Journal of Universal Computer Science, vol. 20, no. 14 (2014), 1964-1986
submitted: 24/6/13, accepted: 14/12/14, appeared: 1/12/14 © J.UCS

Context Classification Framework for Handset-based End User Studies

Tapio Soikkeli
(Aalto University, Department of Communications and Networking, Espoo, Finland
tapio.soikkeli@aalto.fi)

Juuso Karikoski
(Aalto University, Department of Communications and Networking, Espoo, Finland
juuso.karikoski@aalto.fi)

Heikki Hämmäinen
(Aalto University, Department of Communications and Networking, Espoo, Finland
heikki.hammainen@aalto.fi)

Abstract: Utilizing rich end user context information is viewed as one of the necessary approaches in developing more personalized mobile services and user experiences. The practical impact of end user context research and new opportunities in the field provided by emerging data collection methods such as handset-based measurements (i.e., collecting usage data directly from the end users' devices) have inspired new highly interesting large scale empirical context studies, but also brought quite diverse usage of the term context itself. Proper discussion and usage of context requires an unambiguous statement of how the term is understood in the particular case. On one hand the term should be positioned with the existing and commonly understood general definitions, but on the other hand it should also be acknowledged that especially an empirical research paper, or a context-aware service, can grasp only some specific aspects or elements of context. This paper proposes a context classification framework that aims to clarify the use of the term context in handset-based related end user studies. The framework is partly based on the experimental experience accumulated in our own handset panel studies. While helping researchers to plan context data acquisition and communicate and position the end user context elements used, the framework helps other stakeholders, such as application developers and service providers, to identify and utilize research and data most relevant for their particular needs. The paper also demonstrates the expressivity of the framework by examples.

Keywords: context, end user, handset-based measurements, framework, privacy
Categories: H.1.0, H.4.0, K.8.0

1 Introduction

End user context information is regarded as one of the key components in developing services and user experiences that adapt to mobile users' capabilities and needs. Mobile services and applications that are more personalized, and respond properly to the changing usage situations of an end user, i.e., are context-aware, require richer context information to work with and adjust to. It is, however, wise to take a step back and take a look at what is actually meant by the term context in these mobile and context-aware computing domains. In short, context is often regarded as something that characterizes the situation of a user. However, many differing definitions and interpretations of context [Dey, 01a], [Dourish, 04], [Tamminen, 04] exist. A single commercial application or an empirical research paper usually grasps only some specific aspects of context such as location, semantic place and/or surrounding people. Thus, on practical level and in the case of single studies it is often useful to define context quite case specifically as suggested in [Barkhuus, 03]. On the other hand, however, if a single context study wants to be recognized as a part of the whole context/context-aware research the case specific definition needs to be positioned with the more general definitions used in the domain. Previously, many empirical studies have, indeed, defined and used the term context very case specifically, but without putting much weight on positioning their definition with the more general definitions and categorizations. Examples include [Verkasalo, 08] using the term context while meaning semantic place, [Liang, 11] meaning physical place and task at hand, [Do, 11] meaning semantic place and proximity of other people, and [Böhmer, 11] meaning time of day and location.

One of the emerging methods of acquiring rich context information is through handset-based measurements, i.e., collecting data directly from the smartphones of opted-in users. In the recent years researchers, not only from purely technical fields, but also increasingly from the fields related to human behavior, have started to realize the potential of smartphones in acquiring rich behavioral data [Raento, 05]. As stated in [Miller, 12], smartphones are ubiquitous, unobtrusive, intimate, sensor-rich, computationally powerful, and remotely accessible. For these reasons smartphones are highly suitable for gathering precise and objective data on the real-world behavior and surroundings of millions of people. The potential of acquiring rich context information from the handsets is evident, but due to limitations in data collection platforms, devices at hand and other restrictions in the study/research setups, also the handset-based studies have to settle for a limited and case specific view on the broad context or situation of the end user – at least for now.

This paper proposes an end user context classification framework for handset-based studies. We consider such a framework necessary and possible because of the experimental experience accumulated in our panel studies (e.g., [Karikoski, 13], [Soikkeli, 11], [Verkasalo, 08]) and in those of others (e.g., [Do, 11], [Eagle, 06]). The motivation for the framework is to assist in taking into consideration the case specific nature of single handset-based context studies and at the same time position the study according to the more general definitions of context. The paper also demonstrates the expressivity of the framework by examples.

The paper is structured as follows: Chapter 2 provides background for handset-based measurements, and the concept of end user context. Chapter 3 presents the end user context classification framework, and Chapter 4 demonstrates the expressivity of the framework through examples. Chapter 5 concludes the paper and considers future work.

2 Background

This chapter provides the background on handset-based measurements as a data collection method. Then we take a look at some of the previous context definitions and categorizations in an effort to provide background for the term end user context and the framework. We also go through some general approaches for utilizing contextual data. Finally, we introduce some of the most central stakeholders in the handset-based context data “ecosystem”.

2.1 Handset-based measurements

Gathering diverse data from smartphones of individual users for research purposes is a relatively new approach. This is quite obvious since capable enough devices have not been around that long. At the moment, however, interest towards these so called handset-based measurements is increasing rapidly and multiple data collection efforts have been conducted around the world. One of the first platforms for collecting data from smartphones was the ContextPhone, developed by [Raento, 05]. It included a ContextLogger, which collected data related to the user’s location, device interaction, communication patterns and physical environment (namely, surrounding Buetooth devices). For example, [Eagle, 06] used the ContextPhone platform in their Reality Mining project where various aspects of human behavior were measured with the use of smartphones. The topic has been of interest also in the industry. [Nokia, 07] developed SmartPhone360 data collection software with an aim to study mobile service usage in Finland. SmartPhone360 tracked, e.g., the user’s location, communication patterns, app launches and installations, phone charging and network usage. Lately also open data collection platforms, such as the Funf Open Sensing Framework developed by MIT Media Lab [Aharony, 11], have emerged. The Funf framework provides ‘*an open source, reusable set of functionalities, enabling the collection, uploading, and configuration of a wide range of data types*’. A more comprehensive listing of different handset-based data collection platforms can be found, for example, in [Karikoski, 12].

Research utilizing handset-based data collected directly from smartphones include studies focusing mainly on the general usage of mobile services and applications (e.g., [Kekolahti, 13], [Shepard, 10], [Verkasalo, 08], [Xu, 09]); studies focusing on diversity of smartphone usage (e.g., [Falaki, 10], [Shepard, 10], [Soikkeli, 13]); studies focusing on contextual patterns of mobile service and application usage (e.g., [Do, 11], [Karikoski, 13], [Soikkeli, 13], [Verkasalo, 08]) and studies focusing on social networks and social relations among smartphone users (e.g., [Eagle, 09], [Karikoski, 10]).

As noted above, different implementations of handset-based measurement platforms/software exist. In principle, however, many of the platforms use very similar logic in the data collection process. Usually the first step is to distribute a data collection client to the smartphones of target group users. Different approaches have been utilized, ranging from small student/personnel panels inside a university to country-wide panels and all the way to user groups of whoever is interested in downloading the client from an app store or an equivalent. In the first two of the described cases a planned recruiting process is needed. After the client is installed into the users’ devices, configured data types are collected and periodically sent from the devices to a database. Some data collection platforms provide additional features that allow, for example, sending pop-up questionnaires and instructions to the users during the data collection. At the end, the usage of the data is dictated by contracts between the stakeholders participating in the data collection.

Table 1 gives an example of the different data types that can be collected with handset-based measurements. This particular example is based on the capabilities existing at the moment in the Funf framework [Aharony, 11]. Number of data types and the level of detail in the actual data are limited by the capabilities of the device the data collection client is installed into. However, in principle the most advanced devices are able to produce very rich data without any additional user participation.

