THE ROLE OF SOCIAL SOFTWARE ENGINEERING IN THE DESIGN OF PRIVACY-AWARE INFORMATION SYSTEMS

Aikaterini Vgena[*]
Angeliki Kitsiou[*]
Dimitris Kavroudakis[*]
Christos Kalloniatis[*]

Abstract

Social Software Engineering is a new field of study that aims to identify how social aspects interact and affect the design of software during the software development cycle. The identification of social parameters and criteria that affect the elicitation of technical, functional and non-functional requirements is a critical step for the system to be. Protecting users' privacy on the other hand, is also critical since users tend to utilize services and applications they trust especially regarding the manipulation

[*] University of the Aegean, Department of Cultural Informatics
[*] University of the Aegean, Department of Cultural Informatics
[*] University of the Aegean, Department of Cultural Informatics
[*] University of the Aegean, Department of Cultural Informatics

and handling of their personal information. At the same time users are inclined to express their privacy concerns regarding mobile applications that store and disseminate various privacy related information e.g. geolocation data. The social parameters and social criteria which are indicative for setting requirements and designing software focusing on privacy issues by previous literature, are explored, focusing on users' social identity and geolocation data in social media. Drawing on Social, Information and Communication and Geolocation Theories, users' special characteristics of social identity- multiplicity, permeability and overlapping- will be explored, as well as their possible revealing information by utilizing geolocation services - space of actions, frames-, in order to identify users' faces in this field. This work aims to broaden understanding of users' digital identities and geolocation data impact through an interdisciplinary approach, on the setting of requirements elicitation and modelling approaches for designing privacy-aware systems.

1. INTRODUCTION

It is an undeniable fact that over the past few years research related to software engineering have been on the rise. More precisely, research has shown that social aspects should be further investigated in the process of designing systems and setting requirements. In that way, a new field of study started to present itself as a vital part of software engineering, Social Software Engineering (Jenkins, 2008; Lahlou, 2008; Miguel & Medina, 2011; Nario-Redmond, Biernat, Eidelman, & Palenske, 2004; Schwartz & Halegoua, 2015). Social Software Engineering (SSE) proposes the investigation and implementation of social aspects and parameters form an early stage of designing Privacy-aware Information Systems (Lahlou, 2008; Miguel & Medina, 2011).

Social Software Engineering is a relatively new branch of software engineering which deals with the social aspects of software development. Social Software Engineering is a key term in order to proceed in discussing the protection of users' personal information from the point of incorporating social parameters and social criteria in the system's requirements.

More specifically, Social Software Engineering's radical approach represents a new field of study, which aims to identify how social aspects interact and affect the design of software during the software development cycle, providing an interesting brand-new interdisciplinary approach for designing privacy-aware systems.

This paper aims to provide the core role of Social Software Engineering in designing Privacy-aware Information Systems, which will be further analyzed in setting requirements that will meet users' needs in a realistic and adaptive way. Digital identity (part of users' social identity) and user's current geolocation data (users' geographic place) are going to shed light upon the design of Privacy-aware Information Systems in Social Software Engineering. Digital identity can provide valuable information on users' profile (Jenkins, 2008; Lahlou, 2008; Schwartz & Halegoua, 2015) while geolocation will set a specific context to our analysis as it provides a specific space of action for the users.

