universität wien

MASTERARBEIT

Titel der Magisterarbeit

"The Effect of Social Networking on Privacy Management"

Verfasser

Leonidas-Dimitrios Perellis-Konstantinidis

angestrebter akademischer Grad
Magister der Sozial- und Wirtschaftswissenschaften (Mag. rer. soc. oec.)

Wien, 2011

Studienkennzahl lt. Studienblatt: A066922
Studienrichtung lt. Studienblatt: Informatikmanagement
Betreuer: Univ.-Prof. Dipl.-Ing. DDr. Gerald Quirchmayr

THANKS

Before I begin, I would like to thank Prof. Gerald Quirchmayr who, throughout the course of my studies, showed me the relationship that an academic teacher ought to have with his students. Without his contribution and moral support this work would never have managed to escape from its strict technological framework, nor reach its interdisciplinary potential.
I would also like to thank the amazing friends who took care of me through my time in Vienna; guys, you are the best.
Finally, I would like to thank my mother, who has always been trying to do her best for me. During my long journey through Academia she has stood by me to support me, advise me and keep my spirits high. Thank you mother.

Leonidas

Contents

1 Scope and Background of the Thesis 7
1.1 Social Networks and Threats to Privacy in SN Sites 7
1.2 Privacy and Privacy Management 8
1.2.1 Challenges to Privacy Management 9

2 Goals and Planned Contribution 11
2.1 How Online Privacy Management is changing 11
2.2 What is this Thesis aimed at 12
2.2.1 Contribution of the Thesis 12

3 Studying Social Networks 15
3.1 Dominating the latest News 15

4 Foundations that we can build on 19
4.1 Changes in the Use and Perception of Facebook 20
4.1.1 The Perception of Audience in Facebook 21
4.1.2 The Attitude of the Users is changing 23
4.2 The Current Legal Situation in Europe 25
4.2.1 Privacy Protection in Contemporary Society 25
4.2.2 New Technologies and Legal Privacy Debates 26
4.2.3 Efficiency and Security VS Privacy 26
4.2.4 Security and Privacy Legal Framework in Europe 28
4.2.5 Summarizing 28

5 Bridging the Gap 31

6 Identifying Core Requirements 35
6.1 Requirements Identification and Collection 35
6.2 Requirement Analysis 36

7 Structuration Theory and its contribution to this Thesis 39
7.1 The Structure and the Agency 39
7.2 Adaptive Structuration Theory 40
7.2.1 The Concept of Appropriation 40
7.3 Appropriation of Privacy Management in Social Networks 41
7.3.1 Deriving Scales from Appropriation Moves 42
7.4 Testing PM Appropriations in Actual SN Sites 43
7.4.1 Testing Appropriation Measures in Facebook and MySpace 44
7.4.2 Testing Appropriation Measures in StudiVZ (Austria) 44
7.5 Results and Findings of Appropriation Testing 46
7.5.1 Relationships between Familiarity and Use Measures 47

8 Collective Privacy Management 49
8.1 Data Co-ownership in Social Networking Environments 50
8.2 A Collective Privacy Management Algorithm 52
8.2.1 Credit Bargaining in Privacy Contexts 52
8.2.2 Privacy as a Tax Problem 54
8.2.3 Truthfulness and the Importance of Clarke-Tax 55
8.3 Inference Logic in Privacy Reasoning 56
8.4 Experiments and Results on Collaborative PM 57

9 The Combined Model 59
9.1 The Main Concept 59
9.2 Design of the Combined Model 60
9.3 What is to be expected from the Model 61
10 Achievements and Limitations 63
11 Summary and Conclusions 65
12 Appendix 67
12.1 Zusammenfassung (Deutsch): 67
12.2 Abstract (English): 68
12.3 Curriculum Vitae 69
Catalog of Images and Tables 73
Bibliography 75

1 Scope and Background of the Thesis

Social Networking (in short, SN) is undoubtedly one of the major technological phenomena of the new era of Web 2.0, leaving all other features way behind. Social networks enable a form of self expression for hundreds of millions of people, help them socialize and get to know each other and, more importantly in our case, share personal content among themselves. However, despite the fact that content sharing happens to be one of the main features of the prominent SN sites, the latter do not yet seem to support notable Privacy Management mechanism for sharing this sensitive data. User participation in online communities, social networking sites and media-sharing platforms expand for multiple years, during which time, the systems can undergo radical redesign. At the same time, user populations may change and the individual users' social context may evolve. This potential is inherent to long running social computing sites and can affect how members of a site use and perceive it. It has been studied how use changes over time in social computing environments, including early work on Multi-User Domains [1], online discussion forums [2], open-source software [3] and content creation communities [4]. Evidently, a particular type of multi-user platform that has uniquely succeeded in the last years is the Social Networking site [5].

1.1 Social Networks and Threats to Privacy in SN Sites

Although the body of research related to SN sites has been constantly growing over the past several years, no change in the use of these sites has been actually noticed. Boyd and Ellison [6] define three main characteristics of SN sites: such sites allow users to "(1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system."