Data Type	**Details**
Positioning	GPS (Global Positioning System), Bluetooth, WiFi, Cell
Social	Contact, Call Log, SMS (Short Message Service)
Motion (sensors)	Accelerometer, Gravity, Gyroscope, Orientation, Rotation
Environment (sensors)	Light, Proximity, Magnetic Field, Pressure, Temperature
Device	Battery, Hardware Info, Time Offset, Telephony
Device Interaction	Running Applications, Applications, Screen, Browser Bookmarks, Browser Searches, Videos, Audio Files, Images

Table 1: Example of handset-based data types (based on the Funf framework [Aharony, 11])

2.2 Understanding and utilizing end user context

2.2.1 Context definition

The word context has its roots in the Latin word contextus (con- ‘together’ + textere ‘to weave’). In the modern language context is: ‘*The circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood.*’ For example, in linguistics context is: ‘*The parts of something written or spoken that immediately precede and follow a word or passage and clarify its meaning.*’ [Oxford Dictionaries, 14] The linguistic definition of context was brought to the realm of computer science by computational linguistics. It is an important component in the problem of word sense disambiguation, which is highly relevant, e.g., in machine translation and search engine development [Miller, 95]. More recently, the emergence of more portable and capable computing devices has induced the realization that also the computer systems themselves need to be increasingly sensitive to their context [Lieberman, 00]. The advent of the fields of ubiquitous computing [Weiser, 93] and context-awareness [Schilit, 94] brought also the user and her context more tightly into the discussion.

One of the most widely accepted definitions of context, related to ubiquitous computing and context-awareness, is provided by [Dey 01a]. It states that: ‘*Context is any information that can be used to characterize 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*.’ [Dourish, 04] argues that Dey’s definition, among many others in the context-awareness domain, consider context as a representational problem. Dourish himself regards context as an interactional problem where context is a relational (something may or may not be contextually relevant to particular activity) and occasioned (i.e., not stable) property, its scope of features is defined dynamically and it arises from activity. Based on these properties Dourish wants to avoid viewing context as information that can be somehow encoded and represented. Obviously, the whole situation cannot be encoded, but if we are interested in some aspects of it in practical applications, the representational perspective is required. Dey's definition accepts the limited view on context and does not, in fact, rule out any dynamic changes in the represented features of context, e.g., over time or activity. In this paper we follow Dey's definition of context, and because the entity we are interested in is the mobile end user, we use a term *end user context*.

2.2.2 Context categorization

In addition to different definitions of context, existing context-aware computing literature has put some effort in categorizing or classifying context. [Dey, 01b] introduces four categories or characteristics of context information – *identity, location*, *status* (or *activity*), and *time*. *Identity* refers basically to a unique identifier assigned to an entity. *Location* means more than just coordinates; it could include, e.g., orientation, elevation or places. *Status* (or *activity*) refers to the characteristics of the entity. *Time* is used normally in conjunction with other context information to provide a timestamp or time period for the other context information. [Schmidt, 99] proposes a working model of context for context-aware computing. They divide context into two high level categories, namely *Human Factors* and *Physical Environment*. *Human Factors* are divided further into *User*, *Social Environment*, and *Task*, whereas *Physical Environment* is divided into *Conditions*, *Infrastructure*, and *Location*. These features can be then divided even further. For example, one element of *Conditions* can be light and one element of light can be wavelength. In the considerations of [Schmidt, 99], time adds a historical dimension as context history is regarded important for approximation of a given situation or environment. [Xu, 09] use a similar high-level division of context by dividing it into *Personal Context* (profile, emotion, mobility, social state, and intention/task) and *Environmental Context* (where, with whom, what resource, temporal and physical condition). [Hong, 07] divides context into *Computing context* referring to hardware configurations used, *User context* representing all the human factors and *Physical context* referring to the other information provided by the real-world environment. [Henricksen, 04] classifies context based on the principal sources of context information. This leads to three context types, namely *Profiled*, *Sensed*, and *Derived* contexts. *Profiled* context refers to user-supplied context information, *Sensed* to context information acquired from physical and logical sensors, and *Derived* to context information that is inferred from "base" context information (e.g., from *Profiled* and/or *Sensed*) by using some derivation process.

2.2.3 Context utilization

Contextual information has no value if it is not used for gaining valuable insight and then acting upon it. [Perera, 14] presents a context lifecycle model for contextual data in context-aware software systems. The context lifecycle includes: (1) *context acquisition*, (2) *context modeling*, (3) *context reasoning*, and (4) *context dissemination* (e.g., for an application to act upon). As the context definitions and categorizations provide conceptual tools for understanding and representing context, the context lifecycle model takes a firmer stance on practical context utilization. *Context acquisition* phase covers the techniques for acquiring raw contextual data. These might include considerations on different acquisition processes (e.g., the profiled, sensed, derived approach by [Henricksen, 04] mentioned above), data sources (e.g., directly from the device or through some mediator service), different types of sensors to be used (physical, virtual or logical), and data collection frequencies.

Context modeling considers the techniques for representing context. According to [Perera, 14], in a context modeling process context information is defined as attributes, characteristics, relationships with other context information, quality-of-context attributes and queries for context requests. Incoming context information is then accumulated in the defined format and made available for further use when required. The defined formats and parameters in them, however, vary widely from solution to solution. The research domain of ubiquitous user modeling [Berkovsky, 09] aims at reducing this variation. In particular, it has focused on developing (i) generalized models for semantically standardized formats and parameters, and (ii) mediation and hybridization techniques to overcome the heterogeneity in data and modelling techniques. The most used context modeling techniques currently in use, according to [Perera, 14], are key-value modeling, markup scheme modeling, graphical modeling, object based modeling, logic based modeling and ontology based modeling. *Context reasoning* refers to techniques for inferring new, higher level, knowledge from the available context information. [Perera, 14] classifies the reasoning techniques into six categories: supervised learning, unsupervised learning, rules, fuzzy logic, ontological reasoning, and probabilistic reasoning. Finally, *context dissemination* refers to techniques for delivering the context information. Some of the techniques include query, i.e., sending the information upon a formulated request, and subscription, i.e., sending the information upon agreed rules, e.g., periodically or after a certain event.

2.3 Stakeholders

Handset-based context data are seen as a valuable asset in the mobile domain. It can be used to gain useful behavioral information about users, customers, consumers, i.e., about human beings, and to build better, more adaptive and more personalized mobile services. Only the perspectives might vary depending on which stakeholders are asked. Below we describe the most notable stakeholders in the handset-based context data “ecosystem”.

2.3.1 End user

The end user and her device are in the center of the handset-based context data “ecosystem”. The end user is the primary data subject. The data are collected from the user’s device and these data are assumed to describe the situation of the user. In principle the user needs an incentive to participate in this type of data gathering. In an academic setting the benefits for the end user might be monetary compensation, vouchers, “free” devices, or just the pure joy of contributing to academic research. In some service type of settings (e.g., “regular” smartphone apps collecting various data) the end user trades her data for a service. Often the service providers state that the data are used for better services and better user experiences (e.g., personalization). The risks for the end user relate to data privacy issues.

2.3.2 Other users

As described above, the data are collected from the primary data subject's device. However, these data might include information that can risk also the privacy of other users. Bluetooth traces, photos, audio or video might exist which can be linked, e.g., to location information. Also calling and messaging information might exist which by definition involves also other users than only the primary data subject. On the other hand, as user behavior is understood better through handset-based research and better services are provided, all users, including not only the current but also future users, will benefit. Users who manage to avoid disclosing handset-based data can get the benefits without the costs associated to such data disclosing and thus be free riders in this sense.

2.3.3 Universities and research institutes

Universities and research institutes are interested mainly in academic research around handset-based context data. Initially the handset-based research was quite technology oriented; developing handset-based data collection methodologies and measurement platforms. More recently the possibilities provided by such data have awoken also social science oriented researchers. The benefits of this kind of approach include, for example, continuous measurements in the real world, large scale datasets and objectivity (at least in principle) and unobtrusiveness compared to traditional surveys and interviews. The potential risks revolve around acquiring suitable technical and data analysis skills, acquiring representative datasets (e.g., smartphones required by participants), low control over participants, handling the data securely and taking the privacy and anonymity aspects into account.

2.3.4 Device vendors, operating system vendors, mobile service providers, application developers

Here we bundle together industry actors that produce and develop products and/or services through which the actors are usually able to gain access to handset-based data. Modern handsets, mobile operating systems, services and applications can by default have such features that enable collection and sending rich behavioral data back to the actors themselves. The benefits of this include data driven approaches to develop better products, services and user experiences to attract new customers and lock in the old ones. Also using the data for marketing purposes or selling to third parties is quite a common practice. The risks relate especially to privacy issues. Nowadays the companies handle these mainly with terms and conditions contracts, but if the observed actions are deemed unfair or unreasonable the customers might vote with their feet.