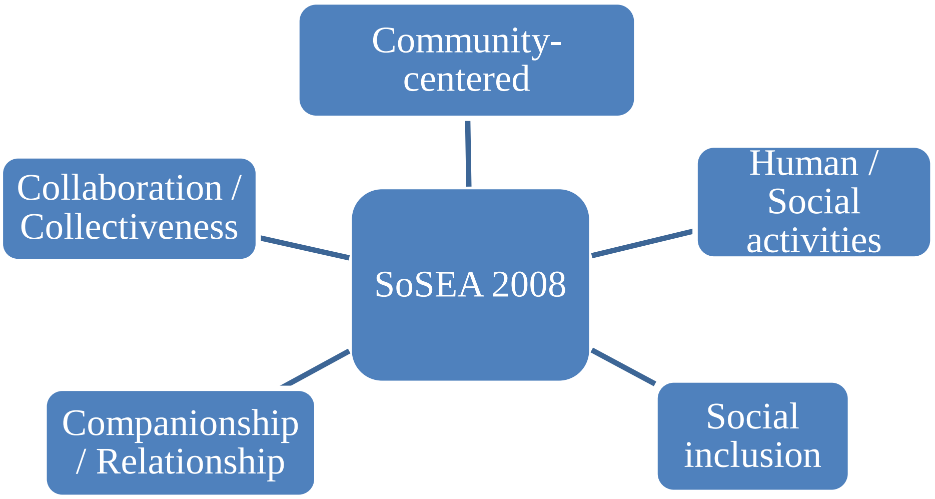
The rest of the paper is organized as follows. Section 2 presents the findings of the 1st International Workshop on Social Software Engineering and Applications (SoSEA) in 2008. Section 3 focuses on digital identity, discussing some general concerns (3.1) and the qualities of digital identities (3.2), which are further analyzed as variables. The example of geolocation in Section 4 will shed light upon digital identities and geolocation data (4.1) and the Privacy Paradox through the geolocation example (4.2). Section 5 drawing on Social, Information and Communication and Geolocation Theories, presents users' special characteristics of social identity, while discussing notions, such as Faces, Frames, Stages and Performances (5.1) and Predictability and Tracking Users' Normativity (5.2) and their elicitation (5.3). Finally, the last Section recalls the main discussion and raises future research objectives.

Designers seem to highlight the importance of Privacy-aware Information Systems in every step of the designing process (Fabio Massacci, Marco Prest and Nicola Zannone, 200AD; Nicola Zannone, Fabio Massacci & and John Mylopoulos, 2006). Satisfying privacy requirements in informatics both in reference to computer security and communication security is of primary importance. Engineers seek new ways in defining and designing systems in order to address possible threats and minimize the risk of a probable privacy leak in the user's personal information in a more effective way, yet, the risk cannot be entirely eliminated. However, setting meticulously analyzed requirements for Privacy-aware Information Systems in Social Software Engineering can cope with potential threats up to a satisfactory level (Fabio Massacci, Marco Prest and Nicola Zannone, 200AD; Nicola Zannone, Fabio Massacci & and John Mylopoulos, 2006).

Privacy and security awareness is a core value in the design of information systems. It's primary concerns are in accordance with the ones of Social Software Engineering. In other words, the 1st International Workshop on Social Software Engineering and Applications (SoSEA 2008) proved to be a cornerstone in designing Privacy-aware Information Systems in

Social Software Engineering. More precisely, Social Software Engineering's principles were set during that conference (Abbattista, Calefato, Gendarmi, & Lanubile, 2008). The principles are schematically represented in Figure 1.

Figure 1. Social Software Engineering Principles, the 1st International Workshop on Social Software Engineering and Applications (SoSEA 2008)


Community-centered
Collaboration / Collectiveness
Human / Social activities
SoSEA 2008
Companionship / Relationship
Social inclusion

As shown in Figure 1, the 1st International Workshop on Social Software Engineering and Applications set the principles for the Social Software Engineering (Abbattista et al., 2008): More precisely, engineers should design systems that are focusing on social aspects, such as:

- Community-centered: Produce software that promotes society's wellbeing, instead of private interest.
- Collaboration / Collectiveness: Reference to the human collaborative and collective abilities.
- Companionship / Relationship: Designing systems supporting human activities in order to target social problems
- Social inclusion: Designing software should promote social inclusion by creating social bonds and mutual trust among its social actors.

