new approach to enable software engineers to extract mobile privacy requirements from end-user reports which were in the form of qualitative data. To this end, a novel distillation approach was introduced which exploited the facilities within the PriF framework to firstly, structure qualitative data, second, to model information-flows and thirdly to perform privacy problem analysis to extract and refine privacy requirements. In the first step of the distillation approach, privacy sensitive contexts were identified and isolated using qualitative data, and then additional contextual details were extracted through the use of privacy facet questions. Further, instances of privacy norms and concerns were extracted and represented as privacy arguments. The second step of the distillation approach, which exploits the information-flow problem patterns of the PriF framework, helped model the information-flows associated with the previously identified privacy norms. In the last step, when these information-flow models were analysed in conjunction with privacy concerns, they produced new privacy requirements. These requirements were represented using the privacy arguments language.

We initially used a scenario to exemplify the PriF framework, and then used qualitative data taken from a Mobile Facebook study to show the workings of the distillation approach. We then validated the distillation approach by using a fresh dataset from another empirical study involving location-tracking of users and successfully derived privacy requirements for a mobile location-tracking system. Finally, we showed how some distillation tasks can be automated using existing off-the-shelf tools.

We believe there are several benefits of using the distillation approach and the PriF framework. First, unlike other top-down approaches, our bottom-up approach provides software engineers with a realistic chance of addressing application specific privacy threats and concerns as and when they happen and are reported by end-users. The distillation approach equips software engineers to spot privacy norms, threats and violations in the qualitative data which was previously thought to be difficult. Further, our approach provided facilities to capture the privacy norms found in the real-world and formalise them in a way they can be understood and reasoned about by software engineers and other stakeholders.

In conclusion, we claim our novel distillation approach along with the PriF framework were successful in achieving their research objectives and can benefit the software engineering community in deriving privacy requirements for mobile applications. We believe our work can also benefit the research community in understanding privacy and privacy threats that apply to mobile computing systems. We have also shown how techniques from traditional qualitative research methods can be successfully ported to enrich and extend the capabilities of existing software engineering approaches.

Appendix A

Mobile Facebook study

In PRiMMA [1], an EPSRC funded research project, we had conducted three user studies that specifically focused on privacy of mobile phone users. The first study was based on the social networking application - Facebook, the second study focused on location-tracking within families. Both these empirical studies investigated a small group of mobile users by employing methods developed by us. Here, we describe the first study on 'Mobile Facebook' and produce partial data obtained through participant interviews.

A.1 Participants and data

To understand how people really feel about privacy, it was critical to understand how their networking practices integrated with their daily life practices and routines. In this study, our aim was to observe how Facebook activities integrated with people's other daily practices, mainly to identify behavioral patterns relevant to any emerging privacy concerns produced by a familiar technology.

In this study, we monitored social networking activities of 6 participants using Facebook. The participants were aged between 21 to 28; they were either studying or working in two universities in the UK. Not only were all of participants experienced and enthusiastic users of Facebook but they also used the social networking application

[1] Privacy Rights Management in Mobile Applications - (http://primma.open.ac.uk)

on their mobile phones. In order to avoid inconsistencies caused by different functionalities and user interfaces of different phones, we selected participants who owned the same type of handset, in this case it was Apple's iPhone.

Approximately a week into the study, the participants were interviewed regarding three aspects:

(i) *The persons habits and routines in daily life:* these included questions about, for instance, household arrangements, work patterns, socialization patterns, etc.; they aimed to develop a rough profile of the participant;

(ii) *Facebook activity for which we had received feedback:* these followed the chronological order of the participants Facebook actions and investigated further the answer they had provided to the experience sampling questions, including the choice of memory phrase; these questions aimed to probe the participants about their Facebook actions and gather further details about the context in which these had been carried out;

(iii) *Use of Facebook and any privacy-related issue:* after going through their specific actions, these questions aimed to explore any issues that had not yet been touched on, expand on issues that had been touched on, and find out about the participants general views on Facebook and privacy.

Although there was a structure to the interview, there was flexibility in order to allow the participants to discuss any emerging issues. The interviewer shared and discussed with the participant any interpretation of their answers. The outcome of this study is available in Mancini *et al.* (2009).

A.2 Technical implementation

In this study, we used a novel approach to elicit the participants' feedback on the different Facebook actions; it combined two methods:

(i) *Experience sampling:* a method which recorded 'structural' information about participant's activities (Consolvo & Walker, 2003).

(ii) *Deferred contextual interviews:* an in-depth semi-structured interview whose purpose was to probe the participant using the experience sampling data as pointers (the answers to the questions) and memory triggers (using a memory phrase).

Whenever participants performed a Facebook action, a short questionnaire (6 objective type questions) was sent to them on their mobile phone which captured the context of the action. For example, the questionnaire would elicit the user's mood, their current location, other people who may be co-located etc. In addition to this short questionnaire, the participants were encouraged to insert a *memory phrase* which could help them recall their experience during the deferred contextual interviews.

To deliver the questions to the participants and collect their answers, we built a User Feedback System (UFS), whose aim was to collect data in an unobtrusive and practical way, via the participants mobile phones. The system consisted of several modules, one of which collected Facebook status updates via the RSS feed at a sampling rate of 10 per hour. Another module, deployed on the same server, detected changes in the status of the Facebook user and sent an SMS message to the participant containing a URL to a web form containing a short questionnaire using their iPhone. If the status update was carried out through a desktop PC, the SMS would have been ignored by participants and no feedback would have been given.

Although feedback was collected for several Facebook actions such as status updates, photo uploads, comments etc., this research focuses mainly on the updating of 'status message'. Thus, there are partial transcripts of the deferred interviews that took place in the study and these are made available for our analysis here.

A.3 Partial transcripts from the interview with Dr

¶ **[1]** I usually don't share my private life on Facebook, or at least what I think is my private life (when I was in a relationship my relationship status on facebook didn't change).

¶ **[2]** *Project officer: intellectual property. Live alone and drive to work, even though live only one mile away. Like triathlon and photography. Goes out only every two weeks.*

¶ **[3]** Started very excited using almost every day within a close group of friends at work, until there was a social accident: we went outside we did some pictures and put them on Facebook and some of the people inside the close group asked why were we not invited to the dinner, so I started to be concerned about security and the fact that you need to share everything with everyone (of your friends)...I got concerned and stop to use it and put the security filters really high so no one was able to find me. The concern is that you have your public life exposed and that you don't have control...the picture was uploaded by another person and she put a tag so everyone had access to the full album so everyone could see what we had done...so I had done something over which I didn't have control.

¶ **[4]** Also at the time there was a rumour that facebook was selling personal details to other companies so I got concerned about that as well...I don't know if it's true or not. I stop uploading everything for one years. One year ago I changed attitude, I don't know why, and I put the filters down and people were able to see me again and I got interested in it again...that (if there was an episode that triggered that) I can't remember, but a year ago facebook exploded in Italy as well...it exploded among my friends so I was quite excited about finding and being found by my friends...yes, I started to use it pretty much every day...and also my sister joined last year and told me, yes, facebook...you hadn't told me that you had it... a bit yes (she resented that I hadn't told her I was on it) because she was excited as well. Now it's almost every day. Usually update my status every day, maybe I'm checking it two or three times a day.

¶ **[5]** Yesterday I found my sister so we started to chat so it was about half an hour...usually is less. Usually it is on the desktop...sometimes on the mobile...usually 90% on the desktop and 10% on the mobile...easier on the desktop for the interface and the speed of connection. 98 friends, I met them all at some point of my life...friendship is a lot for me...I tend to consider friends a small number of them...seven eight...the people I can count on in my life are on facebook are among the people of facebook. Yes (I wouldn't mind my parents being on my facebook) because there is nothing that is really private.

¶ **[6]** *First action (status message update):* Usually I tend to be specific on a certain topic which I'm ok that people know...I'm not happy people knowing about my relationship...or my personal problems, my working problems...I usually don't put these... Was alone waiting to go to the gym...was comfortable...

¶ [7] *Second action:* That was a joke for you guys...was watching tv...on my own...nobody saw me.

¶ [8] *Third action (status message update):* Watching tv ...on my own

¶ [9] *Fourth action (status message update):* Waiting to go to Tesco...was alone...nobody saw me.

¶ [10] *Fifth action (status message update):* I got three comments with that...interesting that I want my friends to know and then I get comments... There is a way if I want to target a specific person on specific things...if I want to have some support from particular people I tend to use that particular thing that relate to that person...that person reacts. If you go back to my status with Karis...I write down "Dario is missing katan...it's basically a war game...and am very close to Karis and her husband Chris and nobody else knows about that and Karis replies "I am missing that too and if you come in September we can play".

¶ [11] *Sixth action (status message update):* In the garden... (garden is overlooked) I don't care...

¶ [12] *Seventh action(status message update):* (same thing happened: he says he wanted his friends to know and he got comments back) I am surprised now because I am noticing that when I want others to know I get a reaction whereas when I say that I am bored I get no reaction...it's like I'm doing but I am unconscious that I am doing it. Was in the gym in the Costa caf.

¶ [13] *General questions:* I don't mind let people see me updating my status on facebook because I quite like technology usually and...the iphone went out November 2008 and I got that iphone nov 2008...but if people see the content I think it's not their business, I get a bit annoyed and frustrated and irritated...it my stuff, I don't want to share my stuff with you that I don't know...I am more concerned about the monetary value of the telephone...and if somebody would like to stole it from me...I am more concerned about showing the technology in public because that can attract attention from people who want to stole it...steal or try to steal...that's going to be a difficult situation...

¶ [14] I was on the train last year...from Birmingham to London...we were considering the train or the bus, definitely we went with the train because I was able to plug my iphone. I didn't no, because most of the time with iphone you can do at least two things at the same time, you can listen to music and you can do other stuff, so usually I listen to music and try to isolate myself from the world. I think this question is pretty much the same as text, like a normal texting in public...I don't have any problem, it's a way of communication...yes, I am worried about the monetary value but...in first I would be irritated...possibly you are too close to me, so you are invading my space...with the text is "we are going to meet in five minutes, I'm going to be late, sorry", with facebook is more general stuff: what do you do, what you are up to...I think if you follow my story if you follow my page, yes you can get a lot of information about people, but it is only one spot on the bus, no...yes (it bothers me in the same way, for the same reasons) because it invades my space...so I get annoyed.

¶ [15] If I am out with friends I don't take my phone out, I don't do facebook...yes, ok, if I am with my sister I keep to read emails, but no I don't use facebook and I tend not to use the mobile...because I am busy with other stuff, talking with them, socialising...facebook tends to fill the gaps...if I am with a person I concentrate with that person...also because the iphone is quite fancy I don't like showing it, something I have bought for me not showing I can afford it...

¶ [16] Not that I consider private: no relationships, no problems, no work...not problems I have with my boss, with my friends...just general things which I don't think are much relevant to other people for security...yesterday I had a meeting with my boss and the boss is really happy with my work and says yes we are thinking about extending your contract, I was very happy but this is something I would never share with facebook, so I called my family and said, hmm, there is this opportunity that has just come I am really excited, but I am not going to share that (on facebook)...when the contract is finalised maybe...but this stuff I would never put on facebook because my stuff, yes, it's mine. Yes (I share it with some people) but with different way of communication, like with email or phone...because you have control on it. There are two other things that I am comfortable sharing: what I am doing now, it's ok; the triathlon and the photography is fine, while other definitely are not. I don't have a filter that I can choose what yes and what not.

¶ [17] Do the checking on the laptop and the status update from the mobile. I look at all the status update form all the friends, all the pictures from all the friends, this is what I usually look at...and then there is my status update, which is form the mobile. It's for the interface, because with the mobile it's easier to filter just the status update or just the photos from the PC.

A.4 Partial transcripts from interview with E

¶ [18] General questions

¶ [19] On the laptop from home, on the mobile from the uni. Check updates on what's going on with people I know. Over 400 friends. Met and spoken to them all but many only acquaintances. 50 are good friends.

¶ [20] *First action:* I wanted my friends to know what I was thinking. Trying to organise another bbq, because we had a very successful one. In the kitchen in my house with friends. I let my friends see what I posted on facebook.

¶ [21] *Second action:* Was in boring lecture. I tend to check it a lot when I am in lectures. We were sat together, they were close (friends). They probably didn't see it. You could see that they were in action (too). It a sort of knowledge (that everyone does it). I didn't mind, if they saw it (me doing facebook) they wouldn't mind. Probably, there is a chance definitely 9the the lecturer would see me. No (I don't mind) if the lecture was more interesting I would pay attention.

¶ [22] *Third action:* It was the rugby grand slam, three games in a row. I was with my housemates and other friends come over. I got the results I was waiting for...it was like a notification of my feelings really...sort of sharing that I was happy. Couldn't be bother to go upstairs to my laptop so I did it on the mobile, which is always with me. I don't think they saw my facebook action, I didn't share with them....it was a self expression. I was just sort of putting it out there. I would have felt comfortable even if they had seen me, they would have seen it later anyway.

¶ **[23]** *Fourth action:* We were trying to put things on my friend while he was sleeping. It was quite funny because he didn't realise until the next day. Yes, (I did that communally with one of my other house mates). I was comfortable because it was quite funny.

¶ **[24]** *Fifth action:* The computer of my friend broke. I came up here for a meeting, a group meeting with people I don't really know, colleagues...and we had a free moment, so I just wrote on his wall...winding him up a bit...it was more just a joke...we had a quick break during this meeting so I just quickly...yes (I waited for the break, because I don't really know the people in the group...no (they didn't see me) I sort of like, I was a bit sort of like, I kept it a bit more personal, sort of like held it close to my body, yes sort of like, sort of really not out there, they wouldn't really understand it...(if they had seen it) I thought they might have thought I was a bit harsh... also I would have raised quite a few questions, I didn't want to answer because I don't know them, so keep it a bit more personal...I want a little bit more, yes, closer, keeping it in on hand. I felt comfortable because they didn't see it. It was more an injoke...sort of a bit more...injoke. (if I had seen it I would have felt) probably more uncomfortable because I don't know them and also they don't know my mate, they don't know the back story...it probably come across differently if you don't know the back story...

¶ **[25]** *General questions:* It's been very work heavy this week...so I didn't have a chance to do much...I would expect it to double or even triple on holiday...I do a lot of checking on the bus...more through boredom...when I check as I am working I use my laptop, because it is already on...I actually prefer using the phone...you can do a lot more on the laptop but I prefer the layout of the mobile...it's got more like the keys going on...the laptop displays more but it's almost a little too much...

¶ **[26]** I don't really tend to do anything on the bus because the journey is not long. Things like buses and trains I don't feel so comfortable, because, I don't know, people, people I don't know, lots of people I don't know...so people, if they for example read some of the posts I have done...because they don't know the people that they are aimed at or the backstory...they'd probably come across quite differently and they would not understand them, it would look a little weird...yes (they would get) the wrong sort of, almost the wrong first impression...but yes, (I do check on the bus)...but good for checking not so good for actions.

¶ **[27]** I am quite happy to do it inmost situations, like at least to check...aside places where I wouldn't do it...not because I wouldn't feel comfortable doing it...would be like at family events, family meals and things like that, cause it's more rude, so it's more the people...yes, I wouldn't want to be rude...yes (with my friends) if something comes through I would be happy to take it...friends would be fine, it all very open...more because of the context...when you are with friends is more relaxed, whereas with family there is more of an inbuilt strictness...yes you have to appear to enjoy yourself...yes I would say that (the don't understand the technology)...with my family, they don't understand the technology and don't see the point, they don't see the need...it's also especially with some members of the family, like the older one, like a generation gap.

¶ **[28]** For me I would say I use facebook more like a social constructor, I use it to find out events and see what people want to do...my family would say why can't you use the phone or a letter, but facebook allows me to do that in 20 seconds whereas a letter would take two days. My close family, my parents and grand parents. Everything is a bit more forced, like the conversation. Ever since being at university, facebook is has become a very easy way of displaying events, what's going one, what people are thinking...also it allows you a lot of forward planning because it has an inbuilt calendar, it reminds you things, oh, I'm going out today...

A.5 Partial transcripts from interview with Dv

¶ **[29]** Transcription

¶ **[30]** (For the past week little mobile Facebook activity because there was so much work to do for the course assignment that I was) ...completely unavailable for anything, including eating....if you were even seen to be near your phone by another member of your group...you would have probably been frowned at...not even taking personal calls.

¶ **[31]** I am far more inclined to use my laptop just because it's so much easier. If I happen to be near my computer the mobile takes a back seat...for practical reasons. On the computer you can see more contextually and you have a keyboard to type. You can do multiple things at ones (have the news open, as well as facebook and your work file).

¶ **[32]** *First action:* This time last week I was about to start my final year project and I wasn't looking forward to it. I thought that a lot of my friends on facebook were in the same position as me, they were all about to start and I thought that was a feeling that they were also probably feeling so I just wanted to share. I was kind of relaxed looking at what I was going to do the following week and probably was on my bed. I was alone in my bedroom watching tv during an advert break.

¶ **[33]** *Second action:* To let my friends know that I won't have my mobile...so they would be aware is I didn't answer the phone. Also as a kind of a joke because my friends don't use the mobile quite so much and think that people with an iphones...are always using them. A bit of a joke at my own expenses. It was to give a bit of information and make a joke. A news update because you want people to know. In the living room with friends, I can't remember what I was doing. We were probably sitting around...I wasn't feeling we were socialising right then. They were also my facebook friends and was communicating with them also in real life at the same time.

¶ **[34]** *Third action:* Morning after finishing my FYI again I just wanted to share...because I was there with my mobile I did it on the mobile. I wanted my friends to know that I was relaxing, it was a shared experience... (they would have been finished work as well). I wanted them to leave me alone. My status was due for an update...

¶ **[35]** *Fourth action:* I was in a pub socialising with B. other did see my action and I didn't even think about it. It doesn't occur to me. I'm quite happy to use my phone wherever whenever. Even in places where I know people don't enjoy other people using mobile phones I don't care I have total disrespect for the rules. Recently if I see something I want to share I use twitter because it's very easy to update photos so I don't do it through facebook anymore. When there is an experience I want to share I'd post a photo rather than writing about it (on facebook). I do it because of the novelty, I have never been able to do this with a mobile before.

¶ **[36]** *Fifth action:* I didn't actually post this one, this is one of my friends. I don't know how they got it (the phone), probably when I looked away. Just before we went out. It was one of the girls that did it. That didn't bother me at all. It's kind of an ongoing joke, we update (each other's) statuses. You can do it to anybody but probably wouldn't say something like this. If someone left their laptop open. I wouldn't do it to someone I didn't know. I wouldn't let some of my friends have my phone because they are clumsy (and also because the might post something I don't want). I don't know many people who don't have facebook, partly because the uni of bath uses facebook to post events.

¶ **[37]** *Sixth action:* I thought I would let the world know of my disapproval (of the weather, which was bad just as I finished my assignment). I was thinking about something in the future (playing a game outdoors, which had to be cancelled).

¶ **[38]** *General questions*

¶ **[39]** I feel like I am in many relationships because the relationships with my friends are...