2.3.5 Market research companies

Some market research companies[1] exist that have put effort in capturing the mobile device market, mobile media consumption and mobile user behavior, at least partially, through handset-based measurements. These kinds of companies have usually proprietary handset-based measurement platforms and established channels for participant recruitment. The data are collected and analyzed to gain market insight information to be sold to advertisers, marketers and in general to those who want to utilize information of mobile consumer behavior in their business. The benefits include direct and continuous data from the users and services under interest, large scale or targeted datasets depending on the purpose, and relatively easy management of user panels when in place. The risks relate to developing the technical platforms, acquiring users for the panels, and data security and privacy issues.

2.3.6 Legislator

The legislator is not directly involved in collection, analysis or usage of handset-based context data, but it needs to be familiar with the pros and cons of such efforts. Handset-based context data are very privacy sensitive. The legislator has to balance between social and economic benefits and the risks associated to possible privacy infringements arising from these kinds of data and related analysis. Finding this balance requires also extensive cooperation between all of the stakeholders.

3 End user context classification framework

In this chapter we propose an end user context classification framework for handset-based studies. The framework attempts to take into consideration the case specific nature of single handset-based measurements related context studies and operational context-awareness efforts, but also helps in positioning the case specific contexts with the broader meaning of context. The work towards the framework is essentially based on the traditional analysis-synthesis approach [Ritchey, 91]. The analysis leans on experience from our own handset-based studies and on literature review of the domain. It examines characteristics of and linkages between different end user context elements deemed identifiable in a handset-based study setting. In its essence the synthesis is constructing a structured whole from the individual context elements in order to represent the end user context.

We start the positioning effort from the widely used context definition of [Dey, 01a] [see Chapter 2]. In the case of handset-based measurements the entities whose context can somehow be described are the handset or smartphone itself and the end user of the smartphone. Our main interest is the end user and thus the framework itself is positioned so that it considers *any information available through handset-based measurements that can be used to characterize the situation of an end user*. It should be acknowledged that the end user's handset works as a proxy in determining the end user's situation. Given the ubiquitous and intimate nature of the modern smartphone, we believe this assumption is justified in the majority of cases.

[1] For example Arbitron Inc. www.arbitron.com, comScore Inc. www.comscore.com and Nielsen Company www.nielsen.com are such market research companies.

The proposed framework is depicted in Figure 1. The framework is based on and combines the context categorization approaches of [Schmidt, 99] and [Henricksen, 04]. The former is one of the first works that articulated clearly the distinction between context related to human factors and context related to physical environment. This is now quite a commonly used distinction and integrating it into the framework grounds the framework well with previous work on context/context-awareness. As explained in the previous chapter, the latter context categorization approach is interested in the context information sources. The handset-based measurement method enables a diverse set of context information sources. We feel that the division to profiled, sensed and derived contexts suits well the nature of the handset-based measurements. Usually a handset-based data collection effort includes a pre and/or post questionnaires assigned to the users, and in addition many platforms provide a possibility for pop-up questionnaires (i.e., experience sampling) to the users' devices in order to collect profiled context information. The handsets are sensor-rich devices and thus very suitable for collecting sensed context information. Finally, many research efforts have tried to analyze and combine data from different sensors and user provided data to come up with derived context information. Rather than categorizing context, [Van Bunningen, 05] classifies context categorization schemes into two categories: conceptual and operational. Conceptual categorizations are based on the meaning and conceptual relationships between the context, whereas operational categorizations are based on how context were acquired, modeled, and treated.

Conceptual perspective

Operational perspective

	Personal Context			**Environmental Context**		
	User	**Social Environment**	**Activity/Task**	**Conditions**	**Infrastructure**	**Location**
Manually Provided Context	identity age gender nationality	phone contacts	meeting, etc. (from calendar entries)	visible conditions (from user taken photos, videos)	device operating system	semantic place (from user tags) place of an event (from calendar)
Sensed Context	heart rate body temp. blood pressure	surrounding devices noise level/voices	calling messaging using apps	temperature noise level pressure light level	battery level signal strengths access alternatives	coordinates network cell altitude velocity, acceleration
Derived Context	at sleep/awake health condition physical state mental state	surrounding people co-location social network	commuting studying shopping	weather day/night summer/winter	device orientation network quality device operating range (in time, in distance)	semantic place indoors/outdoors stationary/moving transport mode

Time

Figure 1: End user context classification framework and several example context elements placed on the framework

As can be seen from Figure 1, the main construct of the framework is a two dimensional matrix where, inspired by [Van Bunningen, 05], we have *Conceptual perspective* on the horizontal axis and *Operational perspective* on the vertical axis. Following [Schmidt, 99] the *Operational perspective* of the end user context is a combination of *Personal* and *Environmental* contexts. *Personal* context includes information about the end *User*, the end user's *Social Environment* and the end user's *Activities*/*Tasks*. *Environmental* context includes information about the physical *Conditions* and computational *Infrastructure* of the environment the end user is in and information about the end user's *Location*. Similar to [Dey, 01b], information about the location is deemed to be more than just a point in a two-dimensional space.

Following [Henricksen, 04], from the *Operational perspective* the end user context is a combination of *Manually Provided*, *Sensed*, and *Derived* contexts. Related to handset-based measurements, *Manually Provided* context information can be collected from paper- or web-based questionnaires assigned to participating users, so called pop-up questionnaires pushed to the users' handsets, or from applications such as calendar or phonebook which have information added by the user. *Manually Provided* context information requires an extra effort to collect and will have inaccuracies due to human errors. *Sensed* context information can be acquired from the data provided by different sensors on the handset, such as coordinates from GPS, acceleration from accelerometer or a list of surrounding WiFi beacons sensed by the device's WiFi antenna. *Sensed* context information is relatively easy to collect after the data collection platform is up and running. Inaccuracies result, e.g., from sensor errors and failures and network disconnections. *Derived* context information can be acquired by applying different derivation methods to the raw sensed and/or manually provided context information. For example, [Reddy, 10] uses a combination of accelerometer and GPS data and a decision tree classification approach to derive different transport modes (still, walking, running, biking, motor vehicle) of users. Another example is inferring semantic place information (e.g., at home or at work), by using mobile network cell ID and WiFi data and a heuristic place detection algorithm [Soikkeli, 11] or cell ID data with an additional user input [Bayir, 10]. *Derived* context information is probably the most difficult to produce since it requires the development and usage of some derivation method. Inaccuracies rise not only from inaccuracies of the underlying raw data, but also from the derivation process itself.

The framework has also a third dimension composed solely of one context category, *Time*. The whole situation of the end user, including all elements in the framework, can be described as a snapshot in time or as a time period. Nature of some of the elements is more dynamic than that of others and their state changes more frequently. As handset-based data accumulates during the data collection period, the time dimension provides also end user context history. In principle the time dimension can be thought to consist of the past (context history), the present (current context), and the future (forecasted context based on the past and present context information). In practice every handset-based data point is accompanied by a timestamp. Out-dated or incorrectly time-stamped data can cause inaccuracies in context inference.

To put shortly, in the proposed end user context classification framework a single context element is positioned based on its nature related to the (i) *Conceptual* and (ii) *Operational* (i.e., information source) perspectives of end user context. Figure 1 shows examples of different end user context elements that can be acquired already or quite possibly in the foreseeable future with handset-based measurements. The elements are positioned according to the framework. It is possible that data of namely the same context element are pursued through different sources. For instance, semantic places can be derived or asked directly from the user. In some cases an element can have also a personal and environmental aspect at the same time. For example, noise level can describe the environmental conditions, as well as the social group dynamics related to the end user. Objectives of a practical use case determine then the final position.