Social Software Engineering and Privacy-aware Information systems seem to be interconnected, as there is a common effort for designing systems which intend to protect users' privacy. Their main target is to produce software focusing on social well-being. Social actors' satisfaction, cooperation and privacy protection are of utmost importance as the main objective of Social Software Engineering and Applications is to create a generalized sense of belonging in a group (strengthen people's sense of belonging).

2. Digital Identity

2.1.General Concerns

Researchers argue that digital identity is an important part of social identity, as nowadays users' daily lives seem to have both an offline and online dimension. Taking both ways in which people's identity is shaped nowadays into consideration, it is not surprising that people tend to spend a lot of their time online, for various reasons, i.e. work, spend their free time, listen to music, watch a movie or socialize. Interestingly enough, social media applications and social media users are on the rise, while at the same time the majority of internet users also possess a social media account on Facebook, Tweeter, Instagram or LinkedIn.

Therefore, in order to examine the importance of online / digital identities, it seems valid to proceed first in investigating a number of social aspects, parameters and criteria, which would have a minor or major influence on the construction of users' online / digital identity in their social media accounts. Those social aspects, parameters and criteria are also going to allow a number of further assumptions before setting specific technical, functional and non-functional requirements.

Specifically, the main characteristics of social identity, namely multiplicity, permeability and overlapping, which have already been widely explored, are to be meticulously analyzed from the point of the digital identity through social media representation while at the same time additional key elements and qualities of the digital identity are going to be introduced as variables to our analysis. Last but not least, users' digital identities are also going to be analyzed as they are intertwined and identified as dialectical and reciprocal (Miguel & Medina, 2011).

2.2.Qualities of Digital Identities: Variables

Social aspects, parameters and social criteria are to be defined in order to proceed in creating a valid methodology in designing Privacy-aware Information systems using Social Software Engineering focusing on privacy and security issues in an adaptive way. Setting specific technical, functional and non-functional requirements should be introduced from an early stage of the system's design in accordance with a number of social identity characteristics that are going to be investigated and verified. In order to set specific technical, functional and non-functional requirements, it is important to understand a number of social aspects parameters and criteria.

The basic characteristics of online identities mentioned by previous literature are, namely: gender, age group, national group/identity, permanent address, religious beliefs, political orientations, education and field of study, free time activities, hobbies, branding, likes, mentions and traffic flows (Balicki, 2014; Jenkins, 2008; Nario-Redmond et al., 2004).

What is equally important is to investigate users' intimacy, which is willingly converted to extimacy (public content which is shared and uploaded online) in order to be consumed as entertainment content through social media (Miguel & Medina, 2011).

3. Geolocation

3.1. Digital Identities and Geolocation data

Location data can prove to be another quite descriptive aspect of users' digital identity. This kind of information provide not only a specific setting (place) for users' digital activity but also information concerning user's activity at a specific geographic location at a specific time of the day (Barkhuus & Dey, 2003; Bettini, Wang, & Jajodia, 2005; Myles, Friday, & Davies, 2003). In that way, geolocation data can provide sensitive information on user's activity space or space of action. Revealing user's activity space places them in a specific context, providing their social status, geographic location or intentions (Lahlou, 2008).

Space of action is often referred to as Frame by Social Theory. That is to say, users tend to cover different needs in different social settings, as places function as stages (Barkhuus & Dey, 2003; Duckham & Kulik, 2006). In that way, users' social media representation at a specific space makes them follow certain social norms, which in Social Theory are referred to as Faces (Jenkins, 2008; Lahlou, 2008).

As a result collecting geolocation data provide information on tracking users' trajectory and thus being able to proceed in making assumptions about user's normativity. In that way, users' trajectories can predict their future choices as social status, geographic places and users' ambitions will have already defined them as biographical subjects (Miguel &

Medina, 2011). Keeping that in mind, it is understandable that privacy aware systems are vital n protecting users' online activity from malware or private profit.