¶ **[40]** I don't add people who I don't actually know (200 roughly), some people add friends randomly

¶ **[41]** Yes (I tend to use my mobile more to check than to do actions). I use my mobile to check facebook when I am on the bus on my own. I may share something I have seen on facebook with my friends, but it doesn't happen often because if my friends are there I am interacting with them. It takes a while so if your friends are there you interact with them. (concern about being seen by others) not necessarily because I don't want them to see what I am sharing, but because it is not relevant to them. If you are telling a friend about what you are going to do in the weekend...you don't necessarily want the world to know what you do at the weekend. I wouldn't say I feel uncomfortable with people seeing what I am reading or writing, but at that point you are more aware of people who are watching, at that point, if there was anybody, but equally you don't do that on your phone because reading and replying to what is essentially a long email isn't possible. But equally if you are using your mobile you don't have much problem with people or feel uncomfortable because people cannot actually see what you are doing unless they were watching over your shoulders. At times you may do it during lectures (don't mind being seen). At times you feel that the lecturer hasn't put much effort and you think it's dull or you didn't need to turn up for that so you zone out (other people do that as well).

¶ **[42]** People would feel that you are more antisocial if you do facebook on your laptop (when you are with people) whereas with your mobile is more accepted, you can pose to check a text, etc. whereas my parent's generation would find it rude if I stop a conversation to take a call.

¶ **[43]** I don't keep secrets and I haven't got much to hide so there is no point (in being concerned with privacy on facebook). Anything I feel is private to myself I keep it to myself. I have a lot of good friends so if I want to share it I am happy to share it with all my friends. If there was something private, that is more close to me, like a girl that I liked and I wanted to share it with a friend I would do that in person rather than on facebook. If I've got something pressing on my mind I want to share it verbally and not wait around for a reply. If a relative was ill...something that was troubling me.

¶ **[44]** My parents probably I wouldn't want to check my profile...there is a separation between what you want people to know. If I talk about going to a disco with my friends I don't necessarily want my parents to know, not because here is something wrong with it, but because it's not important to them. Equally my parents wouldn't have a hope of getting around either facebook of twitters, so I've got nothing to worry about.

A.6 Partial transcripts from the interview with R

¶ **[45]** Account for 3 and years. Checking on average 10 times a day, of which 4 on mobile. On laptop during work, on iphone otherwise. First time in the day 10 minutes, after that only a couple of minutes. 80% checking and only 20% actions. Around 470 friends, of which 300 I say hello to if I see them. People I accept I don't necessarily say hello. I tend to accept people who ask me to be friends. Not every time, but I find it kind of rude if I don't, but I'll only ask them to be a friend if I have some kind of motivation to do it.

¶ **[46]** *First action:* A friend asked me if I wanted to go to a summer festival, I responded through facebook...like I was responding an email. That would have been actually in a lecture...because we had a lecture (at that time). It was a two hours presentation that dragged on a bit so by that time I guess I was just you know I was looking down like that checking my facebook...yes (I do that at times), if I get bored. With colleague...oh yes, all my classmates...they are on facebook. People sitting next to me, I don't think they did (see my action) because they would be concentrating on the presentation. I was feeling comfortable because what I was talking about wasn't really you know that private, I would have felt comfortable anyway.

¶ **[47]** *Second action:* We were watching a film downstairs. Someone added a new photo and I commented on it...I think it was my friends drunk. I was at home enjoying leisure time. They didn't see me...we were watching tv. I was at a different angle.

¶ **[48]** *Third action:* I went onto the festival site to see if I was interested and there was this game on the site. If you get a top score over a million you score. I only got...so I couldn't understand how they do it. I guess I was frustrated about the game so I wanted some of my friends to get interested in the site and come to the festival...so I was promoting the festival. I think indirectly I was promoting it, but this was just to show there is a good game on the site I don't seem to make a good score. I was in the library...didn't know anybody...the only reason I would have to be concerned I if I was writing something private but I wouldn't write something private on my status...in my regards anything that go on my status is public so I shouldn't have a reason to be concerned about somebody behind me seeing what I am writing when 400 of my friends are going to see it directly...if it was something private I would write it in an email...quite often if there is something going on and only a small number of people involved we'll do it in a 5/6 way emails...generally anything that is not publicly open has to go private.

¶ **[49]** *Fourth action:* (struggles to remember the details without the memory phrase). I would have just been after a check...it might have been...I check it so many times a day because there are so many updates and things change so quickly.

¶ **[50]** *Fifth action:* My flatmate is irish and I am welsh so...I thought we were having a good game and wanted mainly to show my support for wales for the team. I thought maybe my friends are watching I thought I would show my support and also maybe increase the rivalty with my flat mate. Yes I was at home, I nipped up to my room (to do that) but (were watching tv) downstairs. People were going back and forth (up and down). I don't know why that was the memory phrase...oh...may be we were watching Notting Hill before so that was the connection.

¶ **[51]** *General questions*

¶ **[52]** Yes (I do facebook on the move) if I am on the train, but on the bus here just check my email and the signal is not always guaranteed...it depends on the time of the journey as well...but also it does restrict me because at times I tried writing a message on my phone and the signal wouldn't work...I think if there is a stranger sitting next to me on the bus and he can see I would feel uncomfortable I wouldn't want him to see because regardless of what I'm posting you know my friends messages they are still private and I wouldn't want a stranger to see them...I don't know I just think is a bit intrusive, it's like someone looking over your shoulders at a book you are reading...it's not what they might see (because all my friends can see directly what I write anyway it might as well be cast to everyone) (it's the intrusion of space)...but I think like in library setting there is an unwritten rule that you don't look at anyone else's facebook...I don't know is like a kind of etiquette that you don't look at anyone's facebook, but on the bus...it's polite to look into anyone else's facebook...(I rely on the to feel less concern)...yes on the bus you don't know who they are is less of an uncontrolled setting as opposed to everyone knowing the etiquette in the library...the majority of people in the library knows that it's impolite to look at anyone facebook whereas the public on a bus they might never have used facebook before they may not know the etiquette...the library is more controlled.

¶ **[53]** I don't like my girlfriend to look on my facebook...she is among my friends...if she is not friend with some of my friends she cannot see what I have written on their wall...if she is watching me I could have said my girlfriend and I had an argument...so I have to be careful with that.

¶ **[54]** I don't think I want my mom and dad to see what I have done on facebook. Also my family friends...I have had a few requests from family friends, I don't like to accept them because that pus more of a restriction on what I can actually say on my facebook because that exposes me eve more to my family. If I did let my family see it it makes my facebook activities more and more public so it restricts what I can write...(if they don't see it) it gives me a wider scope as to what I can write...

¶ **[55]** Facebook is good for...two things really...the first is the student version of email so you don't have to have a direct phone call...wherever you are you can get a facebook message and it gives freedom to respond in your own time and think about what you are going to say, whereas on the phone you have to think on the spot...secondly I think that it lets you know and keeps you in touch a lot closer with friends you wouldn't know what's going on I have got a lot of friends I may not have the chance to speak to that often but still gives me the opportunity to know what's going on in their lives through status update...so I think status update (is the big plus of facebook) without that facebook wouldn't be the same.

¶ [56] My brother is of a similar age to me so things are a bit more public with him...my sister is a bit younger so she falls in the same category as my parents.

Appendix B

Location Tracking Study

A location-tracking study was the second of the three studies carried out in the context of a broader EPSRC research project PRiMMA [1]. The aim of the study was to investigate how privacy was managed within highly trusted groups such as families, and the study is briefly described below.

B.1 Participants and data

The study comprised two groups of family members, one with 7 and the other with 5 participants. The first group (F1) was a family from [another country] who had relocated to UK two years ago after suffering an abduction and an armed robbery. They consisted of a married couple in their fifties, their three daughters and the partners of the two older daughters. The husband, a lawyer who worked locally, and his spouse, a housewife, lived with their youngest daughter, who was in her late teens and attended high school. The middle daughter, in her early twenties, lived and attended university. Her boyfriend of six months, in his late twenties, lived in a city and worked in a restaurant. The eldest daughter, in her mid twenties, worked in a city and lived with her long-term partner, who was in his early thirties. In [another country], where crime levels are high and abductions frequent, it is common for families to use tracking

[1] Privacy Rights Management in Mobile Applications - http://primma.open.ac.uk

technology for security purposes. This family had also done so but none of them had used it after relocating to UK, prior to the study.

The second group (F2) was a family from UK, and consisted of a married couple, their son and younger daughter, plus a family friend and housemate of the son. Husband and wife, both in their forties, worked locally, the former as a company director, the latter as a social worker. They lived with their younger daughter, who was in her early twenties and worked as a nurse and child caretaker at various locations out of town. The older son, in his mid twenties, was a PhD student and lived in another town with his long-term friend, also in his mid twenties, who worked as a business consultant at various locations around the country. None of them had ever used tracking technology prior to the study.

Each group was asked to use a custom-built location-tracking application installed on their mobile phones for a period of three weeks. The study consisted of four different phases, each lasting 5 days. During each phase participants were allowed or asked to do different things such as remaining exposed, carrying out tracking tasks on other participants or using location-sharing preferences.

Similar to the first study, this study also used experience sampling (Consolvo & Walker, 2003) to elicit the participants' reactions with deferred contextual interviews at the end where the participants were debriefed on the experiences faced in different contexts. The data collected from these interviews were in the form of audio scripts which were later transcribed to be used as an input for the distillation approach. In addition to the audio scripts, other artefacts such as location-tracking data and select screen shots from the participants mobile phones were also available for analyses. The outcome of the study is described in detail elsewhere (Mancini *et al.*, 2011).

B.2 Technical implementation

The study deployed a custom-built location tracking software on the mobile phones of participants. Key functions of the location tracking application included an interactive map which displayed (i) the location of all members in the group (ii) current location of single member (iii) history of past location visited by each member of the group and

(iv) settings to control the level of location details. The Figure B.1 shows a high-level architecture of the location tracking application used in the study.

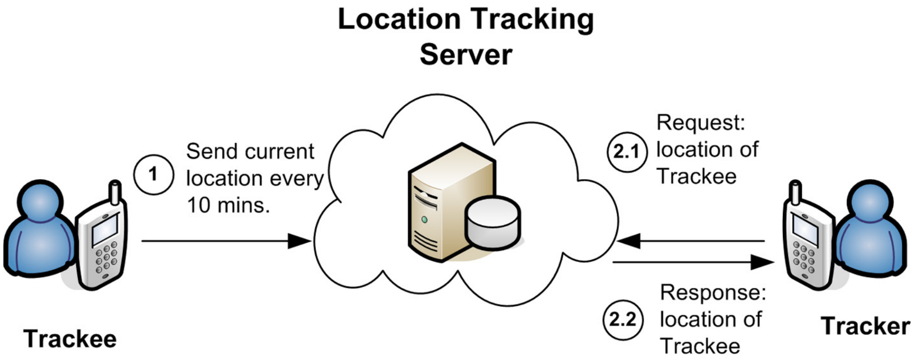

Location Tracking Server
1
Send current location every 10 mins.
2.1
Request: location of Trackee
2.2
Response: location of Trackee
Trackee
Tracker

Figure B.1: Location-tracking application

In the location tracking application, the user could take two possible roles: *tracker* or/and *trackee*. Tracker is a user making a request to determine the location of another user. Trackee is the user who is being tracked and is the source of location information (information subject). As shown in the Figure B.1, a client application on the trackee's mobile device sends the location coordinates to a location tracking server every 10 minutes (1). When a tracker makes a request for the location information of a trackee (2.1), the location-tracking server responds by sending the location details of the trackee (2.2).

Trackees were provided with facilities to control their location-sharing preferences on the application, for example, the application had a button which enabled the user to become invisible for a few hours (time-sensitive visibility settings) and another interface which allowed a group member to see one's location only at city level rather than street level (coarse or fine-grained settings).

For the purposes of this evaluation, only a subset of the use-cases from the location-tracking application have been selected. Table B.1 contains a brief description of these use-cases.

For a detailed description of the technical features, see Jedrzejczyk *et al.* (2010).

The software system for the study consisted of a client system on an Apple iPhone which allowed users to track their co-participants in the study. For this, the client

Table B.1: Initial requirements from the location-tracking application

ID	Requirements	Description
R1	Create location information	Location-tracking client application uses the facilities on the mobile phone to determine the current location and sends it to the location-tracking server.
R2	Switch-off location creation	The trackee (user) is able to switch-off the location creation functionality on the client application
R3	Change precision of location information	The trackee (user) is able to change the granularity or level of precision at which location information is created
R4	Show the location of trakee	The tracker (user) makes a request for the location information of trackee and the information is displayed on the tracker's mobile screen display.
R5	Specify who should be able to see location of the trackee	When the tracker makes a request for the location information of trackee, the query is answered depending on the control set by the trackee.

application made use of the Google Maps API to display the location of users on a map. In the third phase of the study, the client application supported additional features to provide real-time feedback on who is tracking whom and to provide fine-grained privacy-management controls.

An integral part of the software system was a centrally hosted Server which polled for specific events. When an event was detected, the server generated and sent an SMS to the user with an embedded URL (via 3rd party SMS gateway). By opening the SMS and clicking on the URL, participants connected to a User-Feedback System (UFS). The UFS was a web application optimised for display on a mobile i.e. iPhone. Once connected to the UFS, participants are prompted to answer experience sampling questions and provide a suitable memory-phrase.

The following are the transcripts from interviews with three participants whom we call 'CW', 'JW' and 'AW' for short. All three participants were siblings from the first family 'F1'. In the first interview, CW recollects her experiences with using the location-tracking application on her mobile device to locate her family members, especially her boyfriend 'RN'.

B.3 Transcripts from interview with CW

¶ **[1]** *Project officer:* OK. Now. So these are all the events recorded with you as a tracker. Basically they correspond to these, the feedback that you sent. We will go through these and I will ask you questions

¶ **[2]** *CW:* Yeah sure

¶ **[3]** *Project officer:* So one question here. How do you feel about seeing yourself? Like that?

¶ **[4]** *CW:* It is quite cool. It is quite interesting to see that I actually do move round quite a bit during the day. Kind of feel like I am only studying and then... I go to the gym and it is quite nice to see that. I don't really mind this project at all. I quite liked

¶ **[5]** *Project officer:* And what about the exercise of having to remember, you know, I was here and I was there

¶ **[6]** *CW:* It's difficult. Even though you see a pinpoint. So I know I was definitely there at that time, on that day and I still, some of the time I didn't even know what that was. It was quite weird to think that - that you forget so much as well

¶ **[7]** *Project officer:* Right, Right. OK. So this is the first, your first tracking exercise and you tracked, you looked at RN. You said you just thought of him. He was exactly where you expected him to be and you felt neutral about it. Do you remember anything about this episode?

¶ **[8]** *CW:* Well the first time I thought who shall I look up? Look up RN. Oh he is at work, that's nice. Nothing

¶ **[9]** *Project officer:* How do you feel about looking him up?

¶ **[10]** *CW:* I think it was fine providing we told each other during the day what we were going to be doing...As soon as you get to that point where you look and oh what is he doing there? He didn't tell me that he is going out...I think it can, as long as it always working fine and you are always where you say you are going to be...Then that is fine but as soon as there is something different then suddenly you get a bit suspicious. You are like "right. Oh. He is not at home. What is he doing and why didn't he tell me he was doing that." But generally with RN he was always at work or, I never really had any of those problems

¶ **[11]** *Project officer:* OK. And about the exercise of peeking into his, you know, into his life if you like, without him knowing. How do you feel about that?

¶ **[12]** *CW:* You do feel like, if he couldn't look me up I would feel like I was doing something wrong...It is the same thing as maybe going through his phone and reading his texts...You kind of feel like you are invading his privacy...But because I knew he could look me up it was different...I didn't really feel like that and also he knew that I would be looking him up anyway...So because it was like a short period of testing. So I didn't really feel like I was invading his privacy in any way

¶ **[13]** *Project officer:* OK. If this wasn't a testing... if this was real life, you know, situation and you would be looking him up, you know, not for a test but because you had this application on your phone and you just

¶ **[14]** *CW:* I don't know

¶ **[15]** *Project officer:* Would you have done it and if you had done it how would you have felt it? You know, how, sort of how would you think, because now you have, obviously this is a bit fake. This is a bit artificial because it is a study but because you have done it, even though it is artificial you can get a sense of what it might feel like in real life

¶ **[16]** *CW:* Generally I didn't actually think it was too invasive. The only thing, I absolutely don't mind anybody looking my location up at any time...Like I didn't really mind it at all. It is just with RN there was one night, I think it will come up later, when I was really ill...And I know he had gone out and I wanted to know where he was, so that I could know if he was at home and I could phone or if he was still out and wouldn't answer his phone, so I looked him up but it said it was last tracked four hours ago which was when he left work. SO I though OH did he turn his phone off

¶ **[17]** *Project officer:* Or did he turn the tracker off and why did he do that because I now he is going out, and so then and also I was ill so I start thinking so why didn't he tell me? Who is with? And why is he there

¶ **[18]** *CW:* And that was the thing. Actually he was at home and asleep. I phoned him and his phone was on silent so he didn't answer...Right so I phoned about ten times like why is the ticker off, I am ill. I need you to phone me so that was the only incident where it caused a problem...And I think with something like this if you had it for the long term I think it could cause o lot of problems like that...If you not where you say you were going to be, and his phone was off, actually then I was like it is 2 o'clock in the morning, I don't know where you are and I can't get hold of you, where are you and who are you with, and then in the morning he was like "Oh I am sorry". He, I felt bad because I had completely over reacted...Just because I thought he had turned his tracker off. I thought he didn't want me to see where he was

¶ **[19]** *Project officer:* OK. OK. And that did that feel

¶ **[20]** *CW:* Oh it felt horrible and I really stressed myself out because I was like "Oh well, if he has turned it off the obviously he... and then I though" I don't really know any of his friends live anyway so even if I had tracked him it wouldn't have made any difference. He would just tell me somewhere in [city2]...So I thought that was a bit weird, but still, but still I am still mad at you

¶ **[21]** *Project officer:* I guess, I guess you still would have seen, so is it the fact in your mind, because you didn't know, in your mind possible he had, hidden where he was from you

¶ **[22]** *CW:* Yes

¶ **[23]** *Project officer:* So wherever he was,

¶ **[24]** *CW:* Just the fact that it was off, yeah even if it was work and it was off or at home and it was off, you know, I would still think well, why don't you want me to see it and that was one time where I got really upset it...And actually what had happened, I think on his phone because he never charged it properly so when it dies he had to reinstall it so that is why it was off. The tracker was off but his phone was still on so I thought maybe his battery has just died and his phone was off but it was ringing so the fact that he had turned off I really like a bit cheated. Why wouldn't you want to know where you are?

¶ **[25]** *Project officer:* Yes. OK. OK. So, we will come to that later but you have told me. It is good that these things come up. So then you looked at RN 2 days late, so you were checking on him. He wasn't exactly where you thought he would be so you felt cheated. Is this the episode you were talking about?

¶ **[26]** *CW:* No *Project officer:* It was another one?

¶ **[27]** *CW:* It was another one *Project officer:* OK. So what is going on?

¶ **[28]** *CW:* I that day. I think it was right in the beginning so I know at Westfield the reception is not great so it doesn't pinpoint where exactly where work was...Sometimes it was just around. I remember first times I saw it when wasn't at work and I thought he was at his friends house, and I thought that's fine, I don't mind but you said you couldn't phone me because you were at work...Well you are not speaking to me because you are at work but you are not at work, you are down the road and I can see where you are. Then I realised he was at work so then I felt a bit bad about that...Then I realised it didn't always pinpoint the exact spot all the time. It was kind of a little dotted around

¶ **[29]** *Project officer:* How did, you said you felt bad, can you tell me more? First you said you felt bad because you thought he was not honest with you about where he was

¶ **[30]** *CW:* Yes

¶ **[31]** *Project officer:* Then you felt bad because you thought you thought he was being honest

¶ **[32]** *CW:* Yes

¶ **[33]** *Project officer:* So what went through your mind when you realised that actually, no, it wasn't his fault, that is not...