The proposed framework attempts to clarify the use of the term end user context in handset-based studies and bring the term to the practical level. It leans against the most used general definitions of context related to context-awareness research, but also takes into consideration the operational and practical side of acquiring context information from handset-based data. Unlike the previous approaches, our framework brings together the conceptual and operational perspectives of end user context into a coherent whole. While the conceptual perspective facilitates understanding the conceptual relationships within end user context, the operational perspective permits us to assess the challenges and opportunities related to the data acquisition and utilization. For example, the quality, cost and usability of contextual information depends on the acquisition technique. The effort of acquisition, subsequent data/information processing and related privacy issues might vary considerably. By integrating the perspectives, it is possible to model end user context and its utilization more accurately.

4 Expressivity of the framework

In this chapter we showcase the expressivity of the end user context framework through three example types: (i) assessment of end user privacy, (ii) view on the mobile end user context research domain through the framework, and (iii) two individual handset-based study cases positioned on the framework.

4.1 Privacy assessment

End user context in mobile services is one of the most privacy sensitive issues in the evolving information society and a potentially large legal problem. The same properties that make the data collected from smartphones a very interesting research subject, raise many problems about privacy. Information privacy is inherently bound with (i) control, i.e., who controls what information [Ackerman, 01], and (ii) users' concerns for privacy depend on what information they are asked to give up [Ackerman, 99]. The end user context classification framework is very suitable for examining the privacy issues related to contextual end user data. The *Operational perspective* of the framework can be used to shed light on the ownership and control dynamics related to the data. The *Conceptual perspective*, on the other hand, is concerned on what information can be collected and thus enables examining privacy concerns related to the different context types. Together the two perspectives give richer insight on the complex matter of end user privacy than the perspectives alone.

4.1.1 Operational perspective

A report related to personal data by World Economic Forum [WEF, 12] brings up three types of personal data: volunteered (comes directly from the individual), observed (a result of a transaction between an individual and an organization, e.g., the individual's location during a phone call), and inferred (output of data analysis, combination and mining, i.e., derived from volunteered and observed data). The types are nicely comparable with the *Operational perspective* of the end user context classification framework where manually provided, sensed and derived correspond to volunteered, observed and inferred, respectively. The report discusses the different personal data types from privacy, and especially, control perspective. Many of these insights are in line with the *Operational* categories in the context framework. Table 2 describes the ownership and control dynamics related to contextual end user data.

Operational perspective		
Context	**Description**	**End user's sense of ownership and control**
Manually Provided	*Manually Provided* context information is produced and given by the end user. Thus the user might feel high sense of ownership and control towards the data (even regardless of who actually legally owns the data). Normally when collecting such data, an explicit consent for further use of the data is asked from the user.	High
Sensed	*Sensed* context information is produced by the sensors in the user's device. The sensors sense the user's context, but when purposefully collecting the data, usually some other party/organization(s) provide(s) the means for collecting and storing the data. In this case the sense of ownership and control moves towards the organization that captured the data. A possibility exists that the users do not fully comprehend the types and amount of data captured about them since they are not involved hands-on in producing the data.	Moderate
Derived	*Derived* context information is produced from *Manually Provided* and *Sensed* information by using relevant deriving methods. In addition to providing means for collecting and storing data, the organization(s) involved provide(s) also the deriving methods and perform the derivation. The methods might be results of extensive research efforts. Thus, the sense of ownership and control shifts even more away from the user and towards the organization. For-profit organizations generally see the analytics and insights derived from the data as a proprietary asset. The derived information can reveal even highly intimate insights about the user, but often the user's sense of direct control and awareness remains quite limited. A concern also arises if the derivation produces false results and the user is described incorrectly. This might lead to difficult consequences if someone then acts upon the incorrect information.	Low

Table 2: Operational perspective for end user privacy assessment, and description of ownership and control dynamics related to contextual end user data

4.1.2 Conceptual perspective

From privacy perspective the different types of data that can be used to describe an end user's context are relatively diverse. Table 3 describes, through the conceptual context classification, the differing privacy issues, and with the help of previous research considers especially end users' willingness to disclose the different types of data.

Conceptual perspective		
Personal Context	**Description**	**Willingness to disclose**
User	Direct information about the *User* is somewhat bipartite. According to [Ackerman, 99] and [Phelps, 00], demographic information such as age, name or marital status are given out with smaller concerns whereas, e.g., the social security number and information on income or health had much higher thresholds for disclosure. Normally the identity of the user is anonymized, but as argued, e.g., by [Ohm, 10] sophisticated deanonymization methods can undermine this approach.	High or Low (bipartite)
Social Environment	Information about the *Social Environment* of the user raises the issue where the end user is not the only one whose privacy is at stake. If, for example, Bluetooth scans, video clips or audio records connect the information also to other individuals, how much should these individuals have ownership or control over the information? According to [Wagner, 10], users are more cautious to explicitly disclose others' location information than their own, however the users might not be aware of all the information they disclose of others. Based on [Ackerman, 99] users were more uncomfortable disclosing any type of information about a child in their care compared to their own information.	? (Lower than for directly own information)
Activity/Task	Information about *Activities/Tasks* of the user can shed light on the user's lifestyle. According to, e.g., [Ackerman, 99] and [Phelps, 00] information on lifestyle, such as favorite hobbies, magazines, TV shows, snacks and leisure activities had quite a low threshold for disclosure (comparable, for example, to age information).	High
Environmental Context		**Willingness to disclose**
Conditions	As stand-alone information the environmental *Conditions* do not tell that much about the user herself. It is mainly environmental information that cannot be influenced by the user. Combined with other information it, however, can reveal some of the user's habits, but the assumption is that this information has a relatively low threshold for disclosure.	High
Infrastructure	Information on computational *Infrastructure* is privacy-wise presumably quite close to the environmental *Conditions*. Information related, for example, to the state of the cellular network is something that the user cannot directly influence (except by choosing the service provider). Information about the user's device is closer to the user and can reveal at least some preferences of the user. According to [Ackerman, 99] users were relatively willing to disclose information about their computer.	High or Moderate
Location	Privacy issues related to *Location* information are considered quite pressing, since by just tracking a user's whereabouts surprisingly much can be inferred regarding the user and her way of life (see, e.g., [Gasson, 11], [King, 11]). Based on previous research on location sharing such as [Consolvo, 05] and [Wagner, 10]: (i) Inquirer of the location information matters. This boils down to whom to trust, and to perceived harms to the user from disclosing. (ii) If the user perceives the information useful for the inquirer she discloses, otherwise not. (iii) If the user decides to disclose location information, she prefers precise information (more useful for the inquirer). (iv) Users would actually want to disclose additional information of, e.g., activities to explain why they are at a certain place.	High or Low (bipartite)

Table 3: Conceptual perspective for end user privacy assessment, and a brief examination on end users' willingness to disclose contextual data

4.1.3 Temporal perspective

Temporal information, depicted by the *Time* axis in the context framework, is also an important aspect when considering information privacy issues. It defines the point in time or the time period through which the context information on the other axes is valid. As the temporal information gets more granular, it combined with the other context information reveals more and more of the habits of the user. Temporal information is often a necessary component in turning *Manually Provided* and *Sensed* information towards *Derived* information, and thus provides means for loosening the user's grip of control to the personal information. In addition, temporal information enables the construction of a user's context history, which in turn can be used in attempts to predict the future contexts of the user.