In other words, geolocation data are descriptive in unveiling digital and thus social identity, as people act as biographical subjects (Lahlou, 2008). That is to say, users, are affected by their past experiences as social actors while making upcoming decisions. In a nutshell, actions that are located at a specific place and time of the day in a repetitive manner can create affiliations and correlations not only for users' identity but for their future choices, too.

Those conclusions can lead in tracking user's trajectories and spotting their online normativity. That is why, there is a variety of protection mechanisms, providing solutions for handling user's geolocation data. Location-Privacy Protection Mechanisms (LPPMs) utilize different ways of protecting these sensitive personal data (Beresford & Stajano, 2003; Consolvo et al., 2005; Duckham & Kulik, 2005; Liu, 2007; Shokri, Theodorakopoulos, Boudec, & Hubaux, 2011; Snekkenes, 2001).

3.2.Privacy Paradox and Geolocation

Taking into consideration the aforementioned discussion on geolocation data and the increasing number both of social media applications and social media users , research findings often seem perplexed as far as a conflicting, complex phenomenon is concerned, the "Privacy Paradox".

Previous research/literature investigate user's tendency to reveal personal information online while at the same time expressing their fear about probable privacy leaks. The recent example of Facebook data leak and the debate about it, increased users' awareness on the topic. However, users tend to share their personal and geolocation data while online. The phenomenon which describes the difference between users' online behavior and their beliefs is known as the "Privacy Paradox".

In any case, the protection of users' personal information is a matter of serious investigation. This approach tries to identify how users' different identities interact simultaneously while being affected by the social media's settings, forming the online identity of each user. Further analysis of the aforementioned theories will allow a deeper understanding of users' digital identities and geolocation data impact in their everyday social media practice.

4. Faces, Frames, Stages and Performances

Drawing on Social, Information and Communication and Geolocation Theories, users' special characteristics of social identity, namely, multiplicity, permeability and overlapping, will be explored along with the aforementioned social identity variables. In this analysis, geolocation services will provide the space of action setting for users' activity. Frames are going to be vital in identifying users' faces in this field.

According to Jenkins (2018), researchers can identify users' specific space of actions which are characterized by a set of frames, using location data given at a specific time. In other words, location data provide specific space of action information (frame). Thus, frames can function as stages, as they give additional meaning to social representations by setting norms to be identified, applied and followed by the social or digital actors (Jenkins, 2008; Lahlou, 2008).

As users follow certain social norms, they tend to cover different needs in different social settings through their social media representation in their activity space. At that point, each user seem to utilize a specific face in order to communicate appropriately within a specific context (social setting). The user expresses a part of his or her personal identity by taking part in an online social performance.

4.1.Predictability and Tracking Users' Normativity

Due to the Privacy Paradox phenomenon, users are inclined to express their privacy concerns regarding applications that store and disseminate various privacy related information e.g. geolocation data, while at the same time revealing willingly this sensitive data. This practice, which is available in the most social media applications, can create new affiliations by representing one's digital self, reshape social media practice and transform the way geographical locations are perceived in digital world.

However, disseminating geolocation data, a quite common practice among social media users, may intensify possible threats as far as privacy and security are concerned, as revealing user's current geographical position can be indicative for one's digital identity (Barkhuus & Dey, 2003; Cramer, Rost, & Holmquist, 2011; Tang, Lin, Hong, Siewiorek, & Sadeh, 2010). Thus, spotting oneself at a specific place may create a link between the place (frame) and the current user's face. According to Schwartz and Halegoua's "spatial self"

(2015), describes the process of online self-representation which is primarily based on offline activities.

Furthermore, it is important to underline the necessity of dealing with potential privacy and security risks which are constantly increasing due to the track record of the user's geolocation data dissemination. This kind of data are descriptive enough to enable conclusions or further correlations about user's activities and thus should be taken under serious consideration while setting system requirements (Bettini et al., 2005).