¶ **[34]** *CW:* Then I thought, obviously I did over react and I shouldn't have. I didn't say anything to him obviously

¶ **[35]** *Project officer:* And this would be confidential

¶ **[36]** *CW:* Oh yeah I don't mind. So yeah I really did feel like "I don't know why you are not telling m me the truth" because we actually had an issue like right in the beginning where his ex girlfriend, she was kind of in the scene ...So I had already got my guard up about that...And then because he wasn't exactly where he was going to be then I immediately assumed oh well he is obviously doing something else and didn't tell me about it ...So, but then that was an irrational thought and it did feel bad about it because he was in the right place doing work, when he was supposed to so you get quite mixed feelings. You get angry and you feel bad because you were angry

¶ **[37]** *Project officer:* Yeah. Yes. OK. OK. Thank you. Then this is, then you looked at JW on the same day and you were curious to see where she was and she was exactly where you expected and you felt neutral about it. Do you remember anything about this episode?

¶ **[38]** *CW:* Also in the beginning, you know, I wanted to see where JW's work would be on the map because I didn't exactly know, you know, in the beginning just finding out where everyone was, where work would be for them...and where home would be so yeah, I think that was definitely be at work

¶ **[39]** *Project officer:* What feeling does it give you seeing where everybody is?... You know. How..?

¶ **[40]** *CW:* It's, you do feel a lot closer to people...And also I know when my mum was coming to visit me on one of those days. She was supposed to arrive at eleven and she only arrived at like an hour later...So instead of phoning to say where you are, are you ok? ...Was there an accident or what. Obviously as she was drinking she couldn't answer her phone so I searched her and saw that she was close by so I thought Oh that's fine ...So that was quite nice and I think in that respect?? my parents I wouldn't mind having a tracker because that is different, and also like with my sister

¶ **[41]** *Project officer:* How is that different? So, OK this sounds interesting so you can tell me more. So with your parents it is O, because it sounds like there would have been about with somebody else

¶ **[42]** *CW:* With RN I think it is different because for my parents...It is actually quite useful for them to know where I am. If I am driving, if I am a train, to see if I am close by or far away...Same as my parent, same as if they were visiting me...But then if I searched my mum and she was not at home and she said she was going to be at home then it doesn't really make a difference to me because she is doing her own thing...Whereas if RN said he is at home and he is not at home...You immediately think I am not a jealous kind of person really but then you just have a slightly response because I don't really care what my parents are doing during the day

¶ **[43]** *Project officer:* So it is a different emotional investment?

¶ **[44]** *CW:* Yeah

¶ **[45]** *Project officer:* Different expectations of openness

¶ **[46]** *CW:* And if my parents aren't where they say they are going to be it doesn't matter but If RN is not where he says he is going to be it matters more to me

¶ **[47]** *Project officer:* And it matter because ...?

¶ **[48]** *CW:* It matters because he is my boyfriend and I want to know where he is. I don't really know what the difference is. Maybe because we had that problem in the beginning I am kind of more aware of I need to know where he is and also, especially with RN, because he has got a static...I think because RN at the moment, he has got quite a static life style...He is at work. He is at home he goes off with friends and stuff...And I usually know about all that...So then if he does something completely different I don't mind, it is just more curious and

¶ **[49]** *Project officer:* It is so easy...

¶ **[50]** *CW:* And I don't want to have to jump to conclusions if he is not in the same place

¶ **[51]** *Project officer:* Yes. So it is the mismatch really between what you think you know about what he is up to

¶ **[52]** *CW:* Yeah

¶ **[53]** *Project officer:* And what you might find with the tracker?

¶ **[54]** *CW:* And also what I didn't like about the tracker was I like surprising people and I like going to [city2], surprising him and like coming to [city] and surprising me...But you can't do that because "Oh you are on train on the way to [city], I know, you are not surprising me"

¶ **[55]** *Project officer:* Also with his birthday I wanted to organise a lot of things during the day as a surprise?? To get his cake and special balloons but then he could see where I was and was "oh what are you doing in this part of [city2]?" "I am getting you a surprise Birthday Cake". Those kinds of things it is difficult if everyone know where you are all the time

¶ **[56]** *CW:* Do it, and I don't really surprise my parents that much. Maybe I should surprise my parents more, I don't know...OK. OK. Good, good. OK So you say that with your sisters is similar as it is with your parents because you said they do their own thing so it is like your lives are not necessarily sort of connected or bound together

¶ **[57]** *Project officer:* And, like the fact that my sister has a day of or doesn't

¶ **[58]** *CW:* Yeah doesn't impact on me but if RN has a day off I like to see him, or if he has an evening off we like to do something...When he has time off from work I am involved in that where as I am not involved anything??

¶ **[59]** *Project officer:* Yes, OK Good. AW. You said, this is the same day. You said you were curious to see where she was she was exactly there and you felt neutral again

¶ **[60]** *CW:* Mmm, she was at school

¶ **[61]** *Project officer:* She was at school. So how did that feel looking ... OK a question? The question is so you said that with RN, it matters more to you to know what he is up to so it might be a bit tricky if you find that he doesn't seem he is exactly where he said whereas it doesn't matter so much with your other family members so how does it feel to look, to peek into RN's life and to peek into your mums and your sisters and dads life? So the fact, just the fact, not so much the reaction when you find what you find, I would like to know how it feels when you decide to do it, the moment when you decide to do it, how you feel about your action of, your decision to check and to look?

¶ **[62]** *CW:* Oh right

¶ **[63]** *Project officer:* So how is that different, if it is different, and how is that different?

¶ **[64]** *CW:* I think my decision to look; it depends on who the person is I guess?

¶ **[65]** *Project officer:* So my decision to look where my mum my dad, JW or CP are

¶ **[66]** *CW:* For me I don't feel, I do feel completely neutral about it...I know because it is a study it is slightly different...But even if it wasn't a stud. Slightly different with AW though, I think because she is the youngest sister...I look where she is on a Friday night for example...You have a slightly more protective feeling...Oh, she is out. I wonder if her boyfriend driving her home tonight?...That is slightly more protective but actually looking to see where they are, I didn't really find it too bad. I didn't feel bad about doing it so...

¶ **[67]** *Project officer:* But...

¶ **[68]** *CW:* No, there is no but...

¶ **[69]** *Project officer:* I mean with respect did you because I didn't feel bad about looking up them, you know, AW, JW but did you feel that about checking looking at RN

¶ **[70]** *CW:* I feel I would only feel bad about checking him because my response would be different...So but the actual checking on him, where he is was fine...I didn't really think it was...I didn't really feel like I was doing something wrong...Did that answer your question?

¶ **[71]** *Project officer:* Yes, Yes, I think so. And then you looked on your dad on the same day, same month, and same set of answers. So you were curious where he was and he was exactly where you'd expect and you felt neutral about it. So is this the same with your dad as with your mum and your sisters?

¶ **[72]** *CW:* Yes, yes. It doesn't really matter what they are up to during the day

¶ **[73]** *Project officer:* OK. So then y our mm and again the same response

¶ **[74]** *CW:* Yeah. I was just trying to find out that the same time

¶ **[75]** *Project officer:* OK so same answers as well for everybody. Then we go down to 29 January and you looked at your dad that was a Friday

¶ **[76]** *CW:* OK

¶ **[77]** *Project officer:* And you looked up your dad. Again the same answers. You were curious to see where he was. He was exactly where he said. You felt neutral about it. So shall I take that this is all the same that you explained to me?

¶ **[78]** *CW:* Yeah

¶ **[79]** *Project officer:* The this is?? because on the same day you were checking on your mum, looking up your mum, when you say you were missing her. She was more or less where you expected her to be, you felt closer

¶ **[80]** *CW:* I looked up my mum, I was studying and it was Oh, there's my mum, and she was I think driving between the shops and home. So I could see her on the road and it was Oh, she has obviously gone shopping...And it was nice you did feel a bit closer because you knew what she was doing...Yeah. That was good

¶ **[81]** *Project officer:* OK. OK. When you check, you look at somebody because you are curious to see, you wondered where they are or whatever and when you were sort of missing them, what, do you feel different, was it different when you check for the different reasons?

¶ **[82]** *CW:* Yeah I suppose it is. I think because one; sometimes I might be bored studying and I am not sure what to do, Ah check where everybody is so it is kind of prompted by boredom or,...Whereas this time it was prompted by me saying I really miss my mum...So, you have the initiative as to why I looked was actually because I did miss her. I wanted to see what she was doing, didn't have time to call or...

¶ **[83]** *Project officer:* Did you feel, did you feel different because you were doing it for that reason

¶ **[84]** *CW:* Not particularly, I mean I don't think it made much of a difference. When I got the result it kind of, instead of it being like "ok she is driving" Ok the result evoked different emotion where ah that is nice and I kind of maybe a bit homesick, made me feel a bit closer. So ...Looking them up wasn't that much different. Maybe a different prompt but the emotional response when a got the answer was different

¶ **[85]** *Project officer:* Yes it was different OK. OK. The RN again. So, oh, you were curious to see where he was he was exactly where you expected and you felt reassured

¶ **[86]** *CW:* Yeah. He said he was at work and he was at work. That's good

¶ **[87]** *Project officer:* OK. So he definitely used a stronger set of answers. These feelings here in this bit. So when you felt reassured it is like everything as it should be. And is good

¶ **[88]** *CW:* Especially with RN as our relationship is difficult because I am in [city] and he is in [city2]. It is not only that but we do manage to see each other but he never has weekends off so even when we d see each other at his restaurant studying on a weekend or when he comes to visit me he is just around [city] while I am at lectures...So we never really have time off together and it is even more difficult because he has such late hours...He will finish work at one/two in the morning and you are asleep by then...So we go for days when we don't even have a chance to have a proper conversation...A few texts or a few quick phone calls so for us it is quite difficult, so then of course I couldn't speak to him for a few days, I knew it was a Wednesday, he is probably at work, let me check. OK he is at work he is still working that that's fine. Because our relationship is a bit strained because of the different response but definitely from that. I knew he was working Wednesday He said he was working I could see he was working. OK. That's good. No worries. I can carry on with my day

¶ **[89]** *Project officer:* Everything is fine. OK. Ok.

¶ **[90]** *CW:* Even though he hasn't text me I can see he is at work. He is probably busy

¶ **[91]** *Project officer:* OK. OK. So if you, another question for you. I am sorry I am asking load of questions. So if you see, because you said something interesting here. Even though he hasn't texted me I can see he is at work. So how would he feel, do you think, how does he feel if you check a lot? Not a lot. You check him one time, another time and he is, you now, you can see what he does. So you can see if he is at work, at home. So regular places. Nothing, you know, crazy, but he is not text you. He hasn't text you yet so how does the fact that you can see where he is sort of, but he is not getting in touch with you

¶ **[92]** *CW:* I think...

¶ **[93]** *Project officer:* How does that feel?

¶ **[94]** *CW:* It probably makes it, like I said again at the time; it's different because you think I know you are at home. I have sent you a text 20 minutes ago and you have still not replied. You are watching TV. You can't be bothered to reply, you don't have your phone on you...So you always, you come up with different scenarios in your mind, and not that that is usually a problem. I know when he it as work we don't chat. When I am at university we don't chat ...So, but then if he was at home, if he had a day off, if he hadn't spoken during the day that wouldn't be normal and I would be upset but I suppose because I could see, because we have got such regular routines I could see he was at work. Yes that is fine he is busy, I am busy, but if I had seen that he was at home and he hadn't texted me I would again feel like maybe hurt, maybe I would have texted you if I wasn't doing anything. I think it is quite hard having it with your boyfriend

¶ **[95]** *Project officer:* OK. OK. Thank you

¶ **[96]** *CW:* I hope I am giving you some interesting feedback.

¶ **[97]** *Project officer:* yes, yes, very interesting and I need to, how can I say, I need to probe because I am trying to find out as much as I can from these, maybe a bit heavy because I need to get all these questions over. And then on the 31st, that is 2 days later you looked up CP and say you wanted to check if your tracker was working because your ?? three hours

¶ **[98]** *CW:* This was the night this was the

¶ **[99]** *Project officer:* Ah, this was the night that you got upset...2.48 am... So you checked, you looked at CP to verify that the system was working?

¶ **[100]** *CW:* Yeah I thought maybe all the systems was down

¶ **[101]** *Project officer:* So you tried to look at all the possibilities to exclude, I see. So he was exactly where you expected. So that was 2.48am. Wow. OK

¶ **[102]** *CW:* That was the night when I was. And then I was like Wow. You obviously turned your phone off

¶ **[103]** *Project officer:* So you felt cheated obviously. Not because you found CP was, you were still thinking of RN

¶ **[104]** *CW:* Wow, I must seem such a horrible girlfriend

¶ **[105]** *Project officer:* No you know we all have these fiends and it was a stressful situation for you

¶ **[106]** *CW:* I was really ill.

¶ **[107]** *Project officer:* That's is interested because you thought. He was exactly where you expected but you felt cheated. Obviously because you weren't sort of responding to where CP was but

¶ **[108]** *CW:* Memory Phrase-3am I remember that one, OK

¶ **[109]** *Project officer:* 3am yes. Ok and then on the same day, this was 7 in the morning really and so you didn't get much sleep that night and AW, you looked at AW. You were curious to see where she was and she was exactly? Where you expected. So is this a similar situation n to the ones you have described before?...So this is a similar situation to those you described before. You checked on your sister your email, it feels neutral. It doesn't really matter but with AW you feel a little more protective. But this doesn't seem to be complicated? Oh, RN again on the 31. So, ah, OK. This is because I sent you a...So he was exactly where you expected and you felt comforted by it?

¶ **[110]** *CW:* Yeah

¶ **[111]** *Project officer:* OK. So anything that you remember particularly about this?

¶ **[112]** *CW:* No Nothing

¶ **[113]** *Project officer:* Then back again. Oh, prompted again. It was more or less what was expected so you felt pleased so was the different between you feeling comforted here and at work it was exactly what you expected and feeling pleased more or less?

¶ **[114]** *CW:* I think it probably depends on the time of day that I looked because if I was at home, what time what that. 6 o clock.

¶ **[115]** *Project officer:* This was the next day. This was nearly 6 o clock and this one was sort of lunch time

¶ **[116]** *CW:* I think if I was at home and I was alone...And I am getting ready to go to bed and I know that he is at home or he is working...Then that kind of made my response like OK everything is fine I can go to sleep now where as during the day it's OK, or that's nice. He is at work...I suppose please and yeah I supposed it is not that much different to neutral but a happy neutral if that makes sense...So possibly the time of the day does ...

¶ **[117]** *Project officer:* OK, and perhaps the day because this was a Sunday so

¶ **[118]** *CW:* Yeah

¶ **[119]** *Project officer:* He wasn't then, I meant for you

¶ **[120]** *CW:* For me?...I suppose because I was in a comfortable situation I kind a little bit more reassured, whereas if I am out during the day studying or writing an exam the fact that he is at work at home, it is a lot less relevant

¶ **[121]** *Project officer:* A lot less relevant, OK. So this is the one ?? and then your mum again

¶ **[122]** *CW:* Ahhh

¶ **[123]** *Project officer:* On the 1 February 6ish pm. You were missing here. She was exactly where you thought she was and you felt comforted about mum

¶ **[124]** *CW:* I was at phone and I could just picture her having dinner, watching TV and...

¶ **[125]** *Project officer:* OK. Is that to do with the fact that kind of, everything is where it should be in the right place and everything is in order?

¶ **[126]** *CW:* Yeah and I think especially during my exam time...Because I am under a lot of pressure I am up every morning and I am studying because I am doing my Masters now so it is a lot of pressure for these exams, my final years. So because I was under quite a lot of pressure, for everything else to be OK, in the right place. I don't need to worry about anything else

¶ **[127]** *Project officer:* OK

¶ **[128]** *CW:* That helps a lot as well. The fact that mums was at home, AW was at home. I kind of think OK. I have almost finished these exams. Soon I can go and be with them...Yes...So it definitely does kind of help when I am in a situation because everybody else is where they should be

¶ **[129]** *Project officer:* OK. Yeah. And the RN again, on the same day. A bit later, in the evening

¶ **[130]** *CW:* Ah, in the evening

¶ **[131]** *Project officer:* You were curious to see where he was. He was exactly where you expected and you were pleased about. Ah you wrote a memory phrase her because you don't use... you did?? January. Here. Monday 1st. You were waiting for RN to visit

¶ **[132]** *CW:* Yeah. I was

¶ **[133]** *Project officer:* OK so what..

¶ **[134]** *CW:* OK. He was coming from work and because he finishes late he was on the train on the way to see e. I could see oh he was half way. So I could see where he was, what time... because I wasn't exactly sure what he would be getting in so I could see he was wherever and maybe it would be another hour or so

¶ **[135]** *Project officer:* OK. OK. RN again on the 2nd. So at that point he was in [city] on the 2nd?

¶ **[136]** *CW:* Yes

¶ **[137]** *Project officer:* Yes. OK. You say you just thought of him. He was exactly where you thought he was and you were comforted by it. So you remember where he was? Where he would have been?

¶ **[138]** *CW:* Yes because I was, I was just about, just before my exam I was waiting to go into the exam and I just wasn't sure if he was going to be at home or if he was going to town. And then that I thought, I wasn't sure if he was going to go out or not and then I knew he was going to actually wait at home for me so I checked and he was at home because I didn't leave keys for him...Then I thought he is at home he will probably stay there. That is fine because he doesn't have keys. Then I thought I don't need to stress out about that. I am going into my exam. I don't need to phone him and say, don't go back. I don't keys for you...OK he is still at home he hasn't gone out so that is fine

¶ **[139]** *Project officer:* OK. So that was more of a practical thing. So does it feel different when you check, when you look him up for practical reason or for an emotional reason if you see what I mean

¶ **[140]** *CW:* Yeah, I suppose practical reason you are like, you just don't even have a second thought about it at all

¶ **[141]** *Project officer:* Yeah

¶ **[142]** *CW:* I need to see where you are. Do you need keys? No you are at home. You don't need keys, that's fine. I suppose with an emotional one I was missing him or I was checking on him. I suppose it is a bit different. You don't have that second thought, or you don't think what the answer is going to be...You don't have that trepidation or apprehension...Before

¶ **[143]** *Project officer:* OK. And when you u were asked to check how does it, how is that different? When it was me asking you to check on RN? How was that different from checking him by your, you know, your own initiative. Not for practical reasons?

¶ **[144]** *CW:* I suppose because you were prompted you don't have a second thought about it actually. You are like OK let me check where they are. And when you get the result, you are like OH, Well that's nice. You are at home and I didn't need to be a snoopy girlfriend...So it is kind of like "That's nice, I can see where mum is, I don't need to be there"

¶ **[145]** *Project officer:* OK but is it more, when you are asked to do is it more

¶ **[146]** *CW:* It's easier

¶ **[147]** *Project officer:* You are fulfilling... easier? Mmm You are fulfilling a task or something. OK. Good. Let's move on. Your Dad. Same day. Just thought of him. He was exactly where you expected. You felt comforted by it. 92 was the memory phrase that you chose.

¶ **[148]** *CW:* 92? What the hell was I doing? Was this the same day as the exam? I think so because it was the same day you were worried about the keys, yes. So the same day

¶ **[149]** *Project officer:* 11:15 am. This was same time

¶ **[150]** *CW:* I don't remember. A phrase 92?

¶ **[151]** *Project officer:* 1:00...Don't worry if you can't remember. I was just curious about 92? I there anything different to say about this episode. You were doing exams, you thought of your dad. You checked and felt comforted by it. Is it different from the other episodes you told me about, where you were checking on your family?