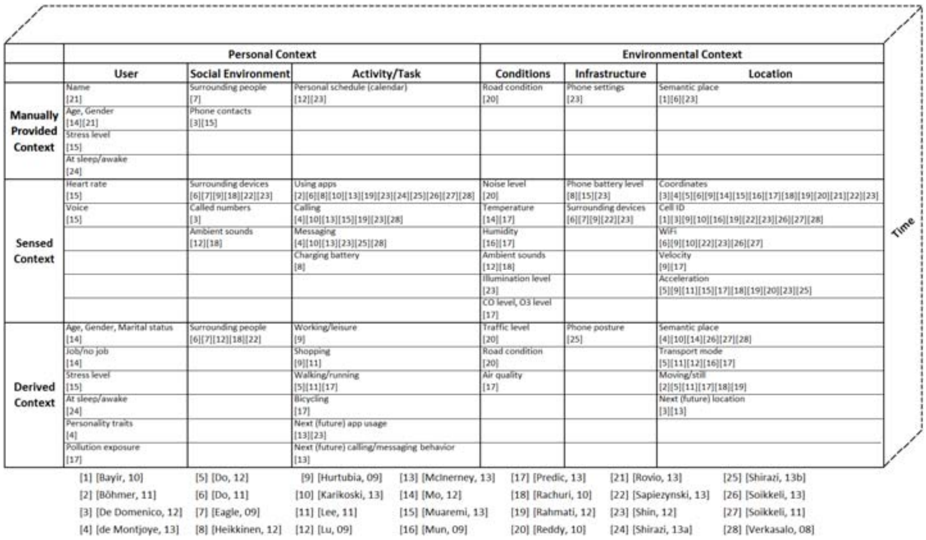
Privacy related to mobile end user context information is a sensitive and complex issue. As a whole it is an ambiguous entity, which needs to be split into parts for analysis. The end user context framework helps in this by considering the important aspects of data ownership and control, and end users' concerns regarding what to disclose. Together these give guidance on what to privacy wise expect when considering collecting certain types of data with certain methods. This is more diverse insight than can be achieved with many of the previous frameworks or conceptual models (e.g., [Henricksen, 04], [Schmidt, 99]). Of course the important questions related to the inquirer of data and perceived usage of the data [Adams, 00], [Lederer, 03], as well as the users' data-for-service trade-off preferences [Smith, 11] remain largely unanswered here. These, however, well largely from the perceived objectives of the data collecting and using entities.

4.2 Previous handset-based studies placed on the framework

The framework is also suitable for examining the broader state of the end user context oriented handset-based research activities in the light of context elements covered. Figure 2 shows a set of studies or handset data acquisition and usage activities (indicated by the numbers linked to the references) placed on the framework. The positions indicate which types of context elements the study has contemplated. Thus, one study appears in as many categories as there are context elements considered in the study. This is not an attempt to fully cover the research in this particular domain, but to just display some of it on the framework.

The positioning shows how some studies cover only a few categories whereas some others are slightly more ambitious in terms of the context coverage. However, none of the works does not, and obviously cannot, cover the context or situation of the user as a whole. On the broader level, it is quite evident that, at least among these studies, the *Location* context is the most popular topic of handset-based context research. This is not a surprise since the whole context-awareness research initially started more or less just as a location-awareness research [Schmidt, 99] and location is one of the easiest context elements to identify and measure [Kaasinen, 03]. It is also quite noticeable, and not surprising, that the main body of the studies shown here center around the *Sensed* context. The nature of handset-based measurements is automated sensing and collecting objective data without much user participation. *Manually Provided* context data has often been the result of adjunct basic information questionnaires to the users, and only more recently the data collection platforms have started to provide better, inbuilt, tools for added user participation. *Derived* context is the result of a derivation process and thus it is not as straightforward to acquire as the basic sensed data. However, *Derived* context is probably the most interesting area research-wise.

Positioning the studies both from the *Conceptual perspective* as well as the *Operational perspective* provides a possibility to examine what has been data acquisition wise attempted, and by what means. In several studies the context(s) of main interest are more high-level than the basic *Manually Provided* or *Sensed*. These contexts appear, as mentioned, on the *Derived context* part of the framework. It is interesting to see with one glance from what type of base data different studies have derived the higher-level contexts. This kind of knowledge on the state of the art helps, for example, in planning similar future studies.



	Personal Context			Environmental Context		
	User	Social Environment	Activity/Task	Conditions	Infrastructure	Location
Manually Provided Context	Name [21]	Surrounding people [7]	Personal schedule (calendar) [12][23]	Road condition [20]	Phone settings [23]	Semantic place [1][6][23]
	Age, Gender [14][21]	Phone contacts [3][15]				
	Stress level [15]					
	At sleep/awake [24]					
Sensed Context	Heart rate [15]	Surrounding devices [6][7][9][18][22][23]	Using apps [2][6][8][10][13][19][23][24][25][26][27][28]	Noise level [20]	Phone battery level [8][15][23]	Coordinates [3][4][5][6][9][14][15][16][17][18][19][20][21][22][23]
	Voice [15]	Called numbers [3]	Calling [4][10][13][15][19][23][28]	Temperature [14][17]	Surrounding devices [6][7][9][22][23]	Cell ID [1][3][9][10][16][19][22][23][26][27][28]
		Ambient sounds [12][18]	Messaging [4][10][13][23][25][28]	Humidity [16][17]		WiFi [6][9][10][22][23][26][27]
			Charging battery [8]	Ambient sounds [12][18]		Velocity [9][17]
				Illumination level [23]		Acceleration [5][9][11][15][17][18][19][20][23][25]
				CO level, O3 level [17]		
Derived Context	Age, Gender, Marital status [14]	Surrounding people [6][7][12][18][22]	Working/leisure [9]	Traffic level [20]	Phone posture [25]	Semantic place [4][10][14][26][27][28]
	Job/no job [14]		Shopping [9][11]	Road condition [20]		Transport mode [5][11][12][16][17]
	Stress level [15]		Walking/running [5][11][17]	Air quality [17]		Moving/still [2][5][11][17][18][19]
	At sleep/awake [24]		Bicycling [17]			Next (future) location [3][13]
	Personality traits [4]		Next (future) app usage [13][23]			
	Pollution exposure [17]		Next (future) calling/messaging behavior [13]			

[1] [Bayir, 10] [2] [Böhmer, 11] [3] [De Domenico, 12] [4] [de Montjoye, 13] [5] [Do, 12] [6] [Do, 11] [7] [Eagle, 09] [8] [Heikkinen, 12] [9] [Hurtubia, 09] [10] [Karikoski, 13] [11] [Lee, 11] [12] [Lu, 09] [13] [McInerney, 13] [14] [Mo, 12] [15] [Muaremi, 13] [16] [Mun, 09] [17] [Predic, 13] [18] [Rachuri, 10] [19] [Rahmati, 12] [20] [Reddy, 10] [21] [Rovio, 13] [22] [Sapiezynski, 13] [23] [Shin, 12] [24] [Shirazi, 13a] [25] [Shirazi, 13b] [26] [Soikkeli, 13] [27] [Soikkeli, 11] [28] [Verkasalo, 08]

Figure 2: Handset-based studies placed on the framework

4.3 Single study cases

In this section we pick two of the studies classified above for a closer inspection in order to examine the benefits of the framework on a single study level. First, we take a look at [Karikoski, 13]. According to the authors: "*The purpose of this article is to study how use context affects the usage patterns of smartphone communication services.*" However, the authors do not claim to capture the whole situation or context of the end user. The context element of main interest in the article is semantic place, which includes: home, office, other meaningful place (e.g., a place for a hobby), elsewhere and abroad. Also, the activities of app usage, calling and messaging are monitored. The work utilizes handset-based measurements to collect data on mobile users' locations and app usage and communication patterns. Figure 3 shows, on the end user context classification framework, the context elements acquired and used in the study.

To get a hold on the semantic place of the users, *Sensed* context data regarding the approximate location of the users were collected. Namely, the network cell IDs the users' phones were connected to, and nearby WLAN access points were sensed. Then a special purpose algorithm was utilized to derive the semantic place from the sensed (and time-stamped) data. Only after these operations, the article finally concentrates on its main problem, i.e., analyzing smartphone communication patterns in the different semantic places.