As geolocation data enable conclusions or further correlations among faces, stages and social performances it is possible to track users' online trajectories and proceed in identifying normativity patterns, increasing the accuracy of predictability results (Jenkins, 2008; Miguel & Medina, 2011). In this way, setting requirements and modelling approaches for designing privacy-aware systems in Social Software Engineering is of utmost importance as data correlations and potential conclusions about users' activity could attract malware or lead to private profit.

4.2.Elicitation of Faces, Frames, Stages and Performances in Tracking Users' Normativity (Predictability)

At this point, Figure 2 would shed light upon the use of the aforementioned notions of Face, Frame, Stage and Public Performance. More specifically, users utilize different faces in order to be effective in different social encounters (Lahlou, 2008). Those faces which can take numerous values, for example, “Father”, “Employee”, “Patient, “Dancer”, “Tourist” etc. can be put on or dropped by social actors in order to perform appropriate public performances.

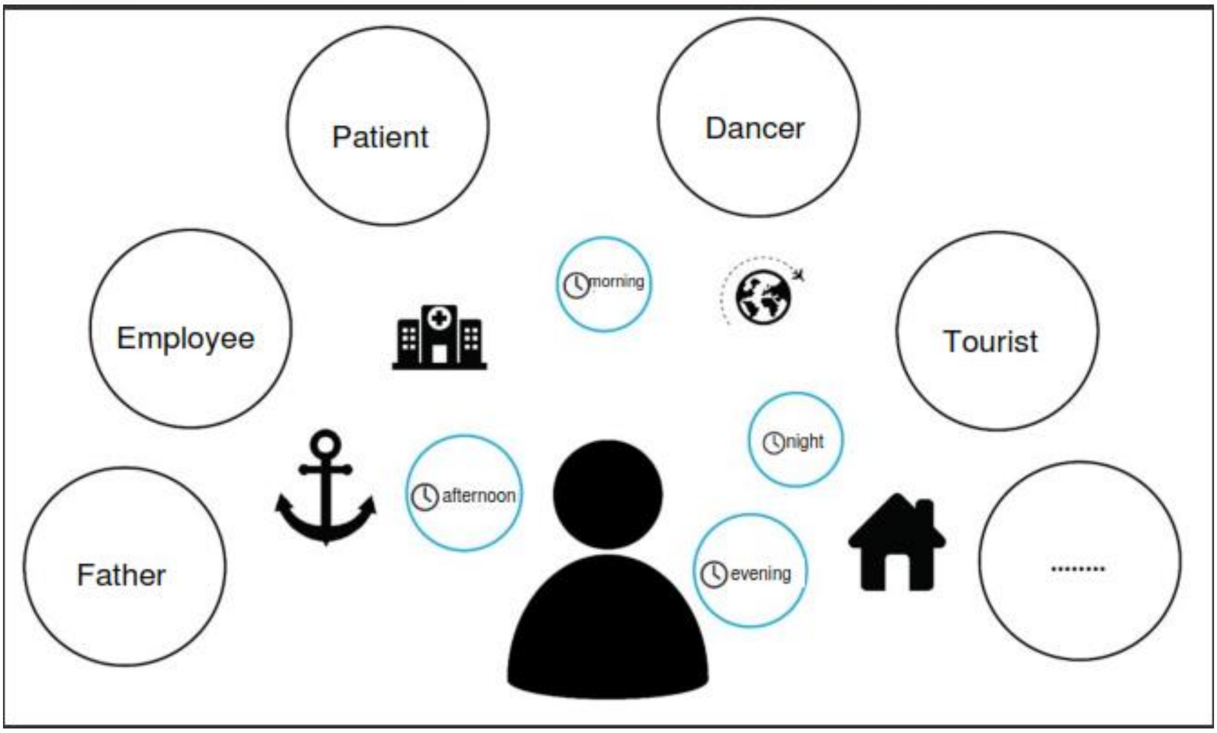


Figure 2. The User’s Face in Different Frames and Stages, by Georgia - Aikaterini Mavroeidi, 2018

As shown in Figure 2 the notion of time, “Morning”, “Afternoon”, “Evening” and “Night” seems to play an important role in placing a social actor in a specific stage. Finally, frames characterize possible locations when social actors tend to behave in an almost definite way, i.e. “Home”, “Port”, “Hospital” or “Travel Agency”. For example, a social actor who “wears” the “Father” face may perform an appropriate public performance by driving his child to school, or a social actor who “wears” the “Dancer” face by enrolling to the next dancing festival of his town.