¶ **[152]** *CW:* This one would be where I was missing my mum. I just wanted to feel a bit closer. You should look at my report mum. Glowing. This was the same, I was missing my dad, and I wanted to see what he was doing

¶ **[153]** *Project officer:* OK. OK. RN again. You just thought of. This was 4th, a Thursday. You just thought of him and he was exactly where you expected and you felt comforted.

¶ **[154]** *CW:* Ah. Thursday. He was still, he was still, he was in [city]

¶ **[155]** *Project officer:* Ah, he was?

¶ **[156]** *CW:* This Thursday. I wrote

¶ **[157]** *Project officer:* So do you remember what is was?

¶ **[158]** *CW:* Let me just check what. Thursday... no I finished... no wait... what day did I finish the exams? I finished my last exam on the Tuesday...And then we went out Tuesday night and then I drove to [city2] on Wednesday morning. I was at [city2] on Thursday...So I was at his house waiting for him to finish work...and then I saw he was on his way home...I was just waiting at his house...Thursday. RN's house

¶ **[159]** *Project officer:* OK and so, you felt comforted because he was coming

¶ **[160]** *CW:* I was home alone at his house. His flat mates weren't in so I was watching TV by myself. I was fed up of being in the house by myself so when I knew; I thought OK he will be home soon. I thought that's fine

¶ **[161]** *Project officer:* OK. Same day. 11.14 here. OK let me look at the chart. That was the night you were still, at this point you weren't just thinking of him you were checking on him.

¶ **[162]** *CW:* Oh no, this must have been,, because I was at his house and maybe he was still at work...Yeah... and then is must be he is on his way home because it is 11:15. It is 11.45, now I am by myself. Come home. Where are you? Then I was checking

¶ **[163]** *Project officer:* OK Yes

¶ **[164]** *CW:* At 9.30. ah, let me think about my boyfriend. 11:15 Where are you?! That must have been it. He must have been coming home late a bit. I was definitely home alone that night

¶ **[165]** *Project officer:* OK. So, OK. So he was exactly where he was supposed to be, so he was on his way home?

¶ **[166]** *CW:* Yeah

¶ **[167]** *Project officer:* So was OK. So you felt neutral about it? That's interesting because you were checking but you felt neutral.

¶ **[168]** *CW:* I think because, to be fair I think he did say he was going to be home at like 11.45, but I was just checking because I knew he was going to get home a bit later.

Maybe wanting to see if he was finishing earlier, maybe he was staying later. I don't know...Neutral because I knew he was going to get home at a certain time

¶ **[169]** *Project officer:* OK. OK. So iIs this like saying...?

¶ **[170]** *CW:* Well I shouldn't really be checking

¶ **[171]** *Project officer:* Yeah or invest too much because I know it might be one way or the other

¶ **[172]** *CW:* To be fair I shouldn't have even checked but I did

¶ **[173]** *Project officer:* Why did you?

¶ **[174]** *CW:* I don't know. I think because I was alone and bored and...

¶ **[175]** *Project officer:* OK. OK. RN again. This was on the 5th. But you were in [city2] at this point were you? Or not?

¶ **[176]** *CW:* Memory

¶ **[177]** *Project officer:* This would have been a Friday . Had you just arrived?

¶ **[178]** *CW:* I was out. No Friday

¶ **[179]** *Project officer:* So you had just arrived the evening before to [city2]

¶ **[180]** *CW:* Yeah, no. I had actually arrived on the 3rd...And I was there Wednesday Night, Thursday night. Oh Frankie and Bennies. I was here in [city3]

¶ **[181]** *Project officer:* So you had come back up?

¶ **[182]** *CW:* I had come from [city2] to [city3]...And I saw a film with my parents. We were out for dinner...And he was just, he was just at work

¶ **[183]** *Project officer:* Ah, you were prompt, so again, as you said before it was more like...?

¶ **[184]** *CW:* Completing a task. I wouldn't have checked on him because I knew he was working. I was out to the film, so I wouldn't have unless I was prompted that night

¶ **[185]** *Project officer:* Ah. OK. OK. So let's see, so that is done. We are close to the end. RN again it was on the 7th, a Sunday. You were prompted. He was exactly where you expected and said neutral. You were in [city3]. So you were in [city3] still?

¶ **[186]** *CW:* Yes

B.4 Transcripts from interview with JW

¶ **[1]** *Project officer:* this is the first feedback form recorded in our system. This was 18th January

¶ **[2]** *JW:* so that was right when we started?

¶ **[3]** *Project officer:* yes. So you looked at all your friends, all your buddies. Do you remember anything about how you felt. It was the first time you've seen that.

¶ **[4]** *JW:* Just interesting to see they all popped up, it all worked. Everyone was kind of where I expected them to be. Yeah nothing significant I don't think.

¶ **[5]** *Project officer:* what did you feel finding everybody more or less where you expected them to be.

¶ **[6]** *JW:* You know it is always nice to know that you kind of know, but I must say I don't feel very comfortable with this technology so I don't like using it very much at all.

¶ **[7]** *Project officer:* so how did you feel about checking or looking at everybody in the first place

¶ **[8]** *JW:* I wouldn't have checked it if you hadn't asked me to, so it is not something, if I want to know where somebody is they'll tell me. That's how I feel. If there is some reason I need to know why, where somebody is, then I'll ask them. But to have a running total of where they are and what they are doing it's a little too intrusive for me. So when someone says check out where they are I'll do it of course no problem but it is not something I'd do anyway.

¶ **[9]** *Project officer:* so what do you feel you are intruding in. I have to dig into what you are saying and...

¶ **[10]** *JW:* because I don't like people checking up where I am, I assume that other people don't like me checking up where they are

¶ **[11]** *Project officer:* why don't you like it then?

¶ **[12]** *JW:* Because part of what is so nice about living in a big city, is you get to be anonymous, so knowing that there is someone actively monitoring where you are and what you are doing, even if they really know because I talk to them all the time. So they know I am at home or at the gym but the knowing that in the background there are people who are actively checking up on you it's not how I like to run my life.

¶ **[13]** *Project officer:* what do you think that takes away from you? Why don't you like it?

¶ **[14]** *JW:* It's a bit of freedom, you know. If I want people to know where I am I will tell them. If I don't then I won't and like I was interested to hear that you might think there is some crossover from my experiences in [another country] because in [another country] it was essential to know where everybody was all the time and if you didn't it was a problem. If someone said I'll be home at 2pm and they weren't home at 2pm you started to worry about them. So we were constantly checking up on each other all the time. It's why I moved here so I don't have to do it any more. So I don't want to live in an environment where it is necessary so I sort of felt that the freedom that you get from being an independent and running around [city2], now just the fact that people can follow me if they want to makes me uncomfortable.

¶ **[15]** *Project officer:* what do you think, and this may sound like I am asking silly questions, but I am trying to understand the reasons behind. What do you think you need that freedom for and why is that freedom precious to you?

¶ **[16]** *JW:* I think it is precious because intrinsically it is something you only get in a safe and free society. It is one of the perks so in a lot of places in the world people are tracked and monitored on a constant basis. Either by the state or other agents or families, you know with social pressure a lot of people don't have that independence, that individuality, so when you come here it's one of the bonuses of living in a society like Britain. To give that up means you are giving up one of the major benefits of living here so why is it important. It's like being able to vote, it's like being able to choose what I wanted to study it's one of these essential elements of living in a society so I see it as a key thing that you have to have and I wouldn't live somewhere where I didn't have it.

¶ **[17]** *Project officer:* if somebody was monitoring you all the time, the choice they are taking away from you is to share or not where you are. Is that a choice that they are taking away from you?

¶ **[18]** *JW:* That's exactly it. They are taking away that choice. They are taking away my control over that information and I know they have CCTV cameras and I know all the information is out there but it is that habit of monitoring. Here you know that if someone wanted to find out about you they could, but they are not following you all the time and it's quite an important difference.

¶ **[19]** *Project officer:* OK yes. Would there be a difference you think if somebody you didn't know, a stranger, looked up your location, as opposed to members of your family or friends or people you have a relationship with.

¶ **[20]** *JW:* I would still feel uncomfortable with it.

¶ **[21]** *Project officer:* what would make you feel more uncomfortable?

¶ **[22]** *JW:* More uncomfortable, that's a good one. I'm not sure. On one hand people you know that are checking up on that information is immediately useful to them in some way, good or bad. Someone you don't know it can't really impact on you at all. I'm not sure if that is slightly more scary because what are they going to do with it, at least they're family I trust them. So there is that, even if my mother knows where I am, she is not going to try and sell that information to an advertising company. So I am not sure what would make me feel more uncomfortable and I guess if this is this anonymous person and you don't know its happening. It's probably happening all the time because everything you do is recorded, every time you withdraw money it's tracked and stuff. So I guess it would just be part of that general living in the modern society thing where everything you do gets traced at some point. I might have to give you an I am not sure.

¶ **[23]** *Project officer:* don't worry because there are more things to go through, to think about whilst we do this.

¶ **[24]** *Project officer:* the next thing was also on 18th January you were curious to see where these people were. You didn't put a memory phrase. You checked on AW, you said she was more or less where you expected and you said you felt pleased.

¶ **[25]** *JW:* She was at home I think this was in the evening and this one also was in response to a question. I only used it off my own initiative once when my father was really late and we had a lunch date. I know he gets as lost as badly as I do, so I was trying to find out where he was but he wasn't answering his phone so I was like he, is still on the tube. All the other times were on request.

¶ **[26]** *Project officer:* this is 18th January. This is before we started giving you location tasks.

¶ **[27]** *JW:* Was it?

¶ **[28]** *Project officer:* yes. Ok then I must have checked off my own back which is great. It was in the evening and I think I was quite pleased because she has such a lot of homework and so OK she's at home. I hope she is doing her homework.

¶ **[29]** *Project officer:* can we stop here for a second. You said you felt pleased to find that your sister was doing their homework. Why did that make you feel pleased, you knew she checked and she had a lot of homework to do. I was pleased to see she was doing it. What's going on there?

¶ **[30]** *JW:* That sort of overprotective, big sister thing where truthfully I probably wasn't going to give her a call. If I did she may not have answered. It's like poking your friend on facebook it is an easier slightly lazier way of staying in touch, so oh yes she's at home she must be fine, great. It's always good to know that your family is where you expect them to be.

¶ **[31]** *Project officer:* was that to reassure yourself that she was OK or she's at home and doing what's she's supposed to do. Because you mentioned her homework.

¶ **[32]** *JW:* I assumed that if she was at home she was probably doing her homework so it was really

¶ **[33]** *Project officer:* what made you feel pleased or good

¶ **[34]** *JW:* it is good because she had been quite stressed out and I know my mum comes down on her hard when she doesn't do her homework

¶ **[35]** *Project officer:* OK

¶ **[36]** *JW:* so I see AW's getting on with it, that's fine. It was OK

¶ **[37]** *Project officer:* so because she was at home you inferred that she must be doing homework.

¶ **[38]** *JW:* Yes I assumed so, in the evening like that, chances are she's doing her homework.

¶ **[39]** *Project officer:* 18th January again so you were testing and checked on your mum and said she was more or less where you expected to find her and you were pleased about it. Do you remember anything about this episode.

¶ **[40]** *JW:* It was also the same day and I think it was in the evening. We were just sitting here and probably talking about the app and how it works and all that. Probably trying to see how it works. Again it's always good to know that everyone is home safe and sound. It is not such a major security thing here but it is always nice to know your folks are at home. It's comforting.

¶ **[41]** *Project officer:* you said you were here with CR, I assume you did that together with him, or did you do it on your own.

¶ **[42]** *JW:* I probably did it on my own but we have talked about it quite a lot, how we feel about it, it's jail breaking and he's got a lot of techie friends and he got quite excited about jail breaking. I don't recall us ever doing it together but we did it while we were together.

¶ **[43]** *Project officer:* so was he aware that you were checking on your mum for instance.

¶ **[44]** *JW:* I am sure he was, I probably would have said, let's see where my mum is.

¶ **[45]** *Project officer:* do you remember if you felt any different when you were checking on your own or when you were checking with somebody else's awareness, with CR say, did it feel the same or different.

¶ **[46]** *JW:* I felt the same always, I felt like a bit of a voyeur so even if you like the answer, never feeling entirely comfortable by asking it and I don't think CR shares my reservations. He was more like let's try this, let's try that.

¶ **[47]** *Project officer:* so he was more involved with the technology

¶ **[48]** *JW:* yes absolutely

¶ **[49]** *Project officer:* you were more focused on the relationship

¶ **[50]** *JW:* yes I think so

¶ **[51]** *Project officer:* the interaction with people behind it. So you said I didn't like doing it I felt like a voyeur, how did that make you feel

¶ **[52]** *Project officer:* I don't know if you have a Facebook page, have you ever checked up on some ex boyfriend or something from 20 years ago, you're interested but you feel slightly embarrassed and you feel, I should be a bigger person than this, but I have to know, I cant resist. So I am gong to check anyway. It is kind of like that, I don't think it makes me a better person to be checking up on people.

¶ **[53]** *Project officer:* why is that?

¶ **[54]** *JW:* I think it implies a lack of trust, it implies some sort of insecurity on my part. If I had spoken to you today or yesterday and you said I am going to play golf, now I am checking [INDICIPHERABLE]

¶ **[55]** *JW:* I think what drives these questions, where are you, they are not positive things. It is either I am afraid for you or I don't trust you, and that's probably just totally personal to me but

¶ **[56]** *Project officer:* that's what I want to know

¶ **[57]** *JW:* for me I assume everything is fine unless I hear different so I am only going to check up on someone when there is a question something's not fine. It's never really a positive thing that's driving it, even if yes he's playing golf just like ???????????? it doesn't come from the better part of your nature I don't think.

¶ **[58]** *Project officer:* now do you remember when you decided to check on something was it more to do with say, I am doing this or I am thinking of this person, now I move on to this. Do you remember what you were doing when you checked, what you moved on to afterwards

¶ **[59]** was there a trigger?

¶ **[60]** *JW:* Yes

¶ **[61]** *Project officer:* in your situation, or was it more to do with what you thought the other person might be doing at the time. Was it more to do with what you might have been doing that made you think OK I'll do this and if so what did you move on to afterwards, if there is a pattern. Or is it I wonder what the person is doing. Like before with AW I knew she had a lot of homework to do so I checked basically, that sounded like what was going on with her.

¶ **[62]** *Project officer:* I think it would start with thinking about that person, having some random thought like my dad's having a meeting in [city2] today, then you start thinking what are they doing, where are they going, ah I wonder where they are. I don't think there would be anything that I was doing that would make me want to, except for that instance when my dad was late. Where I was waiting for him and then it was the fact that I was waiting. I think with everything else it would have started with thinking about them.

¶ **[63]** *Project officer:* do you remember if there was any pattern in you checking in the sense, I usually checked after breakfast or after lunch or I usually do it when I start work. Was there anything like that?

¶ **[64]** *JW:* nothing like that

¶ **[65]** *Project officer:* OK. Now we jump to 29th January and you checked on CR, or request, you were asked to check and he was exactly where you expected him to be.

¶ **[66]** *JW:* Yes he was at home, I knew he was at work.

¶ **[67]** *Project officer:* anything else you recall

¶ **[68]** *JW:* it was just before lunch, we email each other throughout the day so I already knew what he was up to and how his day was going. I think I might have made a joke about stalking him. I don't remember anything unusual, he was at work and we always talk a lot through the day. It didn't tell me anything I didn't really know.

¶ **[69]** *Project officer:* next, still on 29th January and that was you checking on your dad and you were requested. He was exactly where you expected him to be but this time you said you were pleased, so if that was a difference, why, is there something different.

¶ **[70]** *JW:* It was just early evening and I think he was at home that time and my dad works from home but he travels out to go to meetings every now and then. I think he may have been out that day and I thought he's back at home with my mum, that's nice.

¶ **[71]** *Project officer:* is it that same feeling of reassurance

¶ **[72]** *JW:* very much so, they're safe and sound

¶ **[73]** *Project officer:* did you feel strange checking on your dad, how did that feel

¶ **[74]** *JW:* it feels slightly different checking on my parents to checking on anybody else.

¶ **[75]** *Project officer:* so what is the difference that you feel there

¶ **[76]** *JW:* Your parents have less privacy towards their children I think. You can take more liberties with your parents. There is that understanding that they are always at your disposal, especially my folks they are always available for us, so it has always been anytime we needed anything they are there for you. It's been all through my life, if I ever needed to know something from them or about them I just needed to ask and I would probably have been told. So it doesn't feel as evasive. I don't feel as bad checking up on them because I know my mum welcomes it. She loves any form of contact, she always wants to talk to her girls, I know if she saw that I checked up on her she would be thrilled. My dad he would be quite pleased but he is much more chilled out about these sort of things. I know that they welcome the contact so I don't feel like I am invading any of their privacy.

¶ **[77]** *Project officer:* the difference here would be that whereas with a phone call or something then there is a direct communication, in this case if you check on them they don't necessarily know unless when we get to a further phase of the cycle if you remember with the feedback, the real time notification, but at some point there was a time when they did not necessarily know. Did that make a difference?

¶ **[78]** *JW:* I feel with my parents, I feel entitled to know, I don't know if it is reasonable or not, but even more so than my sisters. I think it is because your parents are always your parents and that's what they are, I feel entitled to know where they are all the time and what they are doing whatever.

¶ **[79]** *Project officer:* with your sisters is it a bit different?

¶ **[80]** *JW:* It is, especially with CW now that she is a bit older she's her own independent person, she's living on her own. She now has her own life and I need to respect that.

¶ **[81]** *Project officer:* so it is different to track CW as opposed to tracking AW?

¶ **[82]** *JW:* Slightly, I still feel, because I feel more protective of AW than I do of CW I feel as though I am entitled to know what she is up to, more than with CW because I know CW is grown up, she'll be fine, she's running her own life. AW - there is still a bit of that mothering thing going on even though it is probably a bit, kind of past its prime. I feel more again entitled to be on that level, but with CW...

¶ **[83]** *Project officer:* and what about CR, what's the difference again. Is it different with CR

¶ **[84]** *JW:* I don't feel the same sense of entitlement, he is his own independent person I need to respect his space and his privacy and at the same time I trust him implicitly. If he tells me something I know it's the truth so I would never want to betray that by double checking on him. So I always felt a bit uncomfortable because I am not that sort of person but I was always very happy to know he was fine. So it is again that combination of, you know it is not totally something I am comfortable with doing but if you like the answer, that's great

¶ **[85]** *Project officer:* what do you think would happen if you didn't like the answer, if you felt something like, oh I didn't expect that, so your reaction would be rather so how would you feel towards the sisters or towards CR

¶ **[86]** *JW:* do you connect the two?

¶ **[87]** *Project officer:* how do ??????? to thinking, oh I am sure CR has an explanation or the system must be wrong. The answer the system gives you affects you in your thinking, is there something I don't know, something I am not aware of?