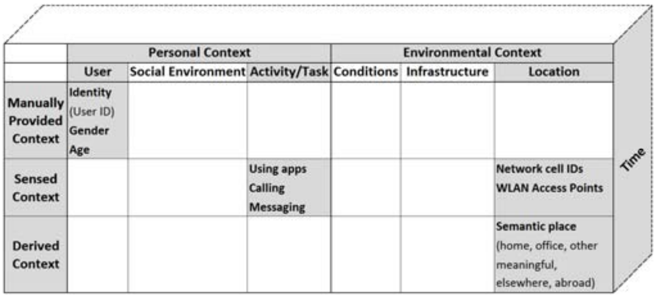
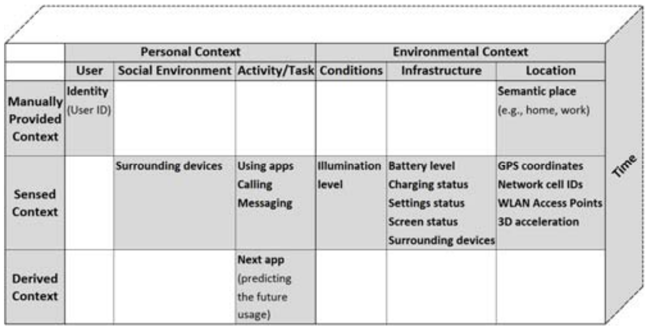


Figure 3: End user context elements used in [Karikoski, 13]

Next, we take a look at [Shin, 12]. In the article the authors "*first analyse the context in which mobile apps are used… [and] then create an application that predicts app usage based on [contextual information]"*. The authors claim to capture a wide range of contextual information and this is, indeed, true. However, the view to the end users' whole situation is still limited. The context elements acquired for the analysis include, for example, GPS coordinates, cell IDs, WLAN access points, acceleration and phone battery status, among others. Figure 4 shows all the context elements acquired on the end user context framework. In the article the authors are interested in a wide range of *Sensed* context elements. They use handset-based measurements to acquire data on these elements and then study the importance of different elements for predicting future app usage. Building on this knowledge, the authors construct context-aware naïve Bayes based predictive models for predicting which apps the end user is likely to use next. Finally, the authors develop a context-aware home screen application, which dynamically organizes the apps on the phone's home screen based on the likelihoods of user using the apps next.

The two studies considered are partly similar and partly quite different. As similarities go, both are interested in end user context's effects on smartphone usage. [Karikoski, 13] has mainly an interest towards examining user behavior in certain pre-defined contexts. As can be seen from Figure 3, the work is rather selective from the *Conceptual perspective*, i.e., on "general" context types. The research setting is such that it examines the effect of one type of *Environmental context* (*Location*) on one type of *Personal context* (*Activity*). The *User* context has only a supporting role. From the *Operational perspective* the work relies on raw *Sensed* data and computational methods for deriving the desired high-level context. By filling gaps appearing on the framework, future work could gain more insight and reliability. The *Social Environment* in the form of nearby people could be an important factor in the usage of communication services. A training set of *Manually Provided* semantic place data, on the other hand, could help validate the semantic place derivation methods. However, limitations on resources and technology always exist.



	Personal Context			Environmental Context		
	User	Social Environment	Activity/Task	Conditions	Infrastructure	Location
Manually Provided Context	Identity (User ID)					Semantic place (e.g., home, work)
Sensed Context		Surrounding devices	Using apps Calling Messaging	Illumination level	Battery level Charging status Settings status Screen status Surrounding devices	GPS coordinates Network cell IDs WLAN Access Points 3D acceleration
Derived Context			Next app (predicting the future usage)			

Figure 4: End user context elements used in [Shin, 12]

The end goal in [Shin, 12] is to build a context-aware application utilizing the most suitable contexts from the application's perspective. Reflecting on Figure 4, the work is rather comprehensive from the *Conceptual perspective*. The research setting is such that it tries to cover a wide range of general context types for evaluation. From *Operational perspective* the work narrows down quite heavily towards the *Sensed* context. Obviously, the sheer volume of the context elements makes an automated approach (from researcher and from end user side) most attractive. The authors also asked in interviews some information on the users' semantic places for a bit more insight. The end results of the work, i.e., the app prediction results are acquired by derivation (by machine learning methods). For future considerations, the wide range of general context types and context elements inside them, narrowed to be acquired by sensing, implies a lot of sensor activity. Battery consumption wise this might become a problem in real world applications.

The framework can also help in the planning phase of the studies. The goals of a study in mind, the framework can first act as a checklist for required context elements. Then by linking the elements to the *Conceptual* and *Operational* perspectives it is possible to assess the to-be-acquired data in terms of, e.g., effort and costs of acquisition, accuracy or privacy. *Manually provided* data may be tedious to collect, but sometimes a necessity for validating other data, for example. Acquiring *Sensed* data requires initial technical effort and data processing capabilities, but rewards with granular, and relatively objective and accurate data. If it turns out that the goals of the study require, e.g., more abstract information than directly available from *Manually Provided* or *Sensed* data, necessary derivation capabilities need to be acquired. Also, privacy assessments such as in [Section 4.1] can reveal aspects on efforts and costs of data collection.

Although the end user context framework proposed in this work rises mainly from context considerations linked heavily with the usage and operation of mobile devices (e.g., smartphones) themselves, handset-based measurements and the framework can be utilized also in other fields. For example, mobile marketing [Kurkovsky, 06], social psychology research [Miller, 12], ubiquitous health monitoring [Milosevic, 11], and urban planning [Reades, 07] and transportation studies [Calabrese, 11] are recognized to benefit from handset-based and contextual data. Linking consumers' preferences (manually provided or sensed) to, e.g., location (sensed or derived) and buying habits (derived) opens up new avenues for mobile marketing. Social psychology research has traditionally relied on pen and paper type of questionnaires and laboratory studies. Smartphones can be used to replace pen and paper in acquiring manually provided data, while ubiquitously sensing the context of a study participant provides undoubtedly new opportunities. Ubiquitous health monitoring where the patients, in addition to manually provided data (e.g., with a smartphone), are also ubiquitously monitored by sensors (both personal and environmental contexts) provides new opportunities for preventive healthcare, patient-doctor interaction and reacting on emergencies. Detailed handset-based data on people's locations, movements, transportation modes, and other contexts provide big opportunities for better urban and public transportation planning.

5 Conclusions and future work

In this paper we proposed an end user context classification framework for handset-based studies. The motivation for such a framework rises from the somewhat ambiguous usage of the term context in handset-based and context-awareness research. Our end user context classification framework is positioned with the existing and most used context-awareness and ubiquitous computing related context definitions. It represents a broad picture of handset-based end user context by classifying elements of context we have identified as being possible to infer from handset-based data at the moment and in the foreseeable future. With the help of the framework it is possible to communicate clearly, which particular context elements are under scrutiny in a certain study and how the elements are related to the broader context or the situation of an end user. Also, the framework facilitates more precise planning of context data acquisition activities by relating conceptual context types to operational aspects of context data collection. The framework is intended to help also other stakeholders, such as service providers and application developers, to identify research most relevant for their particular needs.

In the framework we represent the end user context in three dimensions: *Conceptual perspective*, *Operational perspective*, and *Time*. The *Conceptual perspective* of end user context is divided first into *Personal* and *Environmental* contexts. One step further divides *Personal context* into context elements characterizing the end *User* herself, the *Social environment* of the end user, and *Activity/Task* of the end user. *Environmental context* is divided into environmental *Conditions* surrounding the end user, computational *Infrastructure* available, and *Location* of the end user. The *Operational perspective* of end user context is divided into *Manually Provided*, *Sensed*, and *Derived* contexts. *Manually Provided* information comes directly from the end user, *Sensed* information comes from different sensors available in the end user's handset, and *Derived* information is a result of a derivation process, which uses *Manually Provided* and/or *Sensed* information as an input. The *Time* dimension takes into account the temporal dynamics of end user context and provides context history, valuable especially in acquiring *Derived* context information. We also apply the proposed framework, in an example-like manner, to assess privacy issues related to end user context, to show what type of context elements previous handset-based studies have examined, and to study individual research cases through the lens of the framework.

In addition to being end user centric, the present work is also quite heavily mobile handset-centric. Thus far, the smartphone has been on the cutting edge of collecting end user context information. In the future, for example, wearable devices and the Internet of Things (IoT) paradigm might change this. From the perspective of the proposed framework, wearable devices are possibly even a more accurate proxy for measuring the end users' context. However, the wearable devices might be less suitable for *Manually Provided* context. IoT can provide additional external sensors for *Environmental context* sensing. With multiple devices providing the context information, the *Operational perspective* of the framework could be modified to distinguish a set of different devices. Also, in the IoT paradigm the context of the things (the different devices) is seen important. The proposed framework could adapt to this, for example, by replacing the user's *Personal context* by sort of a device (or thing) context, i.e., the device's personal context.

Acknowledgements

This work has been supported by the MoMIE project of Aalto University and the Future Internet Graduate School FIGS.

References

[Ackerman, 01] Ackerman, M., Darrell, T., Weitzner, D.J.: Privacy in Context, Human-Computer Interaction, 16(2-4), 2011, 167-176.