In addition, public performances are also placed at a specific setting, “Frame” and time. The combination of frame and time creates specific stages, where social performances are to take place. In other words, the face “Father” could be utilized in the frame “Home” in the morning as far as the first example is concerned, while the face “Dancer” can be used in the frame “Dancing School” in the evening for the second example.

In any case, faces are considered to be distinctive from one another, as social actors tend not to share details of their private life in one face to any other of their additional faces (Lahlou, 2008). An example of that could be a social actor utilizing the face of “Employee” and the face of "Patient" in different stages. This social actor would not be willing to share any details of his medical record during a business meeting.

5. DISCUSSION / CONCLUSIONS

This paper sheds light upon a relatively new field of study, Social Software Engineering. The main point of this paper is to broaden understanding on users' digital identities and the impact of geolocation data through this interdisciplinary approach. The aforementioned conclusions of the 1st International Workshop on Social Software Engineering and Applications present the principles of Social Software Engineering.

This dimension will provide valuable information on the setting of requirements elicitation and modelling approaches for designing privacy-aware systems. The social parameters and social criteria which are indicative for setting requirements and designing software, putting emphasis on privacy issues as well, are to be further explored, focusing on users' social identity and geolocation data in social media online practice.

Another key point of this paper is that both digital identities and geolocation concerns have been addressed while drawing on Social, Information and Communication and Geolocation Theories. The example of geolocation provided a deeper understanding on users' special characteristics of social identity, namely Faces, Frames, Stages and Performances. At the same time, notions, such as Predictability and Tracking Users' Normativity were introduced so as to raise awareness in the design of privacy-aware information systems. Taking into consideration the number of user's faces in different frames and stages, this paper also proposes an interdisciplinary analysis based on adaptive privacy and security systems which will incorporate social identity variables during the setting of their requirements.

As a result, some of the main challenges of the future research will be to specify the changes in users' needs both in different social settings and around the clock in a way that system requirements would provide an appropriate privacy and security protection based on users' social characteristics. Social characteristics should be defined and analyzed in order to specify their characteristics and grade their importance. Another key aspect would be bridging the social and the technical part while at the same time matching the social characteristics of the digital identity to the technical, functional and non-functional requirements in order to address future research objectives.

Last but not least, future research on this topic will contribute to designing and proposing a user friendly and socially-aware methodology for designing privacy-aware information systems.

REFERENCES

Abbattista, F., Calefato, F., Gendarmi, D., & Lanubile, F. (2008). Incorporating social software into distributed agile development environments. In *Proceedings of the 23rd IEEE/ACM International Conference on Automated Software Engineering* (pp. II–46). IEEE Press. Retrieved from http://dl.acm.org/citation.cfm?id=3107639

Balicki, J., WSEAS (Organization), International Conference on Artificial Intelligence, K. E. and D. B., IEEE International Conference on Fuzzy Systems, & IEEE International Conference on Neural Networks (Eds.). (2014). *Advances in neural networks, fuzzy systems and artifical intelligence: Proceedings of the 13th International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases (AIKED '14) ; Proceedings of the 15th International Conference on Fuzzy Systems (FS '14) ; Proceedings of the 15th International Conference on Neural Networks (NN '14) : Gdansk, Poland, May 15-17, 2014*.

Barkhuus, L., & Dey, A. K. (2003). Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns. In *Interact* (Vol. 3, pp. 702–712).