¶ **[88]** *JW:* I think if he said to me I spent the whole day in the office today and then later on I found out that he didn't, I can't imagine him doing anything that wouldn't be completely reasonable and normal and everything. So if he was out and about and forgot to tell me, it doesn't matter because he doesn't have to tell me everything. So he is perfectly free to do whatever he wants and I know that we are very open with each other so that if he does want to go to the park at lunch time and sit on his own or meet up with a whole load of mates and not tell me that's fine. So I would be interested and I would probably be a bit more alert to, I would hate to put myself in a position where I would then check up on him. I wouldn't want to, you know. It's like reading somebody else's texts on their mobile phone. Of course you are not going to find anything you don't like but now if you do find something you don't like now what do you do? Because I betrayed your trust I found out that you betrayed my trust. It just makes it so difficult then to deal with things, even if it is something totally innocuous well then you felt that you had to check up on my story and even if the checking up is totally, why is she home later. There could be a hundred rational responses or whatever. It would make me feel quite uncomfortable, like I would know that if CR was checking up on me and he said oh I thought you said you were in the office all day but you went to Stratford then I would probably go, oh yeah I forgot I went to meet a friend for lunch or whatever or anything. It's absolutely fine because I would probably tell him anyway but just knowing that there's again someone checking up on you, you go well don't you trust me so I think that there's a lot to do with that trust.

¶ **[89]** *Project officer:* did you feel that you would have to justify then every move you make so that there aren't any misunderstandings

¶ **[90]** *JW:* exactly

¶ **[91]** *Project officer:* otherwise I can't justify it or whatever

¶ **[92]** *Project officer:* and just looking at those maps, I go, what on earth was I doing, I just can't remember

¶ **[93]** *Project officer:* how did that make you feel

¶ **[94]** *JW:* very uncomfortable I must say. Like just going gosh, I was just wandering around having a nice chat, completely unaware that there was some satellite following me around all the time. There's that potential for somebody to use that information in a way that doesn't help anybody that just makes me feel wow

¶ **[95]** *Project officer:* is it the inferences that somebody might make about your behaviour or what you are up to that are sort of out of your control

¶ **[96]** it is the fact they were checking, people all come to conclusions but whether they were accurate or not isn't really the thing that makes me uncomfortable. I think it is having to justify, yeah being in a position where you have to, oh why did you take that route or why didn't you take that route, we said we were going to meet at 10 but I see at 10 o'clock you were still half way there, why did you leave late you know. Stuff that isn't really important suddenly becomes important because it is fixed and I think it makes stuff maybe significant when it is not.

¶ **[97]** *Project officer:* no that's very interesting. Then there is one on 31st January and you checked on CR. You were curious to see where he was. He was exactly where you thought he was, here you said here that you felt close, do you remember anything about that.

¶ **[98]** *JW:* Yes I do. It was a Sunday so I think what I'd done that evening, I think I had gone to meet a friend for lunch and it was an early dinner but I don't think I was with him and he'd been really, oh he'd been really sweet, because I'd been to the second farewell of the same friend who was leaving so it was just a girls night so he took me all the way there on the Sunday night rather than staying warm and close in bed. So he could have stayed here but instead he ventured out in the freezing cold so that we could spend some extra time together and I thought oh you're so cute, so yes, that was so nice of him. Yes I was just feeling very warm and cuddly at that moment.

¶ **[99]** *Project officer:* then obviously you checked on him after he went back home.

¶ **[100]** *JW:* Yes so I think he'd got back home and I think he actually sent me a text to say I'm home now and then I checked because I was thinking I can see him back in that little house.

¶ **[101]** *Project officer:* oh I see, was it one of those things like with AW or with your parents, saying everything is fine or was it?

¶ **[102]** *JW:* our house is really special for us, first time I've lived with somebody, it's like our little special place

¶ **[103]** *Project officer:* your little nest

¶ **[104]** *JW:* yeah, exactly, and I'm like really attached to it, very much in love with our little house and the fact that he is in the house is always really nice. We come home and it's like hello little house. So it is a very special place for us so knowing that he's in the house, if I come out of the tube station and see the lights on I know that he's in the house and it's always nice.

¶ **[105]** *Project officer:* so with this checking you went back in the nest with him for a moment

¶ **[106]** *JW:* I think so, it was kind of like I wish I was back with him, that was quite nice.

¶ **[107]** *Project officer:* so this was your spontaneous use of the...

¶ **[108]** *JW:* I'm not sure, I don't know, it probably was just from a timing point of view, it's just too much of a coincidence otherwise

¶ **[109]** *Project officer:* the feelings you are describing suggest to me this is not just automatically asked, there seems to be some sort of emotional involvement in this particular check. That's just a feeling I get

¶ **[110]** *JW:* I know at this point and just because we are being all cutesy he would welcome me knowing where he is. It would make him very happy to know that I was thinking about him so

¶ **[111]** *Project officer:* so you felt good about checking on him this time

¶ **[112]** *JW:* because I knew that he would be happy with me knowing that

¶ **[113]** *Project officer:* so there was some connection between you two at the time through this. There is another one on 31st January and that is CW, you wanted to feel closer to her and you say she was more or less where you expected her to be and you felt comfortable. And you said CW was sick

¶ **[114]** *JW:* she was sick, she wasn't well. Around this time, I said to CR when you fill these things in what do you put, how are you choosing, how are you doing this, and he said I don't know. I said when I put these things I put tested or requested and he said maybe you should do it properly. So I think around about this time I started

¶ **[115]** *Project officer:* being a bit more precise?

¶ **[116]** *JW:* Yes instead of just going look you asked me to I sort of said if I had done this on my own what would I have wanted to do to try and be a bit more helpful. So I think going forward there might be any more tested things

¶ **[117]** *Project officer:* don't worry that's fine. I just was wondering did you really want to feel closer to her when you did that?

¶ **[118]** *JW:* Yes, I knew she was unwell and I had spoken to her earlier in the day and she was quite stressed out she had exams on so it was, when we all used to live at home together and I was always studying and she was always studying and we always used to come into each others rooms and have a little moan and drink tea and generally say oh you'll be fine you're smart, no you're smart, OK. So it was kind of like I wish I was there with her because I could bring her tea and ?

¶ **[119]** *Project officer:* so now that she was sick, not a little sister but somebody that needs a little caring, so how did you feel about checking her at the time. Did that feel OK because she was ill?

¶ **[120]** *JW:* when you're sick you want people to look after you, you want them to be there but you know the next best thing I could do and I knew she was trying to study so I wasn't going to call her and disturb her. I'd sent her a text and then when this came through I thought oh that's quite nice.

¶ **[121]** *Project officer:* next one is 1st February, you checked on your dad, you said you were curious to see where he was, he was more or less where you expected him to be and you felt comforted and you put, chilling at home.

¶ **[122]** *Project officer:* where you the one chilling at home?

¶ **[123]** *JW:* I was the one chilling at home and I think dad was chilling at home. Sometimes in the evening him and my mum walk the dogs and my mum loves the dogs desperately and my dad is slightly less enamoured. So she loves dog walking and he is a bit reluctant so when I saw that he was home warm and cosy and stuff I thought, OK well maybe my mum has gone off on the dog walk and left him to snuggle at home or something.

¶ **[124]** *Project officer:* how did it feel to check on him in this case, was it in the category, it's OK, so-so, voyeurism

¶ **[125]** *JW:* so-so to OK, because he's my dad and it's always kind of OK with them

¶ **[126]** *Project officer:* yes as you said. The next one is 1st February that was you checked on AW, you said you were missing, you said she was exactly where you expected and you said you felt neutral, and you were chilling

¶ **[127]** *JW:* yes, I was just relaxing at home and AW, despite the fact that she is physically attached to her phone, is not very good at communicating with her sisters. So if I call her or give her a text it is a one in three if I get a reply and I'm sure it's being a teenager. You know 'bloody get off my back' and get on with your own life. So I don't talk to her as much as I'd like. I think I was actually a bit cross with her because I knew she was home, I knew she was sitting with the phone like that and I probably just tried to call her and she didn't answer again and I was probably like, fine, whatever.

¶ **[128]** *Project officer:* is it at that time you decided to check on her after she didn't answer your calls or get back to you

¶ **[129]** *JW:* again it was probably in response to a prompt coming at the right time

¶ **[130]** *Project officer:* now there is something else to say, usually you would only get one prompt and there is one at 7.30pm, one at 9.20pm so less than two hours apart. You wouldn't get 2 prompts one after the other like that. So sometimes you must have done this by yourself.

¶ **[131]** *JW:* I must have

¶ **[132]** *Project officer:* it is probably you didn't pay attention to. I could probably track it but it depends on whether you responded to the request which is not necessarily immediately after the prompt and I wouldn't send the request at this time because I would consider this a little late.

¶ **[133]** *JW:* I wish I'd taken notes so that I could remember more accurately what went on

¶ **[134]** *Project officer:* don't worry. I am mainly interested in how you felt in doing certain things. So how did you feel about checking on AW in this situation. She was sort of unretrievable, it sounds like

¶ **[135]** *JW:* yes a little bit and I know she is always on Facebook, always on Myspace. They are quite comfortable with letting everybody know where they are and who they are with all the time and it's not a problem for her. So I know she doesn't really care one way or the other whether people are checking up on her or not because everyone is always checking up on everybody. So I know for her if I am saying where are you, she's probably at that moment checking on 20 of her friends so you know for her it is not an issue which kind of makes it feel a bit better if I do check up on her.

¶ **[136]** *Project officer:* then 3rd February you checked on CW, you said you just thought of her, she was more or less where you expected her to be and you said you felt neutral about it. You said CW and RN on hols, so she had done her exams and they were doing something together, so how did you feel checking on her at the time?

¶ **[137]** *JW:* Are these confidential between family?

¶ **[138]** *Project officer:* I will not tell your family or CR what you have said and vice versa, I will not tell any of the others what you have said, it is confidential, unless you say I want this to be known, it's all confidential

¶ **[139]** *JW:* OK that's fine, that's good because my relationship with CW's boyfriend is quite fraught. I don't like him. It's unfortunate that's what it is, so I don't really like him. I'm glad she's with someone who's not an axe murderer, I'd prefer if she was doing other things. I knew she was hanging out with him for the weekend so when I looked, I was like whatever, I knew they would be together. It's not very nice.

¶ **[140]** *Project officer:* well that's OK, I'm not judging on these things I am just interested in what goes through your head. When you checked on her so you knew already that you would be with him, so how did that feel when you did the checking?

¶ **[141]** *JW:* Again it's that, you kind of really know the answer but you are hoping for a different one, wouldn't it be lovely if I saw she was magically up in [city].

¶ **[142]** *Project officer:* so she was in [city2] ?

¶ **[143]** *JW:* I think she was here in [city2], either she was in [city2] or she was in [city] one of the two because he lives here, she lives in [city]. So it would have been one or the other. February 3 was that?

¶ **[144]** *Project officer:* that would have been a Wednesday I believe

¶ **[145]** *JW:* OK I don't know

¶ **[146]** *Project officer:* don't worry. So you checked hoping that you would find something different to what you were expecting.

¶ **[147]** *JW:* Yes exactly

¶ **[148]** *Project officer:* and how did youfeel about checking with these thoughts in your mind, with this sort of disposition?

¶ **[149]** *JW:* You don't feel like a good person. I wasn't checking because I was hoping for everything to be fine, I was checking hoping that plans had fallen through and that's not very nice.

¶ **[150]** *Project officer:* OK

¶ **[151]** *JW:* I just oh dear. He's pretty much here to stay there are all lovey dovey and that's great but you always hope right? And I knew they were hanging out that's just what it is.

¶ **[152]** *Project officer:* let's turn the page shall we?

¶ **[153]** *Project officer:* now oh I must have asked you for this, I must have knew

¶ **[154]** *JW:* somehow you knew

¶ **[155]** *Project officer:* I get the feeling you wouldn't have done this spontaneously perhaps or perhaps yes. Let's find out, it's 4th February a Thursday you checked on RN. You say checking on them and he was more or less where you expected.

¶ **[156]** *JW:* I think he was at work

¶ **[157]** *Project officer:* so this is interesting from my point of view

¶ **[158]** *JW:* go for it I'll try and control myself

¶ **[159]** *Project officer:* you said you were checking because you don't trust him, so how do you feel given the fact that you don't trust him. How do you feel about that?

¶ **[160]** *JW:* Not like a good person. I feel like because he has shown in the past that he is not trustworthy that kind of he started it and it means I can check up on him and not feel too bad.

¶ **[161]** *Project officer:* so he has betrayed somebody's trust in the past and so therefore, I am putting it very strongly, he has no right to

¶ **[162]** *JW:* to be trusted

¶ **[163]** *Project officer:* which would mean that you don't need to check on him

¶ **[164]** *JW:* because I don't have the same level of respect for RN that I do for other people so when I check up on somebody it shows a lack of respect for their privacy.

¶ **[165]** *Project officer:* so he doesn't deserve that respect, is that what you feel?

¶ **[166]** *JW:* Yes kind off

¶ **[167]** *Project officer:* feel free to say these things if you feel that

¶ **[168]** *JW:* I still feel uncomfortable doing it but I don't feel as bad, it's different to checking up on CW, with CW I am starting on the basis that of course she's doing what she's doing what she's supposed to be doing and what she said she's doing so it is only going to confirm that. Whatever it shows, somehow it will confirm that. With RN it's highly unlikely he is doing what he tells you he's doing.

¶ **[169]** *Project officer:* highly unlikely

¶ **[170]** *JW:* highly. So anything you see is probably just going to confirm that. So when I saw he was at work, I though well obviously he's at work JW, get over yourself. It was sort of a little bit of he's not that bad. If it had shown him anywhere else I would have gone Hah! Even if he was just buying a pie at the corner shop I would have still felt Hah, my worst suspicions are confirmed.

¶ **[171]** *Project officer:* so you made no allowances for him

¶ **[172]** *JW:* yes I don't. I think how you feel about it definitely depends on your relationship with the other person. I definitely don't feel bad about checking up on him I almost feel like I a doing it to look after CW, cos if she's not going to look after herself then somebody must.

¶ **[173]** *Project officer:* OK. So you feel that she has put herself in a very vulnerable position because she is with RN and taking care of her by checking on him.

¶ **[174]** *JW:* Yes, which is not my place to do and I know that so I try and stay as far out of it as I possibly can because it is none of my business at all. But this just makes it a little bit too easy doesn't it. You're only human

¶ **[175]** *Project officer:* so who cares let's do it

¶ **[176]** *JW:* when it is easy to check and when someone doesn't know that you are checking it gives you that freedom

¶ **[177]** *Project officer:* oh we'll come to that later

¶ **[178]** *JW:* otherwise you might not do it

¶ **[179]** *Project officer:* on the 5th which was a Friday you were curious to see where AW was you said, your memory phrase was out to lunch with [friend] and you said she was more or less where you expected and you felt neutral about it. Anything new or different about this?

¶ **[180]** *JW:* No. I think she was at school, I'd just had lunch and yes it was just she's at school that's where she should be

¶ **[181]** *Project officer:* OK so nothing different

¶ **[182]** *Project officer:* and then on the 7th which would have been a Sunday you checked on AW again more or less where you expected and you felt pleased about it. You were clapping? That's your memory phrase, again anything different there?

¶ **[183]** *JW:* If I remember correctly she was at her boyfriend AD's house, AD's quite a sweet kid, he's really nice, so I thought she would probably be at home, she wasn't at home she was AD's which is always the next best guess. I think that's what it was

¶ **[184]** *Project officer:* OK then the next one was the 8th which was a Monday and you checked on your dad and you said you were curious to see where he was, he was exactly where you expected and you felt pleased about it.

¶ **[185]** *JW:* Work. I was dealing with a particularly difficult case at the time ?????? yeah, dad is at home I think and I thought OK it's quite he's probably chilling out, done for the day, yeah.

¶ **[186]** *Project officer:* then the next was on Tuesday 9th you checked on AW again so you remember something particular about this?

¶ **[187]** *JW:* Yes we had a great dinner with some [non-UK country] friends of ours, a couple and the guy is very tekkie and him and CR were talking about this a lot and going into it and really going into the possibilities and potential and whatever so I think this was a response to a request.

¶ **[188]** *Project officer:* so were you showing them how it works and how

¶ **[189]** *JW:* they came about 7.30 so I would have just responded

¶ **[190]** *Project officer:* so you had just done it

¶ **[191]** *JW:* yes, we were cooking and getting ready and stuff and then afterwards when we were talking about this thing. The guys loved it, he's an engineer so they were getting really involved

¶ **[192]** *Project officer:* and then on the last one on 11th February which was a Thursday and you checked, Ah. This is when I asked you to hide.

¶ **[193]** *JW:* It was great

¶ **[194]** *Project officer:* in fact that is when we switched and you become the person who is being tracked you made yourself invisible for 2 hours and you went back and hid yourself for longer. So you don't want to be found, to feel trapped down and you feel reassured about it. Tell me how does that feel then?

¶ **[195]** *JW:* I went for drinks with one of my best friends from [another country] who just moved here and we are very close and we have this, we talk about personal stuff so it's got a long history so there's always lots to catch up on and she's had quite a roller coaster the last few weeks. We hadn't seen each other so we met up and it's always like a very close personal things, silly girly gossip whispering you know all this stuff and I quite liked knowing that it was really private from everybody. Everyone knew I was having some drinks because I tell everyone and she tells everybody, but while we were chatting it was just us and I quite liked that because that's how we open up to each other.

¶ **[196]** *Project officer:* so the fact that you were able to hide yourself on that occasion made you feel more private than you would have been if somebody could have seen you on a map.

¶ **[197]** *JW:* Exactly. Every now and then I would be chatting to CW or CR and they would say what have you been doing and I say I went out to Waterloo station and they would say ah yes I saw you I tracked you. You are conscious that other people, out of love, out of interest, or positive things but they are kind of aware. Which is fine there is nothing inherently wrong about it but sometimes you do just want to close the curtains.

¶ **[198]** *Project officer:* so how did you feel when you started to get these notifications that people had been tracking you. Do you remember when you started getting those?

¶ **[199]** *JW:* Yes I remember going oh my goodness, I was in the office ????? Not surprised but not pleased, sort of it's to be expected and there's nothing wrong about it but it never made me feel pleased that someone had checked up on me because either you really know what I'm doing, so why don't you find something better to do or you don't know and if you want to know anything just give me a call.

¶ **[200]** *Project officer:* so they should ask you directly not go behind your back and looking

¶ **[201]** *JW:* it's almost like I give you permission to know these things about my life one of them is where I am so I was actually I thought I would feel more upset about it that I did

¶ **[202]** *Project officer:* it wasn't as bad?

¶ **[203]** *JW:* Yes maybe by the end I just got used to it as well. But I expected to feel upset about it or something, but at the end of the day no-one really cares as much as you think they will, it is just like well obviously she's at work

¶ **[204]** *Project officer:* now that you started to receive the notifications then knowing that they were tracking you in the same places before, did that make you feel different or was it the same?

¶ **[205]** *JW:* It was just being tracked

¶ **[206]** *Project officer:* so the place wasn't so much, it is just because they are on top of you

¶ **[207]** *JW:* the fact that they are checking

¶ **[208]** *Project officer:* when you were then checking on other people you must have been aware that other people must have been getting the feedback, so how did you feel about that?

¶ **[209]** *JW:* It is not my personality to do that at all so you almost want to have a little disclaimer at the end of it, really sorry, tell me if you don't want

¶ **[210]** *Project officer:* and you did say before if RN could see that you were checking you wouldn't

¶ **[211]** *JW:* I wouldn't want to do it, things are tense enough as it is, we put on a big smile for CW

¶ **[212]** *Project officer:* so it would have been worse with RN than for instance with AW, with AW knowing that you had checked on her

¶ **[213]** *JW:* yes, but RN would go, she doesn't trust and who the hell does she think she is, well I assume so

¶ **[214]** *Project officer:* OK that's your explanation.