[Ackerman, 99] Ackerman, M., Cranor, L.F., Reagle, J.: Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences, In Proceedings of the 1st ACM conference on Electronic commerce, November 1999.

[Adams, 00] Adams, A.: Multimedia Information Changes the Whole Privacy Ballgame, In Proceedings of the 10th conference on Computers, freedom and privacy: challenging the assumptions, April 2000, 25-32.

[Aharony, 11] Aharony, N., Pan, W., Ip, C., Khayal, I., Pentland, A.: Social fMRI: Investigating and shaping social mechanisms in the real world, Pervasive and Mobile Computing, 7(6), 2011, 643-659.

[Barkhuus, 03] Barkhuus, L.: Context information in mobile telephony, Chittaro, L. (ed.) MobileHCI LNCS, vol. 2795, Springer, Berlin, pp. 451–455.

[Bayir, 10] Bayir, M.A., Demirbas, M., Eagle, N.: Mobility profiler: A framework for discovering mobility profiles of cell phone users, Pervasive and Mobile Computing, 6(4), August 2010, 435-454.

[Beresford, 03] Beresford, A.R., Stajano, F.: Location Privacy in Pervasive Computing, IEEE Pervasive Computing, 2(1), 2003, 46-55.

[Berkovsky, 09] Berkovsky, S., Heckmann, D., Kuflik, T.: Addressing Challenges of Ubiquitous User Modeling: Between Mediation and Semantic Integration, Advances in Ubiquitous User Modelling in Lecture Notes in Computer Science, 5830, 2009, 1-19.

[Böhmer, 11] Böhmer, M., Hecht, B., Schöning, J., Kruger, A., Bauer, G.: Falling asleep with Angry Birds, Facebook and Kindle: a large scale study on mobile application usage, In

Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, August 2011, 47-56.

[Calabrese, 11] Calabrese, F., Di Lorenzo, G., Liu, L., Ratti, C.: Estimating Origin-Destination Flows Using Mobile Phone Location Data, IEEE Pervasive Computing, 10(4), 36-44.

[Consolvo, 05] Consolvo, S., Smith, I.E., Matthews, T., LaMarca, A., Tabert, J., Powledge, P.: Location Disclosure to Social Relations: Why, When, & What People Want to Share, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 2005.

[De Domenico, 12] De Domenico, M., Lima, A., Musolesi, M.: Interdependence and Predictability of Human Mobility and Social Interactions, In: arXiv:1210.2376 [physics.soc-ph], 2012.

[de Montjoye, 13] de Montjoye Y.-A., Quoidibach, J., Robic, F., Pentland, A.S.: Predicting Personality Using Novel Mobile Phone-Based Metrics, Social Computing, Behavioral-Cultural Modeling and Prediction in Lecture Notes in Computer Science, 7812, 2013, 48-55.

[Dey, 01a] Dey, A.K.: Understanding and Using Context, Personal and Ubiquitous Computing, 5(1), 2001, 4-7.

[Dey, 01b] Dey, A.K., Abowd, G.D., Salber, D.: A Conceptual Framework and Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications, Human-Computer Interaction, 16(2-4), 2001, 97-166.

[Do, 11] Do, T., Blom, J., Gatica-Perez, D.: Smartphone usage in the wild: a large-scale analysis of applications and context, In Proceedings of the 13th International Conference on Multimodal Interfaces, November 2011, 353-360.

[Do, 12] Do, T.M., Loke, S.W., Liu, F.: HealthyLife: an Activity Recognition System with Smartphone using Logic-Based Stream Reasoning, In Proceedings of the 9th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, December 2012.

[Dourish, 04] Dourish, P.: What we talk about when we talk about context, Personal and Ubiquitous Computing, 8(1), 2004, 19-30

[Eagle, 06] Eagle, N., Pentland A.: Reality Mining: Sensing Complex Social Systems, Personal and Ubiquitous Computing, 10 (4), 2006, 255-268.

[Eagle, 09] Eagle, N., Pentland, A., Lazer, D.: Inferring social network structure using mobilephone data, Proceedings of the National Academy of Sciences (PNAS), 106(36), 2009, 15274–15278.

[Falaki, 10] Falaki, H., Mahajan, R., Kandula, S., Lymberopoulos, D., Govindan, R., Estrin, D.: Diversity in Smartphone Usage, In Proceedings of the 8th international conference on Mobile systems, applications, and services, June 2010, 179-194.

[Gasson, 11] Gasson, M. N.: Normality Mining: Privacy Implications of Behavioral Profiles Drawn From GPS Enabled Mobile Phones, IEEE Transactions on Systems, Man, and Cybernetics, 41(2), 2011, 251-161.

[Heikkinen, 12] Heikkinen, M.V.J., Nurminen, J.K., Smura, T., Hämmäinen, H.: Energy efficiency of mobile handsets: Measuring user attitudes and behavior, Telematics and Informatics, 29(4), 2012, 387-399.

[Henricksen, 04] Henricksen, K., Indulska, J.: Modelling and using imperfect context information, In Proceedings of the 2nd IEEE Annual Conference on Pervasive Computing and Communications Workshops, March 2004, 33-37

[Hong, 07] Hong, D., Chiu, D., Shen, Y., Kafeza, E.: Ubiquitous enterprise service adaptations based on contextual user behavior, Information Systems Frontiers, 9(4), 2007, 343-358.

[Hurtubia, 09] Hurtubia, R., Flötteröd, G., Bierlaire, M.: Inferring the Activities of Smartphone Users from Context Measurements Using Bayesian Inference and Random Utility Models, In Proceedings of European Transport Conference 2009, May 2009.

[Kaasinen, 03] Kaasinen, E.: User needs for location-aware mobile services, Personal and Ubiquitous Computing, 7(1), 2003, 70–79.

[Karikoski, 10] Karikoski, J., Nelimarkka, M.: Measuring social relations: Case OtaSizzle, In Proceedings of the Second IEEE International Conference on Social Computing, August 2010.

[Karikoski, 12] Karikoski, J.: Handset-based data collection process and participant attitudes, International Journal of Handheld Computing Research, 3(4), 2012, 1-21.

[Karikoski, 13] Karikoski, J., Soikkeli, T.: Contextual usage patterns in smartphone communication services, Personal and Ubiquitous Computing, 17(3), 2013, 491-502.

[Kekolahti, 13] Kekolahti, P., Karikoski, J.: Analysis of Mobile Service Usage Behaviour with Bayesian Belief Networks, Journal of Universal Computer Science, 19(3), 2013, 325-352.

[King, 11] King, G.: Ensuring the Data-Rich Future of the Social Sciences, Science, 331(6018), 2011, 719-721.

[Kurkovsky, 06] Kurkovsky S., Harihar, K.: Using ubiquitous computing in interactive mobile marketing, Personal and Ubiquitous Computing, 10, 2006, 227-240.

[Langheinrich, 01] Langheinrich, M.: Privacy by Design - Principles of Privacy-Aware Ubiquitous Systems, In Proceedings of Ubicomp 2001, September 2001.

[Lederer, 03] Lederer, S., Mankoff, J., Dey, A.K.: Who Wants to Know What When? Privacy Preference Determinants in Ubiquitous Computing, In Extended Abstracts on Human Factors in Computing Systems, 2003, 724-725.

[Lee, 11] Lee, Y.,Cho, S.: Activity Recognition Using Hierarchical Hidden Markov Models on a Smartphone with 3D Accelerometer, Hybrid Artificial Intelligent Systems: Lecture Notes in Computer Science, 6678, 2011, 460-467.

[Liang, 11] Liang, T., Yeh, Y.: Effect of use contexts on the continuous use of mobile services: the case of mobile games, Personal and Ubiquitous Computing, 15(2), 2011, 187-196.

[Lieberman, 00] Lieberman, H., Selker, T.: Out of context: Computer systems that adapt to, and learn from, context, IBM Systems Journal, 39(3-4), 2000, 617-632.

[Lu, 09] Lu, H., Lane, N.D., Choudhury, T., Campbell, A.T.: SoundSense: Scalable Sound Sensing for People-Centric Applications on Mobile Phones, In Proceedings of the 7th international conference on Mobile systems, applications and services, June 2009, 165-178.