Beresford, A. R., & Stajano, F. (2003). Location privacy in pervasive computing. *IEEE Pervasive Computing*, *2*(1), 46–55. https://doi.org/10.1109/MPRV.2003.1186725

Bettini, C., Wang, X. S., & Jajodia, S. (2005). Protecting privacy against location-based personal identification. In *Workshop on Secure Data Management* (pp. 185–199). Springer.

Consolvo, S., Smith, I. E., Matthews, T., LaMarca, A., Tabert, J., & Powledge, P. (2005). Location disclosure to social relations: why, when, & what people want to share. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 81–90). ACM.

Cramer, H., Rost, M., & Holmquist, L. E. (2011). Performing a check-in: emerging practices, norms and'conflicts' in location-sharing using foursquare. In *Proceedings of the 13th* *international conference on human computer interaction with mobile devices and services* (pp. 57–66). ACM.

Duckham, M., & Kulik, L. (2005). A Formal Model of Obfuscation and Negotiation for Location Privacy. In *Pervasive Computing* (pp. 152–170). Springer, Berlin, Heidelberg. https://doi.org/10.1007/11428572_10

Duckham, M., & Kulik, L. (2006). Location privacy and location-aware computing. *Dynamic & Mobile GIS: Investigating Change in Space and Time*, *3*, 35–51.

Fabio Massacci, Marco Prest and Nicola Zannone. (200AD). USING A SECURITY REQUIREMENTS ENGINEERING METHODOLOGY IN PRACTICE: THE COMPLIANCE WITH THE ITALIAN DATA PROTECTION LEGISLATION.

Jenkins, R. (2008). *Social identity* (3rd ed). London ; New York: Routledge.

Lahlou, S. (2008). Identity, social status, privacy and face-keeping in digital society. *Social Science Information*, *47*(3), 299–330. https://doi.org/10.1177/0539018408092575

Liu, L. (2007). From data privacy to location privacy: models and algorithms. In *Proceedings of the 33rd international conference on Very large data bases* (pp. 1429–1430). VLDB Endowment.

Miguel, C., & Medina, P. (2011). The Transformation of Identity and Privacy through Online Social Networks (The CouchSurfing case). Retrieved from http://eprints.leedsbeckett.ac.uk/2159/1/The%20Transformation%20of%20Identity%20and%20Privacy_The%20CouchSurfing%20Case.pdf

Myles, G., Friday, A., & Davies, N. (2003). Preserving privacy in environments with location-based applications. *IEEE Pervasive Computing*, *2*(1), 56–64.

Nario-Redmond, M. R., Biernat, M., Eidelman, S., & Palenske, D. J. (2004). The Social and Personal Identities Scale: A Measure of the Differential Importance Ascribed to Social and Personal Self-Categorizations. *Self and Identity*, *3*(2), 143–175. https://doi.org/10.1080/13576500342000103

Nicola Zannone, Fabio Massacci, & and John Mylopoulos. (2006). *Security Requirements Engineering Methodologies.ppt*. Department of Information and Communication Technology University of Trento.

Schwartz, R., & Halegoua, G. R. (2015). The spatial self: Location-based identity performance on social media. *New Media & Society*, *17*(10), 1643–1660.

Shokri, R., Theodorakopoulos, G., Boudec, J. Y. L., & Hubaux, J. P. (2011). Quantifying Location Privacy. In *2011 IEEE Symposium on Security and Privacy* (pp. 247–262). https://doi.org/10.1109/SP.2011.18

Snekkenes, E. (2001). Concepts for Personal Location Privacy Policies. In *Proceedings of the 3rd ACM Conference on Electronic Commerce* (pp. 48–57). New York, NY, USA: ACM. https://doi.org/10.1145/501158.501164

Tang, K. P., Lin, J., Hong, J. I., Siewiorek, D. P., & Sadeh, N. (2010). Rethinking location sharing: exploring the implications of social-driven vs. purpose-driven location sharing. In *Proceedings of the 12th ACM international conference on Ubiquitous computing* (pp. 85–94). ACM.