¶ **[215]** *Project officer:* so look how many people have tracked you, this is a mistake at the end. This intruder is one of our researchers. I don't know why the system put it there. He might have been checking or testing something but this was at the beginning anyway. When you see all these people, you knew it of course, does it feel anything like

¶ **[216]** *JW:* it is quite strange to see it all on one bit of paper

¶ **[217]** *Project officer:* you can see it all, where you were when they checked on you

¶ **[218]** *JW:* it's very so much of interacting with other people is kind of not forgetting stuff, but things move on. Stuff it all gets a little fuzzy after a while when you come with, so I see on 3rd January I see that you were here, you go, gosh was that important should I have remembered that. It is just weird to see it all recorded there now for as long as the records last, potentially forever, you go gosh, now what do you do with that?

¶ **[219]** *Project officer:* but obviously

¶ **[220]** *JW:* that's potentially quite valuable information that now people have it. What? They are always going to kind ways of using it, which probably are going to cost me money. What happens next? It makes me feel a bit weird to see all the many checks. Thing is I really know my family loves me and always found me and wants to go Hi, so it is not as though when they check up on me I go what a relief or now I feel that they care about me that I didn't know before.

¶ **[221]** *Project officer:* let me see if I have forgotten anything. The last thing that I wanted to ask you was, this just concerns the form for the feedback, the notification which is that text. You got the text message. If you were to use such an application and if you were to get feedback would you prefer feedback to come in the form of a text message as it was in this case or in some other way like a little light on your phone or something.

¶ **[222]** *JW:* I think probably if you were someone who wanted to keep that information, keep track, if knowing that someone checked up is valuable for you then to have it in a text message where it is saved, then you can see the sequence that is quite interesting. If it was anything that didn't show that sequence of timeline it might not be as valuable.

¶ **[223]** *Project officer:* you would want to have a record of the history, like you had in the application as well as the realtime notification perhaps. Also a record of who tracked you when.

¶ **[224]** *JW:* I think personally for me I would probably not keep any records but that would be whatever form it came in, I probably wouldn't keep it, but I can see that if it does come in a text it would probably be more helpful.

¶ **[225]** *Project officer:* because there would be a record, in case

¶ **[226]** *JW:* in case you ever need to know.

¶ **[227]** *Project officer:* OK we are done!

B.5 Transcripts from interview with AW

¶ **[1]** *Project officer:* I'll go through the different events that the system picked up and then I'll ask you. This is the first thing you did so it was 10th January and you said you just looked up all your friends. So you just saw them, you have no expectations and you had neutral thoughts. You didn't put any memory phrase.

¶ **[2]** *AW:* I wasn't sure

¶ **[3]** *Project officer:* what to put

¶ **[4]** *AW:* no

¶ **[5]** *Project officer:* do you remember anything about this episode.

¶ **[6]** *AW:* I think it was the first time using it so I think I was experimenting and then just looking where everyone is, that's what I did, instead of just looking up a particular person, because I kind of knew where people should be and if I was more interested I would click on them. So usually I just looked at everyone on the map and I think there is a memory phrase was I think 'wanted to play around'.

¶ **[7]** *Project officer:* the next one you looked at GW and you just thought of him?

¶ **[8]** *AW:* I think my dad was around here ?????????? and he's either going to be at work, at home, so I knew where my dad would be so I was kind of, I wonder where he is.

¶ **[9]** *Project officer:* what made you want to check on him?

¶ **[10]** *AW:* Probably just sitting in a free lesson and I though I wonder where my dad is, because sometimes it comes up with the different distances and I was wondering well is that the distance from school. And then I was probably curious to see how far home is from school and then clicked on my dad because I knew he'd be at home.

¶ **[11]** *Project officer:* so you were more interested in hos distant home was from school?

¶ **[12]** *Project officer:* did you feel anything particular about this type of checking on your dad?

¶ **[13]** *AW:* Not particularly. I think most of the time because it was prompted, would you mind looking up this person, I know that everyone is going to be fine and I don't have the edge that I need to check on them as a way I probably would have in [another country]. So the feeling was kind of curious but not entirely curious, I wonder where my dad is, little bit thinking but not really worried.

¶ **[14]** *Project officer:* now the next one you looked at RN, same day, and do you remember anything about this.

¶ **[15]** *AW:* I think it was being prompted to ask, where is RN.

¶ **[16]** *AW:* Where these before?

¶ **[17]** *AW:* OK then probably just the distances again and probably just a bit curious about RN but I couldn't really pinpoint what it was. I knew where JW and CR was, I knew CW was in [city], I knew he was in [city2], but I was also just seeing

¶ **[18]** *Project officer:* so this was just when you started, you were starting to find where everyone

¶ **[19]** kind of was.

¶ **[20]** *Project officer:* again you looked again at all your friends on the same day, so was this similar?

¶ **[21]** *AW:* Similar. I think to just do the mapping, curious to see if people were home, if JW and CR were at home, something like that. That might have been CW and RN in [city2] or [city] together, I didn't know that they were there but I was pleased to see that everyone was together.

¶ **[22]** *Project officer:* then you looked at RN again on the same day

¶ **[23]** *AW:* I don't remember

¶ **[24]** *Project officer:* do you remember anything about that?

¶ **[25]** *AW:* Curious, oh to see if he was in [city] with CW or to see if she was in [city2] so I saw those two together, then I thought are they are they are his house or work or something. So I was probably a little curious to see if they were together or not.

¶ **[26]** *Project officer:* and then on the same day. You looked at CR.

¶ **[27]** *AW:* How late! I think I was just checking to see where they are, you know thinking are they are home, I think I was just playing around a bit, I was excited. New thing on my phone probably something like that.

¶ **[28]** *Project officer:* did you like having these applications?

¶ **[29]** *AW:* Yes I thought in the beginning it was really new and really different, yeah it was fine, it was quite weird but I didn't feel like people were watching me or anything. It helped a bit, my mother was trying to find me to pick me up from some friend's house, so I enjoyed it.

¶ **[30]** *Project officer:* you said it was a little bit weird, in what way?

¶ **[31]** *AW:* Well weird in that there was something that knew where I was all the time, it wasn't an invasion of privacy, I think it was quite a good idea but at the same time it was just a little bit scary that people know where you are all the time. But that is just a technology wow, nothing major.

¶ **[32]** *Project officer:* on the same day you checked on JW, so after checking on CR you checked on her.

¶ **[33]** *AW:* Probably just clicking to see if that is where their location was.

¶ **[34]** *Project officer:* so you must have seen that CR was at home I imagine and then you checked and you saw that JW was at home, so how did that feel?

¶ **[35]** *AW:* It was good, it was nice because they live together so I was pleased.

¶ **[36]** *Project officer:* now the next day. You ran out of minutes on the next day. You checked on RN, this is still before we started asking you.

¶ **[37]** *AW:* I think 8.02 I would be on the bus so I might be showing someone the application, I did that quite a bit I showed my friends, oh this is where my sister's boyfriend is and all my friend's know my sister's boyfriend so they were curious. I wasn't entirely sure where he worked so probably just checking was he still at home.

¶ **[38]** *Project officer:* so how did you feel about doing that, especially if you were with other people.

¶ **[39]** *AW:* I feel a little bit bad because I was checking on him quite a lot by the sounds of it. Probably because he is a new boyfriend and I am not entirely sure of him yet so probably looking him up for my sister. I am very curious about RN because I don't really know him yet so that is probably an element of these I was curious to see where he was, not so much checking on him because I don't trust but just curious because I don't know much about him. So is he working at this time, is he working at this time, bit stalkerish! I think it is just curiosity. I think my friends just enjoyed seeing it, I quite enjoyed knowing where everyone was all the time. In the beginning just because I don't know RN that well.

¶ **[40]** *Project officer:* so you were exploring his life

¶ **[41]** *AW:* yes

¶ **[42]** *Project officer:* next you looked up all your friends, that was the same day.

¶ **[43]** *AW:* I probably said if we just look everyone up on the map and see everyone and giving an example of looking them up and then looking at them on the map.

¶ **[44]** *Project officer:* so you were still with your friends and showing them the application.

¶ **[45]** *AW:* And sometimes my friends would ask to have a look and just press things.

¶ **[46]** *Project officer:* OK I will ask you about that later, that's interesting.

¶ **[47]** *Project officer:* then you looked at your mum.

¶ **[48]** *AW:* This is Tuesday, so 9.03 I was in form before school started probably just new friends showing them around the system. So just had a look at my mum to see what she was doing not really sure and then I was pleased to see where I thought she was.

¶ **[49]** *Project officer:* is it a way of keeping in touch, keeping close.

¶ **[50]** *AW:* Luckily I do live with my mum and dad so I never really feel that I need to keep close contact with them because I am here all the time. But for my sisters there's definitely, I really felt that I was missing them sometimes, I think I said that once or twice and I really want to feel closer and it did make me feel a bit closer knowing where they were.

¶ **[51]** *Project officer:* OK next. CR. This is Thursday, you thought of him, he was more or less where you expected and you felt pleased about it.

¶ **[52]** *AW:* I got a free, so I was probably just in my free period just probably where does CR work, as I am not entirely sure, looked at what he was doing and just knowing where he was, that was like that's cool.

¶ **[53]** *Project officer:* you looked at him because you like him?

¶ **[54]** *AW:* I do like him and he's also very much part of the family been with us for a very long time, so him and JW are like one person sometimes, just seeing where he is you feel a bit closer to him because I probably thought I hadn't seen them for a while.

¶ **[55]** *Project officer:* OK so was similar to what you do with your sister. And then you looked at your mum on the same day so do you remember anything about that?

¶ **[56]** *AW:* Probably the same thing just wondering where she is, my mum is quite interesting as she is always at B&Q or Ikea or she's out and about so probably a bit curious or I got bored.

¶ **[57]** *Project officer:* the way in which you checked the others was it driven more by what you were doing or what you thought they might be doing

¶ **[58]** *AW:* probably a mixture of both of them. CW, sometimes I was quite interested to see if she was at Uni or at home coz I know she was studying and it is usually I've got two minutes to spare and I just looked, or bored, tinker on my phone. Probably quite a bit more what I was doing, I wasn't really waiting for them to come and meet me anywhere, didn't really have that opportunity I was never waiting for someone and just seeing, oh are they close to meeting me or where I am. So I was never really that much involved with where they are in relation to me. It was usually what am I doing and where are they, kind of thing.

¶ **[59]** *Project officer:* for instance, when you check on somebody do you notice that after that as a result of having seen them, do you then move on to something else. Do you remember what you do after you checked, is there regularity or

¶ **[60]** *AW:* I think it is just usually interest, I like seeing everyone on the map and places on the map, just learning about that and then usually if I was at school I would just have a look, see where they are and get back on with some work. It didn't really pose any questions in my mind about why are they there, what are they doing unless it was something really odd. But I never saw anything that was really different.

¶ **[61]** *Project officer:* and do you know if there were any patterns in the way you were checking, you have described a little bit of that for instance, after breakfast, on the bus every day

¶ **[62]** *AW:* sometimes if I had a free, say Monday 1st and 2nd period and then 5th and 6th period I have a free so maybe if I was on a free it was probably just when I wasn't doing anything. Also to aid the experiment as well to give you guys some information to show that I was using it that type of thing. Then I got frees the rest of the week and sometimes in the morning I think I would just see oh what is everyone doing this morning. Are they going to work, are they at work, instead of sending a text, where is everyone this morning and get back on with my day.

¶ **[63]** *Project officer:* so it would be when you had a free moment.

¶ **[64]** *AW:* Probably the best time that I would do it when I was just not doing anything else.

¶ **[65]** *Project officer:* that is again 21st and you looked at your dad

¶ **[66]** *AW:* free again, I think I did this one by mistake. If the next one is CW

¶ **[67]** *Project officer:* no RN

¶ **[68]** there might be one of CW because I wasn't really missing my dad but there might have been a mistake by saying I was missing them and I thought I was thinking of CW and not my dad. And I was missing them I hadn't seen CW in ages so I wanted to feel closer like I said.

¶ **[69]** *Project officer:* how did you feel about looking at her

¶ **[70]** *AW:* I felt happier, because I have been to her house, I know [city] quite well now so I know if she is at a friend's house oh I can picture her friend's house or I can picture her house or the Uni so I felt a lot closer.

¶ **[71]** *Project officer:* it was OK you didn't feel there was anything strange about taking a peek in her life

¶ **[72]** *AW:* not particularly because I know, the thing about our family is we don't really hide things, we're quite open so it was never like oh I shouldn't be here, there are no secrets or anything. I didn't really feel guilty by taking a peek into her life because I would send a text saying what are you doing today. Also it is not that personal it is just a little dot on the map saying she is in this part of [city], it is not she is in her bedroom doing x,y and z it is just an overview. I didn't feel it was that intrusive.

¶ **[73]** *Project officer:* it wasn't detailed enough

¶ **[74]** *AW:* to kind of feel imposing. Then looking at RN again

¶ **[75]** *Project officer:* looking at RN again. Do you remember what's going on there

¶ **[76]** *AW:* probably got a free, probably finished my work or something. Curious is he at work, I know sometimes he has a later shift at work, so I wasn't really sure he might have been at work, he might not have been at work. Just a little bit curious but

I wasn't really too fussed where he was. I was kind of just learning the process, I do like him but I don't know what he's really about so

¶ **[77]** *Project officer:* so trying to get to know him a bit better this way

¶ **[78]** *AW:* there you go CW. This is the one I did by mistake. Well the one for my dad that's fine. Well she was probably studying at home and I knew what she'd be doing and I was really happy.

¶ **[79]** *Project officer:* this is CW again on the same day.

¶ **[80]** *AW:* Absolutely not. I think I was expecting her to be at home, she might have been in [city2] or she might have been at a friend's house studying, but I wasn't expecting that but I was pleased because if she was in [city2] she would be seeing RN which is really nice for her or if she is studying at a friend's house that's really nice for her. So I was pleased whatever she was doing but I expected her to be at home, I was just assuming, because she's a bit of a geek sometimes so I was just assuming just jumping to a conclusion almost but when I saw she wasn't there I was like oh wasn't where I expected but well.

¶ **[81]** *Project officer:* so how did you feel checking, was it OK?

¶ **[82]** *AW:* I felt fine about it because I know it wasn't, if we have this application in real life and they were checking up on me, I would feel a little bit like checked up on but because I know what we are doing and I know that people are checking up on me as well I don't feel that invaded almost. I think after a long time it can be a little bit like yes you know I'll be at home but in the beginning it's not that bad.

¶ **[83]** *Project officer:* and then you looked at all your friends, I was probably seeing if she was in [city2] or [city]. Oh no that's Sunday, probably just looking to see where everyone was just on the map, at a glance.

¶ **[84]** *Project officer:* and this is again

¶ **[85]** *AW:* just having a look where everyone was. I enjoy seeing all the little things on the map and learning all the things around so that was probably just what it was. Monday just in form class might have been showing someone as well, having a glance to see where everyone was.

¶ **[86]** *Project officer:* OK and then again

¶ **[87]** *AW:* Tuesday. I prefer to looking at everyone because I wasn't that curious to see what road they were on, just to see if they were in [city2] and then I could tell by the distances after a while that JW and CR were at work because they were far apart from each other. Or if mum and dad were some distance, because it is like 11 miles from my home and 19 miles to [city2] and 90 something to [city] so I started learning those and I could kind of assume what they were all doing.

¶ **[88]** *Project officer:* OK so you were reading all the distances and you knew what locations they were. At one go you could see everybody.

¶ **[89]** *AW:* It was quite good that way it was a lot easier, I could look at everyone and then I felt comfortable and then I was learning myself and I was quite happy with that I was learning distances and things, even though it was in miles, I knew where everyone is.

¶ **[90]** *Project officer:* that's good. So this is on Friday and CW, do you remember.

¶ **[91]** *AW:* Yes she was probably in [city], I really didn't have many expectations because she could have been in [city2] or [city] but I wasn't pinpointing her anywhere so I really wasn't banking on her being in [city] or [city2]. So I was happy because she was doing something, I can't remember what she was doing. I always felt quite happy when I looked at where my sisters where because it did make me feel a bit closer to them.

¶ **[92]** *Project officer:* OK this is all your friends again. Excited. This is on Friday 29th.

¶ **[93]** *AW:* That might have been close to RN's birthday, I think it was coz it was a Friday and I was getting excited I was going for the weekend. I was in a free lesson again, not exactly, CW might have been in [city2] or RN could have been in [city]. The reason I didn't look up JW as much is because she is with CR and I don't worry about JW as much for being lonely, whereas CW I might feel that she might feel a bit lonely and want her to feel closer because I know she is by herself a lot of the time whereas JW is with CR and I felt the need to look at CW a lot more just to see if she was with friends or in [city2] because then that would kind of change it.

¶ **[94]** *Project officer:* that would make you feel better to see she was with friends,OK

¶ **[95]** *Project officer:* then you looked at your dad on the 30th

¶ **[96]** *AW:* Saturday. Don't really remember that one, it might have been if I was going to [city2] this weekend I'm not entirely sure I think I might have gone to [city2] or something. Just kind of seeing if my parents are together because they are usually together, they are hardly ever far apart so kind of looking at one parent and kind of looking at both. I probably expected them to be at home, they were at home, didn't really expect that they would be anywhere else. They hadn't spoken about any weekend plans. So as I was probably out somewhere, seeing they were at home was cool.

¶ **[97]** *Project officer:* is it different to look up your sister's boyfriend, well you told me it was different between CR and RN, and then looking up your parents.

¶ **[98]** *AW:* It is different

¶ **[99]** *Project officer:* what are the differences you feel about it

¶ **[100]** *AW:* my mum and dad since I live with them, I don't feel the need to look at them that much because we live together and we know everything what we're going to do. That I didn't feel even intrusive because that I already know, a little bit curious but not even that curious just wondering. It's a bit different looking up RN because I don't really know him that well, I was getting to know him, whereas CR it's different from looking at my sisters because I don't really miss him as much but I do also miss him in his own way. But looking at my sisters is also different, we used to all live together until two years ago and seeing them having their own experiences and stuff is amazing I love seeing that they are doing things but also it does make me miss them. It also makes me feel closer to them, so looking them up is the most different from anyone else, I've got more feelings about that, whereas most other people I am quite neutral and it doesn't really bother me what they are dong. Whereas my sisters, it wouldn't bother me well if they somewhere really weird, I would be like why didn't you tell me, but as we are all quite close it was nice to see where they are all the time. It's just like living with them again.

¶ **[101]** *Project officer:* then you checked up on CW.

¶ **[102]** *AW:* I was prompted.

¶ **[103]** *Project officer:* that was the time when you started to get prompted and then you looked at all your friends again, like before, on the 3rd February

¶ **[104]** *Project officer:* is it different?

¶ **[105]** *AW:* Probably the same thing, might have been maybe my mum was shopping or maybe CW was travelling somewhere, something like that, I probably didn't have an exact idea where people would be.

¶ **[106]** *Project officer:* but you never found on this map

¶ **[107]** *AW:* something shocking. No. my family we kind of all do the same thing so I wasn't too worried ever or shocked to see where people. To be fair we only did move in this little area whether it was JW going to a friends house or CW going to Uni we stayed in this structured thing that we've all been doing for quite a while so if I saw that someone, say CW was in [non-UK country], I never saw anything that was shocking. Yeah 8.27 when I was on the bus just when I have a free moment.

¶ **[108]** *Project officer:* and then RN again on the 3rd. he gets checked a lot!

¶ **[109]** *AW:* I think I might have been prompted, maybe, I got prompted quite a lot, out of all my friends he was the one I got asked the most, it was probably him and CR and my dad, those were the three I got asked to look at and I was always just learning new things about RN. For his birthday, he's been my sister's boyfriend for a couple of months now but because he lives in [city2] and he doesn't really get weekends off so we don't really see him that much. So that might have been his house 8.27 he might have been working so I was just of curious about him still.