[McInerney, 13] McInerney, J., Stein, S., Rogers, A., Jennings, N.R.: Breaking the habit: Measuring and predicting departures from routine in individual human mobility, Pervasive and Mobile Computing, 9, 2013, 808-822.

[Miller, 12] Miller, G.: The Smartphone Psychology Manifesto, Perspectives on Psychological Science, 7(3), 2012, 221-237.

[Miller, 95] Miller, G.A.: WordNet: A Lexical Database for English, Communications of the ACM, 38(11), 1995, 39-41.

[Milosevic, 11] Milosevic, M., Shrove, M.T., Jovanov, E.: Applications of Smartphones for Ubiquitous Health Monitoring and Wellbeing Management, Journal of Information Technology and Applications, 1(1), 2011, 7-15.

[Mo, 12] Mo, K., Tan, B., Zhong, E., Yang, Q.: Your Phone Undestands You, In Mobile Data Challenge (by Nokia) Workshop, June 2012.

[Muaremi, 13] Muaremi, A., Arnrich, B., Tröster, G.: Towards Measuring Stress with Smartphones and Wearable Devices During Workday and Sleep, BioNanoScience, (Open Access), DOI: 10.1007/s12668-013-0089-2, 2013.

[Mun, 09] Mun, M., Reddy, S., Shilton, K., Yau, N., Burke, J., Estring, D., Hansen, M., Howard, E., West, R., Boda, P.: PEIR, the Personal Environmental Impact Report, as a Platform for Participatory Sensing Systems Research, In Proceedings of the 7th international conference on Mobile systems, applications and services, June 2009, 55-68.

[Nokia, 07] Nokia, Nokia Smartphone360 Study - Responding to what the user wants and needs, Nokia Corporation, press release, October 2007, http://www.nokia.com/NOKIA_COM_1/Press/twwln/press_kit/Nokia_Smartphone360_study_ Press_Backgrounder_October_2007.pdf, 2007.

[Ohm, 10] Ohm, P.: Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization, UCLA Law Review, 57, 2010, 1701-1777.

[Oxford dictionaries, 14] Oxford University Press, 2014, Accessed August 25th 2014 from: http://www.oxforddictionaries.com/definition/english/context.

[Perera, 14] Perera, C., Zaslavsky, A., Christen, P., Georgakopoulos, D.: Context Aware Computing for The Internet of Things: A Survey, IEEE Communications Surveys & Tutorials, 16(1), 2014.

[Phelps, 00] Phelps, J., Nowak, G., Ferrell, E.: Privacy Concerns and Consumer Willingness to Provide Personal Information, Journal of Public Policy & Marketing, 19(1), 2000, 27-41.

[Predic, 13] Predic, B., Zhixian, Y., Eberle, J., Stojanovic, D., Aberer, K.: ExposureSense: Integrating Daily Activities with Air Quality using Mobile Participatory Sensing, In IEEE International Conference on Pervasive Computing and Communication, March 2013.

[Rachuri, 10] Rachuri, K.K., Musolesi, M., Mascolo, C., Rentfrow, P.J., Longworth, C., Aucinas, A.: EmotionSense: A Mobile Phones based Adaptive Platform for Experimental Social Psychology Research, In Proceedings of the 12th ACM international conference on Ubiquitous computing, September 2010, 281-290.

[Raento, 05] Raento, M., Oulasvirta, A., Petit, R., Toivonen, H.: Context-Phone: A Prototyping Platform for Context-Aware Mobile Applications, IEEE Pervasive Computing, 4(2), 2005, 51-59.

[Rahmati, 12] Rahmati, A., Shephard, C., Tossell, C., Zhong, L., Kortum, P.: Practical Context Awareness: Measuring and Utilizing the Context Dependency of Mobile Usage, Technical report, Rice University, 2012.

[Reades, 07] Reades, J., Calabrese, F., Svetsuk, A., Ratti, C.: Cellular Census: Explorations in Urban Data Collection, IEEE Pervasive Computing, 6(3), 2007, 30-38.

[Reddy, 10] Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., Srivastava, M.: Using mobile phones to determine transportation modes, ACM Transactions on Sensor Networks (TOSN), 6(2), 2010.

[Ritchey, 91] Ritchey, T.: Analysis and synthesis: On scientific method – based on a study by Bernhard Riemann, Systems Research, 8(4), 1991, 21-41.

[Rovio, 13] Rovio: Angry Birds Privacy Policy. Rovio Entertainment Ltd, http://www.rovio.fi/Privacy, 2013.

[Sapiezynski, 13] Sapiezynski, P., Stopczynski, A., Lehman, S.: Detecting Face-to-face Meetings using Smartphone Sensors, In NetMob13: the Third International Conference on the Analysis of Mobile Phone Datasets, May 2013.

[Schilit, 94] Schilit, B.N., Theimer, M.M.: Disseminating active map information to mobile hosts, IEEE Network, 8(5), 1994, 22-32.

[Schmidt, 99] Schmidt, A., Beigl, M., Gellersen H.: There is more to context than location, Computers & Graphics, 23(6), 1999, 893-901.

[Shepard, 10] Shepard, C., Rahmati, A., Tossell, C., Zhong, L., Kortum, P.: LiveLab: Measuring Wireless Networks and Smarphone Users in the Field, ACM SIGMETRICS Performance Evaluation Review, 38(3), 2010, 15-20.

[Shin, 12] Shin, C., Hong, J., Dey, A.K.: Understanding and prediction of mobile application usage for smart phones, In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, September 2012, 173-182.

[Shirazi, 13a] Shirazi, A.S., Clawson, J., Hassanpour, Y., Tourian, M.J., Schmidt, A., Chi, E.H., Borazio, M., Van Laerhoven, K.: Already up? using mobile phones to track & share sleep behavior, International Journal of Human-Computer Studies, (Available online), http://dx.doi.org/10.1016/j.ijhcs.2013.03.001, 2013.

[Shirazi, 13b] Shirazi, A.S., Henze, N., Dingler, T., Schmidt, A., Kunze, K.: Upright or Sideways? Analysis of Smartphone Postures in the Wild, In the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services, August 2013.

[Smith, 11] Smith, H.J., Dinev, T., Xu, H.: Information Privacy Research: An Interdiciplinary Review, MIS Quarterly, 35(4), 2011, 989-1015.

[Soikkeli, 11] Soikkeli, T., Karikoski, J., Hämmäinen, H.: Diversity and end user context in smartphone usage sessions, In Proceedings of the 5th International Conference on Next Generation Mobile Applications, Services and Technologies, September 2011, 7-12.

[Soikkeli, 13] Soikkeli, T., Karikoski, J., Hämmäinen, H.: Characterizing Smartphone Usage: Diversity and End User Context, International Journal of Handheld Computing Research, 4(1), 2013, 15-36.

[Tamminen, 04] Tamminen, S., Oulasvirta, A., Toiskallio, K., Kankainen, S.: Understanding mobile contexts, Personal and Ubiquitous Computing, 8, 2004, 135–143.

[Wagner, 10] Wagner, D., Lopez, M., Doria, A., Pavlyshak, I., Kostakos, V., Oakley, I., Spiliotopoulos, T.: Hide and Seek: Location Sharing Practices With Social Media, In Proceedings of the 12th International Conference on Human-Computer Interaction with Mobile Devices and Services, September 2010.

[Van Bunningen, 05] van Bunningen, A., Feng, L., Apers, P.: Context for ubiquitous data management, In Proceedings of International Workshop on Ubiquitous Data Management UDM 2005, April 2005, 17 – 24.

[WEF, 12] World Economic Forum: Rethinking Personal Data: Strengthening Trust, http://www3.weforum.org/docs/WEF_IT_RethinkingPersonalData_Report_2012.pdf, 2012.

[Weiser, 93] Wiser, M.: Some Computer Science Issues in Ubiquitous Computing, Communications of the ACM, 36(7), 1993, 75-84.

[Verkasalo, 08] Verkasalo, H.: Contextual patterns in mobile service usage, Personal and Ubiquitous Computing, 13(5), 2008, 331-342.

[Xu, 09] Xu, Z., Yuan, Y.: The impact of context and incentives on mobile service adoption, International Journal of Mobile Communications, 7(3), 2009, 363-381.