¶ **[110]** *Project officer:* so it feels all right to look him up?

¶ **[111]** *AW:* I don't know

¶ **[112]** *Project officer:* your sisters it seems really relaxed and with your parents the same but when it comes to RN you seem to feel unsure about whether it is okay.

¶ **[113]** *AW:* The thing is I don't know whether it is OK to look him up because I don't know him very well and I feel that maybe I shouldn't be looking at him. I should find out about him on his terms as opposed to looking him up all the time. The thing is I know he wouldn't be upset because of this experiment, but if this was a normal application and I was looking him up all the time I don't think it would be OK to look him up to the extent that I was and I should probably get to know him better on his own terms as opposed to me just spying. I get things completely wrong and jump to conclusions which I know I shouldn't be doing and I have no reason because he is a good person but I don't feel relaxed with him because I don't know him.

¶ **[114]** *AW:* I never did a memory phrase.

¶ **[115]** *Project officer:* that's OK. CW again on the same day, so you looked at RN and then you looked at CW

¶ **[116]** *AW:* it might have been sometimes he goes up in the week to visit her so I might have not been expecting him to have gone up. If she is in [city] she is exactly where I expect her to be and sometimes I look at the two together and sometimes I look at him and I won't really understand where he is but then in relation to CW I can tell where he is.

¶ **[117]** *Project officer:* OK. So that's dad again, do you remember anything about this?

¶ **[118]** *AW:* Yes, I was probably curious to see where my dad is, sometimes he goes to [city2] and sometimes he goes to other places for meetings and stuff and it said he was in this part round here and it could have been my house but I am not entirely sure. That's what that was about. I was quite relaxed about it I didn't have any particular feelings.

¶ **[119]** *Project officer:* and then your mum on the next day.

¶ **[120]** *AW:* My mum is probably at home, Thursday at 4pm. She probably would have been at home helping some builders out or something like that. Neutral because I knew she would be here, I wasn't really stressed to see that she was at home or to see where she was. When it comes to my parents I'm quite neutral about it because I have really no huge expectations of where they should be.

¶ **[121]** *Project officer:* and then you looked up all your friends. Have a vague idea, didn't know CW was in [city2].

¶ **[122]** *AW:* Vague idea, she was probably there with RN, pleased to see that everyone is together in [city2]. I was here with my parents, they were all together. Had a vague idea, if it was a Thursday she sometimes goes up to visit him so I assumed she might have been there but I didn't know she was going up to [city2].

¶ **[123]** *Project officer:* and then your mum the last one

¶ **[124]** *AW:* I got prompted.

¶ **[125]** *Project officer:* and again you felt neutral.

¶ **[126]** *Project officer:* OK so this is done. We did quite well. Now this is all what other people have done, looking at you. How does it feel seeing that all those people have been looking at you regularly.

¶ **[127]** *AW:* My mum

¶ **[128]** *Project officer:* your mum, JW

¶ **[129]** *AW:* I quite like that my sisters are looking me up because it means they are interested to see where I am. My mum is a bit mother hennish but I don't really mind. When you started getting the notifications it was a bit like, oh they are looking me up.

¶ **[130]** *Project officer:* that was the next thing I was going to ask you. When you started getting feedback, how did that feel?

¶ **[131]** *AW:* Well because you were doing it you couldn't be too much of a hypocrite - why are they looking me up, but at the same time you feel why do they think I'm here, do they know where I am. What are they filling out about me, do they have any expectations. You feel not paranoid but you do wonder a little bit, wonder what they are thinking about this or why they are looking me up. Towards the end I know sometimes it was because they were being prompted to and I don't have any obligations against them looking me up but it was sometimes like I really behaved myself for these couple of weeks to make sure I'm where I am all the time.

¶ **[132]** *Project officer:* if this was real life were you happy to receive a notification, to be made aware that somebody looked you up or would you prefer not to know.

¶ **[133]** *AW:* I would prefer not to know especially if I was doing the looking, if you have a boyfriend and want to see where they are quickly, I'd rather him not get a notification and I'd rather not know about it. I'd rather it be a bit more anonymous because

¶ **[134]** *Project officer:* there are two different questions, one is if you were the one doing the tracking, would the fact that the boyfriend got the notification stop you,

¶ **[135]** *AW:* yes it would

¶ **[136]** *Project officer:* that's the first. Second question is when you are being tracked would you be happy to know that somebody has tracked you.

¶ **[137]** *AW:* So me being in the position of being tracked, I wouldn't like to know

¶ **[138]** *Project officer:* you wouldn't want to know still

¶ **[139]** *AW:* no because if they want to track me that's fine but I'd rather not know about it because then I would start questioning it. If you don't really know you can't really question the ignorance of this.

¶ **[140]** *Project officer:* when you say start questioning this, is it like before you wonder why are they doing it, what are they thinking.

¶ **[141]** *AW:* It is like when you start getting the notifications you start wondering why are they looking me up, are they judging where I am but then they might not understand why I am here, they might not know why would I be in [city4]. They don't know. Maybe it might unnerve them to see that I am not at home or something like that. It's not only about me but how they are feeling about where I am. I would probably prefer not to get notifications. I feel that being tracked is, in the beginning it was quite exciting, something new, a new toy, but towards the end you start to feel almost oppressed and almost quite bad because your life is quite boring sometimes. Sometimes I was at home all the time, people looking me up must think, boring AW again. After a while you do feel a little bit intruded on, like I probably shouldn't have looked up people as much as I did. You see towards the end I started restraining myself and only doing the whole buddy check as opposed to individuals because I felt that maybe it was being a bit intrusive. But it is not terrible being tracked I wouldn't say it was the worst experience of my life it was fine really I didn't mind it. But sometimes you did feel there were questions.

¶ **[142]** *Project officer:* so you feel that if people made judgements on what you are doing when they don't know exactly, so they jump to conclusions.

¶ **[143]** *AW:* Yes exactly, I'm a person that really cares what other people think about me, always try and have the best public view of myself so people can see the best side of me so I would hate for people to jump to a conclusion. But the good thing, it is my family so they probably would never do something like that, but if this was real life and other people and my friends have this you wouldn't be able to have as much freedom. You would cause a little bit of stress for some people especially if you got a notification that one of your best friends looked you up, it would be like why are you looking me or your boyfriend up well don't you trust me, I think it could cause a little bit of a stir. I'm sure for JW and CW it must have been sometimes unnerving to see well why are they looking me up all the time, if it was real life in the beginning it was fine but towards the end it was the feeling heavier, more oppressive.

¶ **[144]** *Project officer:* I don't think that you ever used any privacy settings.

¶ **[145]** *AW:* No I didn't.

¶ **[146]** *Project officer:* is there any reason why?

¶ **[147]** *AW:* I really didn't feel the need to hide myself for a couple of hours. I only really stayed on the home and I didn't really know we had the privacy settings until my sister told me about it.

¶ **[148]** *Project officer:* your mum told me that you don't really check emails very much

¶ **[149]** *AW:* I'm really bad. So when I did find out about the privacy I didn't feel the need to.

¶ **[150]** *Project officer:* that's fair enough

¶ **[151]** *AW:* I also felt that if I did hide myself people would start asking where is she. Why are you hiding yourself what are you doing, which would probably be worse.

¶ **[152]** *Project officer:* the last thing I want to ask, we talked about the feedback notification. It doesn't sound like it but now having started to receive the feedback when obviously suppose you were somewhere around and you got the notification somebody tracked you and you may have thought, I was here last week somebody might have tracked me here last week, how did that make you feel. Because when you get the notifications you become aware that people have been tracking you?

¶ **[153]** *AW:* Not knowing that people are tracking you can also be quite negative. Sometimes getting a notification at least you know that they have looked you up once but that was also quite annoying almost, you get a notification and you're like yes I'm at school, I promise. But also the feeling that they were looking you up previously was a little bit unnerving but I can't say I was terribly stressed about it. I didn't really feel upset at being looked up.

¶ **[154]** *Project officer:* you got these text messages did you find the notification coming as a text message, did you find that OK, intrusive or?

¶ **[155]** *AW:* Sometimes I got quite a few, sometimes I would get 3 or 4 saying please look up this person because it was always RN and I felt quite bad

¶ **[156]** *Project officer:* no I mean the notification or when you got the notifications with the text message saying this person looked you up.

¶ **[157]** *AW:* It was fine, I had a text message it was open spy, a cute little funny name, it was fine. Sometimes I felt, I got quite a few from CR and my dad and I knew that my dad probably wasn't that curious to see where I was, I was probably in my room. But RN and CR I thought it was kind of nice sometimes, maybe they were doing the same thing checking up to see where I was but then it is not always a negative feeling but sometimes you can feel a little bit intruded on. But I never really did feel too

¶ **[158]** *Project officer:* when you received the notifications that somebody had checked on you, would you have preferred it in a different form, like a little light on your phone rather than a text message.

¶ **[159]** *AW:* You know how you can get a little notification on your phone that you can open but it doesn't got o your text messages, just you see it and it closes like a low battery sign or something like that, I think probably that might have been much better.

¶ **[160]** *Project officer:* better because you would find it less intrusive?

¶ **[161]** *AW:* Not even intrusive, just easier, you can see it once, I wasn't going to go and check them again and it is easier than going and deleting all my messages afterwards

¶ **[162]** *Project officer:* more practical

¶ **[163]** *AW:* yes more practical but deleting a message is not a taxing job, probably a little notification would be more practical. I thought text messaging was fine, I didn't have any particular stresses.

¶ **[164]** *Project officer:* I think we are finished

References

Adams, A. (1999a). The implications of users' privacy perception on communication and information privacy policies. In *In Proceedings of Telecommunications Policy Research Conference*, 65–67, TPRC Press, Washington DC.

Adams, A. (1999b). Users' perception of privacy in multimedia communication. In *CHI '99 extended abstracts on Human factors in computing systems*, 53–54, ACM, Pittsburgh, Pennsylvania.

Adams, A. (2000). Multimedia information changes the whole privacy ballgame. In *Proceedings of the tenth conference on Computers, freedom and privacy: challenging the assumptions*, 25–32, ACM, Toronto, Ontario, Canada.

Adams, A. & Sasse, M.A. (1999). Privacy issues in ubiquitous multimedia environments wake sleeping dogs, or let them lie. In *Proceedings of INTERACT 99*, 214–221J, Edinburgh.

Altman, I. (1975). *The environment and social behavior : privacy, personal space, territory, crowding*. Brooks/Cole Pub. Co., Monterey, Calif.

Anton, A.I. (1996). Goal-based requirements analysis. In *Requirements Engineering, 1996., Proceedings of the Second International Conference on*, 136–144.

Anton, A.I. & Earp, J.B. (2000). Strategies for developing policies and requirements for secure electronic commerce systems. In *1st ACM Workshop on Security and Privacy in E-Commerce (CCS 2000)*, unnumbered pages, North Carolina State University at Raleigh, Athens, Greece.

Anton, A.I., Qingfeng, H. & Baumer, D.L. (2004). Inside jetblue's privacy policy violations. *Security & Privacy, IEEE*, **2**, 12–18.

Arthur, C. (2011). iphone keeps record of everywhere you go. `http://www.guardian.co.uk/technology/2011/apr/20/iphone-tracking-prompts-privacy-fears`.

Bagues, S.A., Zeidler, A., Valdivielso, C.F. & Matias, I.R. (2007). Towards personal privacy control. In R. Meersman, Z. Tari & P. Herrero, eds., *On the Move to Meaningful Internet Systems 2007: OTM 2007 Workshops*, vol. 4806 of *Lecture Notes in Computer Science*, 886–895, Springer Berlin Heidelberg.

Barry, C.A. (1998). Choosing qualitative data analysis software: Atlas/ti and nudist compared. `http://www.socresonline.org.uk/3/3/4.html`, accessed: 2nd October 2012.

Barth, A., Datta, A., Mitchell, J.C. & Nissenbaum, H. (2006). Privacy and contextual integrity: framework and applications. In *IEEE Symposium on Security and Privacy, 2006*, 15–19, IEEE Computer Society, Berkeley/Oakland, CA.

Baxter, J. & Eyles, J. (1997). Evaluating qualitative research in social geography: Establishing rigour in interview analysis. *Transactions of the Institute of British Geographers*, **22**, 505–525.

Bellotti, V. (1996). What you don't know can hurt you: Privacy in collaborative computing.

Bellotti, V. & Sellen, A. (1993). Design for privacy in ubiquitous computing environments. In *Proceedings of the third conference on European Conference on Computer-Supported Cooperative Work*, 77–92, Kluwer Academic Publishers, Norwell, MA, USA.

Benisch, M., Kelley, P., Sadeh, N. & Cranor, L. (2010). Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs. *Personal and Ubiquitous Computing*, 1–16.

Beyer, H.R. & Holtzblatt, K. (1995). Apprenticing with the customer. *Commun. ACM*, **38**, 45–52.

Beyer, H.R. & Holtzblatt, K. (1998). *Contextual Design: Defining Customer-Centered Systems*. Morgan Kaufmann Publishers, San Francisco, CA, USA.

Bjrner, D. & Jones, C.B. (1978). *The Vienna Development Method: The Meta-Language*, vol. 61 of *Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, New York.

Booch, G. (1999). *Object-oriented analysis and design with applications*. Addison Wesley Longman, Reading, Massachusetts, 2nd edn.

Boyatzis, R.E. (1998). *Thematic analysis : coding as a process for transforming qualitative information*. Sage Publications, Thousand Oaks, CA.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, **3**, 77–101.

Breaux, T.D. & Anton, A.I. (2005a). Deriving semantic models from privacy policies. In *Sixth IEEE International Workshop on Policies for Distributed Systems and Networks, 2005.*, 67–76.

Breaux, T.D. & Anton, A.I. (2005b). Mining rule semantics to understand legislative compliance. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, 51–54, ACM, Alexandria, VA, USA.

Breaux, T.D. & Anton, A.I. (2008). Analyzing regulatory rules for privacy and security requirements. *Software Engineering, IEEE Transactions on*, **34**, 5–20.

Breaux, T.D., Vail, M.W. & Anton, A.I. (2006). Towards regulatory compliance: Extracting rights and obligations to align requirements with regulations. In *Requirements Engineering, 14th IEEE International Conference*, 49–58.

Castro, J., Kolp, M. & Mylopoulos, J. (2001). A requirements-driven development methodology. In *Proceedings of the 13th International Conference on Advanced Information Systems Engineering*, 108–123, Springer-Verlag.

Checkland, P. (1999). *Systems thinking, systems practice: includes a 30-year retrospective*. John Wiley & Sons.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. & Efstratiou, C. (2000). Developing a context-aware electronic tourist guide: some issues and experiences. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 17–24, ACM, The Hague, The Netherlands.

Chung, L. & do Prado Leite, J. (2009). On non-functional requirements in software engineering. In *Conceptual Modeling Foundations and Applications*, 363–379, Springer.

Clare, C. (2012). Caqdas: Deconstructing critiques, reconstructing expectations. In *RIBM 15th Annual Doctoral Symposium*, Manchester Metropolitan University.

Clarke, A.M. (2007). Ambient and pervasive technology: designing safeguards for vulnerable users. *Interactions*, **14**, 26–28.

Cockburn, A. (2001). *Writing effective use cases*. Addison-Wesley.

Consolvo, S. & Walker, M. (2003). Using the experience sampling method to evaluate ubicomp applications. *Pervasive Computing, IEEE*, **2**, 24–31.

Corbin, J. & Strauss, A. (2008). *Basics of Qualitative Research, Techniques and Procedures for Developing Grounded Theory*. Sage Publications, 3rd edn.

Cordy, J.R. (2006). The txl source transformation language. *Science of Computer Programming*, **61**, 190–210.

Cote, I., Heisel, M., Schmidt, H. & Hatebur, D. (2011). Uml4pf- a tool for problem-oriented requirements analysis. In *19th IEEE International on Requirements Engineering Conference (RE)*, 349–350.

Crabtree, B.F., Miller, W.F. & Miller, B.F.C.W.L. (1992). A template approach to text analysis: Developing and using codebooks. In *Doing qualitative research*, Research methods for primary care, Vol. 3., 93–109, Sage Publications, Inc, Thousand Oaks, CA, US.

Cranor, L.F. (1998). The platform for privacy preferences. *Communications of ACM*, **42**, 4855.

Cranor, L.F. (2003). P3p: making privacy policies more useful. *Security & Privacy, IEEE*, **1**, 50–55, 1540-7993.

Dardenne, A., van Lamsweerde, A. & Fickas, S. (1993). Goal-directed requirements acquisition. *Science of Computer Programming*, **20**, 3–50.

Davies, S.G. (1997). Re-engineering the right to privacy: how privacy has been transformed from a right to a commodity. In *Technology and privacy: the new landscape*, 143–165, MIT Press.

Davis, A.M. (1990). *Software requirements: analysis and specification*. Prentice Hall Press.

De Florio, V. & Deconinck, G. (2002). On some key requirements of mobile application software. In *Proceedings of 9th Annual IEEE International Conference and Workshop on the Engineering of Computer-Based Systems.*, 3–8.

DeMarco, T. (1978). *Structured Analysis and System Specification*. Yourdon Press, New York.

Deng, M., Wuyts, K., Scandariato, R., Preneel, B. & Joosen, W. (2011). A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. *Requirements Engineering*, **16**, 3–32.

Denning, D.E. (1976). A lattice model of secure information flow. *Communications of the ACM*, **19**, 236–243.

Dey, A.K. (2001). Understanding and using context. *Personal Ubiquitous Computing*, **5**, 4–7.

Dey, A.K., Abowd, G.D. & Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Hum.-Comput. Interact.*, **16**, 97–166.

Dourish, P. (2004). What we talk about when we talk about context. *Personal Ubiquitous Computing*, **8**, 19–30.

Dourish, P., Edwards, W.K., LaMarca, A., Lamping, J., Petersen, K., Salisbury, M., Terry, D.B. & Thornton, J. (2000). Extending document management systems with user-specific active properties. *ACM Transactions on Information Systems (TOIS)*, **18**, 140–170.

Eagle, N., Pentland, A. & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, **106**.

Earp, J.B., Anton, A.I., Aiman-Smith, L. & Stufflebeam, W.H. (2005). Examining internet privacy policies within the context of user privacy values. *Engineering Management, IEEE Transactions on*, **52**, 227–237.

Eason, K. (1988). *Information technology and organisational change*. Taylor & Francis/Hemisphere, Bristol, PA, USA.

Easterbrook, S., Singer, J., Storey, M.A. & Damian, D. (2008). Selecting empirical methods for software engineering research. In *Guide to Advanced Empirical Software Engineering*, 285–311, Springer London.

European Parliament, Council (1995). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML`.

European Parliament, Council (2002). Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications). `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:en:HTML`.

Federal Trade Commission (2010). Fair Information Practice Principles. `http://www.ftc.gov/reports/privacy3/fairinfo.shtm`.

Fereday, J. & Muir-Cochrane, E. (2008). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods*, **5**, 80–92.

FICKAS, S. & FEATHER, M.S. (1995). Requirements monitoring in dynamic environments. In *Requirements Engineering, 1995., Proceedings of the Second IEEE International Symposium on*, 140–147.

FINKELSTEIN, A., KRAMER, J., NUSEIBEH, B., FINKELSTEIN, L. & GOEDICKE, M. (1992). Viewpoints: A framework for integrating multiple perspectives in system development. *International Journal of Software Engineering and Knowledge Engineering*, **2**, 31–58.

GIBBS, G.R. (2007). *Analyzing Qualitative Data.*. SAGE Publications, Ltd, London, England.

GIERYN, T.F. (2000). A space for place in sociology. *Annual Review of Sociology*, **26**, 463–396.

GOGUEN, J.A. (1992). The dry and the wet. In *Proceedings of the IFIP TC8/WG8.1 Working Conference on Information System Concepts: Improving the Understanding*, 1–17, North-Holland Publishing Co.

GOGUEN, J.A. & LINDE, C. (1993). Techniques for requirements elicitation. In *Requirements Engineering, 1993., Proceedings of IEEE International Symposium on*, 152–164.

GOLAFSHANI, N. (2003). Understanding reliability and validity in qualitative research. *The qualitative report*, **8**, 597–607.

GOLDIN, L. & BERRY, D.M. (1994). Abstfinder, a prototype abstraction finder for natural language text for use in requirements elicitation: design, methodology, and evaluation. In *Requirements Engineering, 1994., Proceedings of the First International Conference on*, 84–93.

GSMA (2012). Privacy design guidelines for mobile application development. http://www.gsma.com/publicpolicy/privacy-design-guidelines-for-mobile-application-development, accessed: 5th February 2013.

HALEY, C.B., LANEY, R.C. & NUSEIBEH, B. (2004). Deriving security requirements from crosscutting threat descriptions. 976285.

HALEY, C.B., LANEY, R., MOFFETT, J.D. & NUSEIBEH, B. (2008). Security requirements engineering: A framework for representation and analysis. *Software Engineering, IEEE Transactions on*, **34**, 133–153.

HE, Q. & ANTON, A.I. (2003). A framework for modeling privacy requirements in role engineering. In *9th International Workshop on Requirements Engineering: Foundation for Software Quality, The 15th Conference on Advanced Information Systems Engineering (CAiSE '03)*, 115–124, Klagenfurt/Velden, Austria.

HEITMEYER, C.L., JEFFORDS, R.D. & LABAW, B.G. (1996). Automated consistency checking of requirements specifications. *ACM Trans. Softw. Eng. Methodol.*, **5**, 231–261.

HOLTZBLATT, K. (2005). Customer-centered design for mobile applications. *Personal Ubiquitous Computing*, **9**, 227–237.

HORMUTH, S.E. (1986). The sampling of experiences in situ. *Journal of Personality*, **54**, 262–293.

HUGHES, J., KING, V., RODDEN, T. & ANDERSEN, H. (1994). Moving out from the control room: ethnography in system design. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, 429–439, ACM, Chapel Hill, North Carolina, United States.

HUGHES, J., O'BRIEN, J., RODDEN, T., ROUNCEFIELD, M. & SOMMERVILLE, I. (1995). Presenting ethnography in the requirements process. In *Requirements Engineering, 1995., Proceedings of the Second IEEE International Symposium on*, 27–34.

IACHELLO, G. & HONG, J. (2007). End-user privacy in human-computer interaction. *Found. Trends Hum.-Comput. Interact.*, **1**, 1–137.

INFORMATION COMMISSIONER'S OFFICE (ICO), U. (2012). Key definitions of the data protection act. http://ico.gov.uk/for_organisations/data_protection/the_guide/key_definitions.aspx, accessed: 13th September 2012.

JACKSON, M. (1995a). Problems and requirements. In *Second IEEE International Symposium on Requirements Engineering*, vol. 0, 2–2.

Jackson, M. (1995b). *Software Requirements & Specifications - a lexicon of practice, principles and prejudices*. Addison-Wesley.

Jackson, M. (2001). *Problem frames: analyzing and structuring software development problems*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

Jackson, M. & Zave, P. (1993). Domain descriptions. In *Requirements Engineering, 1993., Proceedings of IEEE International Symposium on*, 56–64.

Jedrzejczyk, L., Price, B.A., Bandara, A.K. & Nuseibeh, B. (2010). On the impact of real-time feedback on users' behaviour in mobile location-sharing applications. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, SOUPS '10, 1–12, ACM, New York, NY, USA.

Jirotka, M. & Goguen, J.A. (1994). *Requirements engineering: social and technical issues*. Academic Press Professional, Inc., San Diego, CA, USA.

Jones, M. & Marsden, G. (2006). *Mobile Interaction Design*. John Wiley & Sons, Ltd., Chichester, England.

Kalloniatis, C., Kavakli, E. & Gritzalis, S. (2007). Using privacy process patterns for incorporating privacy requirements into the system design process. In *The Second International Conference on Availability, Reliability and Security, 2007. ARES 2007.*, 1009–1017.

Khalil, A. & Connelly, K. (2006). Context-aware telephony: privacy preferences and sharing patterns. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, 469–478, ACM, Banff, Alberta, Canada.

Kotonya, G. & Sommerville, I. (1997). *Requirements Engineering: Processes and Techniques*. John Wiley & Sons, Chichester, England.

Kramer, J., Ng, K., Potts, C. & Whitehead, K. (1988). Tool support for requirements analysis. *Software Engineering Journal*, **3**, 86–96.

Kumaraguru, P. & Cranor, L.F. (2005). Privacy indexes : a survey of westin's studies. Tech. Rep. 856, Institute for Software Research, Carnegie Mellon University.

Lahlou, S., Langheinrich, M. & Roeker, C. (2005). Privacy and trust issues with invisible computers. *Communication ACM*, **48**, 59–60.

Langheinrich, M. (2001). Privacy by design principles of privacy aware ubiquitous systems. In *Ubicomp 2001 Ubiquitous Computing*, vol. 2201/2001 of *Lecture Notes in Computer Science*, 273–291, Springer Berlin / Heidelberg.

Lederer, S., Dey, A.K. & Mankoff, J. (2002). A conceptual model and a metaphor of everyday privacy in ubiquitous computing environments. Tech. Rep. UCB/CSD-2-1188, University of California.

Lederer, S., Mankoff, J. & Dey, A.K. (2003). Who wants to know what when? privacy preference determinants in ubiquitous computing.

Lederer, S., Hong, I., Dey, A.K. & Landay, A. (2004). Personal privacy through understanding and action: five pitfalls for designers. *Personal Ubiquitous Comput.*, **8**, 440–454.

Lee, S.W. & Rine, D.C. (2004). Case study methodology designed research in software engineering methodology validation. In *In Proceedings of the Sixteenth International Conference on Software Engineering and Knowledge Engineering (SEKE 04*.

Lemos, R. (2001). Rental-car firm exceeding the privacy limit? `http://news.cnet.com/2100-1040-268747.html`.

Lessig, L. (1999). The architecture of privacy. *Vanderbilt Entertainment Law and Practice*, **1**, 63–65.

Lieblich, A., Tuval-Mashiach, R. & Zilber, T. (1998). *Narrative research: Reading, analysis, and interpretation*, vol. 47. Sage.

Macaulay, L. (1994). Cooperative requirements capture: Control room 2000. In *Requirements Engineering*, 67–85, Academic Press Professional, Inc., 184577.

Madan, A., Moturu, S.T., Lazer, D. & Pentland, A.S. (2010). Social sensing: obesity, unhealthy eating and exercise in face-to-face networks. In *Wireless Health 2010*, 104–110, ACM, San Diego, California.

Madsen, P., Mont, M.C. & Wilton, R. (2006). A privacy policy framework a position paper for the w3c workshop of privacy policy negotiation. Available at: `http://www.w3.org/2006/07/privacy-ws/papers/28-madsen-framework/`.

Mahatanankoon, P., Wen, H.J. & Lim, B. (2005). Consumer-based m-commerce: exploring consumer perception of mobile applications. *Computer Standards & Interfaces*, **27**, 347–357.

Maiden, N. & Rugg, G. (1996). Acre: selecting methods for requirements acquisition. *Software Engineering Journal*, **11**, 183–192.

Maiden, N., Gizikis, A. & Robertson, S. (2004). Provoking creativity: imagine what your requirements could be like. *Software, IEEE*, **21**, 68–75.

Mancini, C., Thomas, K., Rogers, Y., Price, B.A., Jedrzejczyk, L., Bandara, A.K., Joinson, A.N. & Nuseibeh, B. (2009). From spaces to places: emerging contexts in mobile privacy. In *Proceedings of the 11th international conference on Ubiquitous computing*, 1–10, ACM, Orlando, Florida, USA.

Mancini, C., Rogers, Y., Thomas, K., Joinson, A.N., Price, B.A., Bandara, A.K., Jedrzejczyk, L. & Nuseibeh, B. (2011). In the best families: Tracking and relationships. In *To appear in Proceedings of the 29th International Conference on Human Factors in Computing Systems, ACM CHI 2011*, ACM Press.

Margulis, S.T. (2003). On the status and contribution of westin's and altman's theories of privacy. *Journal of Social Issues*, **59**, 411–429.

Massey, A.K. & Anton, A.I. (2008). A requirements-based comparison of privacy taxonomies. In *Requirements Engineering and Law, 2008. RELAW '08.*, 1–5.

Meyer, B. (1985). On formalism in specifications. *Software, IEEE*, **2**, 6–26.

Miner, A.G., Glomb, T.M. & Hulin, C. (2005). Experience sampling mood and its correlates at work. *Journal of Occupational and Organizational Psychology*, **78**, 171–193.

Miyazaki, S., Mead, N. & Zhan, J. (2008). Computer-aided privacy requirements elicitation technique. In *Asia-Pacific Services Computing Conference, 2008. APSCC '08. IEEE*, 367–372.

Mouratidis, H. (2009). Secure tropos: An agent oriented software engineering methodology for the development of health and social care information systems. *International Journal of Computer Science and Security*, **3**, 241–271.

Nissenbaum, H. (2010). *Privacy in Context: Technology, Policy and the Integrity of Social Life*. Standford University Press, Standford, California.

NTIA (2013). Privacy multistakeholder process: Mobile application transparency. http://www.ntia.doc.gov/other-publication/2013/privacy-multistakeholder-process-mobile-application-transparency, accessed: 5th February 2013.

Nuseibeh, B. & Easterbrook, S. (2000). Requirements engineering: a roadmap. In *Proceedings of the Conference on The Future of Software Engineering*, 35–46, ACM, Limerick, Ireland.

OECD (2010). Oecd guidelines on the protection of privacy and transborder flows of personal data. `http://tinyurl.com/msr2nm4`.

Palen, L. & Dourish, P. (2003). Unpacking ”privacy” for a networked world. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 129–136, ACM, Ft. Lauderdale, Florida, USA.

Petronio, S.S. (2002). *Boundaries of privacy: dialectics of disclosure*. State University of New York Press.

Potter, J. & Wetherell, M. (2004). Discourse analysis. *Handbook of data analysis*, 607–624.

Prociow, P.A. & Crowe, J.A. (2010). Towards personalised ambient monitoring of mental health via mobile technologies. *Technology and Health Care*, **18**, 275–284.

Reubenstein, H.B. & Waters, R.C. (1991). The requirements apprentice: automated assistance for requirements acquisition. *IEEE Transactions on Software Engineering*, **17**, 226–240.

Rogers, Y. (2006). Moving on from weiser's vision of calm computing-engaging ubicomp experiences. *UbiComp 2006-Ubiquitous Computing*, 404–421.

ROMAN, G.C. (1985). A taxonomy of current issues in requirements engineering. *Computer*, **18**, 14–23.

RUMBAUGH, J., JACOBSON, I. & BOOCH, G. (1999). *The unified modeling language Reference manual*. Addison Wesley Longman, Inc., Reading, Massachusetts.

SALIFU, M., YIJUN, Y. & NUSEIBEH, B. (2007). Specifying monitoring and switching problems in context. In *Requirements Engineering Conference, 2007. RE '07. 15th IEEE International*, 211–220.

SCHILIT, B., ADAMS, N. & WANT, R. (1994). Context-aware computing applications. In *Workshop Proceedings on Mobile Computing Systems and Applications.*, 85–90.

SCHNEIER, B. (2008). The psychology of security. In *AFRICACRYPT*, vol. 5023, 50–79, Springer-Verlag.

SCHOEMAN, F. (1984). Privacy: Philosophical dimensions. *American Philosophical Quarterly*, **21**, 199–213.

SEYFF, N., GRAF, F., GRUENBACHER, P. & MAIDEN, N. (2008). Mobile discovery of requirements for context aware systems. In *Requirements Engineering Foundation for Software Quality*, 183–197, ACM.

SEYFF, N., GRAF, F., MAIDEN, N. & GRUENBACHER, P. (2009a). Scenarios in the wild experiences with a contextual requirements discovery method. In *Requirements Engineering Foundation for Software Quality*, 147–161, ACM.

SEYFF, N., MAIDEN, N., KARLSEN, K., LOCKERBIE, J., GRUENBACHER, P., GRAF, F. & NCUBE, C. (2009b). Exploring how to use scenarios to discover requirements. *Requirements Engineering*, **14**, 91–111.

SEYFF, N., GRAF, F. & MAIDEN, N. (2010). Using mobile RE tools to give end-users their own voice. In *Requirements Engineering Conference (RE), 2010 18th IEEE International*, 37–46.

SHARP, H., FINKELSTEIN, A. & GALAL, G. (1999). Stakeholder identification in the requirements engineering process. In *Database and Expert Systems Applications, 1999. Proceedings. Tenth International Workshop on*, 387–391.

SHILTON, K. (2009). Four billion little brothers?: privacy, mobile phones, and ubiquitous data collection. *Commun. ACM*, **52**, 48–53.

SOLOVE, D.J. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review*, **154**, 477.

SOLOVE, D.J. (2008). *Understanding Privacy*. Harvard University Press, London.

SOMMERVILLE, I. (2010). *Software Engineering*. Addison-Wesley, 9th edn.

SOUZA, E.S., LAPOUCHNIAN, A., ROBINSON, W.N. & MYLOPOULOS, J. (2011). Awareness requirements for adaptive systems. In *Proceedings of the 6th International Symposium on Software Engineering for Adaptive and Self-Managing Systems*, 60–69, ACM, Waikiki, Honolulu, HI, USA.

SPIEKERMANN, S. & CRANOR, L.F. (2009). Engineering privacy. *IEEE Transactions on Software Engineering*, **35**, 67–82.

SPIVEY, J.M. (1992). *The Z Notation: A reference manual*. Prentice Hall International Series in Computer Science., 2nd edn.

SUTCLIFFE, A., FICKAS, S. & SOHLBERG, M.M. (2005). Personal and contextual requirements engineering. In *13th IEEE International Conference on Requirements Engineering, 2005. Proceedings.*, 19–28.

SUTCLIFFE, A., FICKAS, S. & SOHLBERG, M. (2006). PC-RE: a method for personal and contextual requirements engineering with some experience. *Requirements Engineering*, **11**, 157–173.

TAMMINEN, S., OULASVIRTA, A., TOISKALLIO, K. & KANKAINEN, A. (2004). Understanding mobile contexts. *Personal Ubiquitous Comput.*, **8**, 135–143.

THOMAS, D.R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation*, **27**, 237–246.

THOMAS, P., GELLERSEN, H.W., BRUMITT, B., MEYERS, B., KRUMM, J., KERN, A. & SHAFER, S. (2000). Easyliving: Technologies for intelligent environments. In *Handheld and Ubiquitous Computing*, vol. 1927 of *Lecture Notes in Computer Science*, 97–119, Springer Berlin Heidelberg.

Thomasson, E. (2009). Facebook surfing while sick costs woman job.

Toulmin, S.E. (2003). *The uses of argument*. Cambridge University Press.

Tsai, J.Y., Kelley, P., Drielsma, P., Cranor, L.F., Hong, J. & Sadeh, N. (2009). Who's viewed you?: the impact of feedback in a mobile location-sharing application. In *Proceedings of the 27th international conference on Human factors in computing systems*, 2003–2012, ACM, Boston, MA, USA.

Tun, T.T., Bandara, A.K., Price, B.A., Yu, Y., Haley, C., Omoronyia, I. & Nuseibeh, B. (2012). Privacy arguments: analysing selective disclosure requirements for mobile applications. In *20th IEEE International Requirements Engineering Conference*, Chicago, Illinois.

UK Government (2013). Data Protection Act 1998. `http://www.legislation.gov.uk/ukpga/1998/29/contents`.

van Biljon, J., Kotze, P. & Renaud, K. (2008). Mobile phone usage of young adults: the impact of motivational factors. In *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat*, 57–64, ACM, Cairns, Australia.

van Lamsweerde, A. (2001). Goal-oriented requirements engineering: a guided tour. In *Proceedings of 5th IEEE International Symposium on Requirements Engineering, 2001.*, 249–262.

Van Lamsweerde, A. (2010). *Requirements Engineering: From System Goals to UML Models to Software Specifications*. John Wiley & Sons Ltd., England.

Van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy*. Suny Press.

Want, R., Hopper, A., Falc, V. & Gibbons, J. (1992). The active badge location system. *ACM Trans. Information Systems*, **10**, 91–102.

Warren, S.D. & Brandeis, L.D. (1890). The right to privacy. *Harward Law Review*, **4**, 193–220.

WEISER, M. (1991). The computer for the 21st century. *Scientific American*, **265**, 94–104.

WELSH, E. (2002). Dealing with data: Using nvivo in the qualitative data analysis process. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, **3**.

WESTIN, A.F. (1967). *Privacy and Freedom*. Atheneum, New York.

WHITTEN, A. & TYGAR, J.D. (1999). Why johnny can't encrypt: a usability evaluation of pgp 5.0. In *Proceedings of the 8th conference on USENIX Security Symposium - Volume 8*, 14–14, USENIX Association, Washington, D.C.

WHITTLE, J., SAWYER, P., BENCOMO, N., CHENG, B.H.C. & BRUEL, J.M. (2009). Relax: Incorporating uncertainty into the specification of self-adaptive systems. In *17th IEEE International Requirements Engineering Conference, 2009. RE '09.*, 79–88.

WIERINGA, R., MAIDEN, N., MEAD, N. & ROLLAND, C. (2006). Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. *Requirements Engineering*, **11**, 102–107.

WOOLDRIDGE, M. & CIANCARINI, P. (2001). Agent-oriented software engineering: the state of the art. In *First international workshop, AOSE 2000 on Agent-oriented software engineering*, 1–28, Springer-Verlag New York, Inc., Limerick, Ireland.

YARDLEY, L. (2008). Demonstrating validity in qualitative psychology. *Qualitative psychology: A practical guide to research methods*, **2**, 235–251.

YIN, R.K. (2009). *Case study research: Design and methods*, vol. 5. Sage.

YOUNG, R.R. (2004). *The Requirements Engineering Handbook*. Artech House, Inc.

YU, E. & CYSNEIROS, L.M. (2002). Designing for privacy and other competing requirements. In *2nd Symposium on Requirements Engineering for Information Security (SREIS'02)*, Raleigh, North Carolina.

YU, E.S.K. (1997). Towards modelling and reasoning support for early-phase requirements engineering. In *Requirements Engineering, 1997., Proceedings of the Third IEEE International Symposium on*, 226–235.

YU, Y., THEIN THAN, T., TEDESCHI, A., FRANQUEIRA, V.N.L. & NUSEIBEH, B. (2011). Openargue: Supporting argumentation to evolve secure software systems. In *19th IEEE International Requirements Engineering Conference (RE), 2011*, 351–352.

ZAVE, P. (1997). Classification of research efforts in requirements engineering. *ACM Comput. Surv.*, **29**, 315–321.

ZAVE, P. & JACKSON, M. (1997). Four dark corners of requirements engineering. *ACM Trans. Softw. Eng. Methodol.*, **6**, 1–30.

ZINS, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, **58**, 479–493.