Social Networking sites including Friendster.com, MySpace, Facebook, and LinkedIn have been widely spread over the Internet during the first decade of this new millennium. SN sites have been very successful in attracting new users, as they offer them a form of self-expression and help them interact and socialize with each other in a world where social contact becomes more and more scarce and difficult. Users of those sites are given the option of designing their personal profiles and customizing them according to their wishes. Through these sites, the users can engage in a plethora of activities, many of which include entertainment, business, and/or knowledge sharing. The commercial success of an SN site depends on the number of users it attracts, while at the same time, it is vital for it to encourage them to bring more users to the network and to share data with them within its social environment. However, end users are often not properly aware of the size or the nature of the audience with access to their data, while the sense of intimacy created by being among digital friends often leads to disclosures that may not be appropriate in a public forum such as this.

---

[1] Churchill, E.F. and Bly, S. (1999) "Virtual Environments at Work: Ongoing use of Muds in the Workplace," WACC'99, ACM Press, 99-108, San Francisco

[2] Smith, M. "Measures and Maps of Usenet." In Lueg, C. and Fisher, D. eds (2002) "From Usenet to CoWebs: Interacting with Social Information Spaces." Springer Verlag, New York, NY

[3] Lakhani, K.R. and Hippel, E. (2003) "How Open Software Works: "Free" User-To-User Assistance." Research Policy, 32(6), 923-943

[4] Bryant, S., Forte, A. and Bruckman, A. (2005) "Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia." ACM-GROUP, Sanibel Island

[5] Boyd, D. (2004) "Friendster and Publicly Articulated Social Networks." Conference on Human Factors and Computing Systems (CHI 2004), ACM, April 24-29, Vienna

[6] Boyd, D. and Ellison, N. (2007) "Social Network Sites: Definition, History, and Scholarship." Journal of Computer-Mediated Communication, 13(1), Article 11

Lampe, Ellison and Steinfield[7] have shown that Facebook users connect mainly with people with whom they have already had a previous relationship in the real world, and that they expect (a concept the call "perspective of audience") that they are being observed by their peers rather than by non-peers, i.e. employers, law enforcement agencies etc. This, doubtlessly, creates the illusion of a safe environment. Nevertheless, even if two users know each other, their social relationship does not often imply that they have the same privacy preferences. The average number of friends of MySpace users is 115, which indicates that the friend relationship is being stretched to cover a wide range of intimacy level[8]. Such an exposure of data introduces SN users to a multitude of privacy risks[9].

An additional, yet highly significant, threat that should be considered in this context comes as a result of the alarming increase on the amounts of media content that is being uploaded daily by SN users on their online social profiles. As it has already been mentioned already, these digital images and videos are an integral and rather popular part of the very functionality of these SN sites. In an attempt to be more factual, here are some statistics: As of October 14, 2008, Facebook hosts 10 billion user photos, serving over 15 million photo images per day[10]. These pictures may be tied to the users profile that posted them but they are often, either explicitly (through specific tagging) or implicitly (through simple recurrence), connected to other user profiles [11], and thus to autonomous individuals. Such pictures are made available for other SN users, who can view, add comments and add hyperlinks to indicate the users who appear in the pictures, by using content annotation techniques. It is highly important for someone to notice that, in current SN sites, a picture uploaded by a user is not required to have permissions from other users appearing in the photo, even if they are explicitly identified through tags or other metadata. Although most social networking and photo sharing websites provide mechanisms and default configurations for data sharing control, they are usually simplistic and coarse-grained. Pictures, or in the more general case, data, are usually controlled and managed by single users who may not be the actual stakeholders, thus letting this way serious privacy concerns to be raised. Data stakeholders may be completely unaware of the fact that others are managing data that is related to them. And even when the stakeholders are aware of the fact that their data is managed and controlled by other individuals (transparency), they have limited control over it and cannot influence the privacy settings applied to this data. The discrepancy related to privacy as a result of little or no access control of shared data in Web 2.0 at all, is well documented in the public news media[12].

1.2 Privacy and Privacy Management

As Samuel D. Warren and Louis D. Brandeis first attempted to define back in 1890, Privacy is first and foremost perceived as an individual's right "to be let alone".
From their work, more than a century ago, we read:

---

[7] Lampe, C., Ellison, N. and Steinfield, C. (2006) "A Face(book) in the Crowd: Social Searching VS. Social Browsing." ACM Special Interest Group on Computer-Supported Cooperative Work, ACM Press, Banff, Canada

[8] Hart, M., Johnson, R. and Stent, A. (2007) "More Content – Less Control: Access Control in the Web 2.0." IEEE Web 2.0 Privacy and Security Workshop

[9] Hobgen, G. (2007) "Security Issues and Recommendations for Online Social Networks." ENISA Position Paper N.1

[10] Beaver, D. (14 Oct 2008) "Ten Billion Photos." Facebook Engineering Blog: http://www.facebook.com/note.php?note_id=30695603919

[11] Acquisti, A. and Gross, R. (2006) "Imagined Communities: Awareness, Information, Sharing and Privacy on the Facebook." 6th Workshop on Privacy Enhancing Technologies, 36-58, Springer, Cambridge, UK

[12] Rosenblum, D. (2007) "What Anyone Can Know: The Privacy Risks of Social Networking Sites." IEEE Security and Privacy, 5(3), 40-49

> "Recent inventions and business methods call attention to the next step which must be taken for the protection of the person, and for securing to the individual what Judge Cooley calls the right "to be let alone". Instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life; and numerous mechanical devices threaten to make good the prediction that "what is whispered in the closet shall be proclaimed from the house-tops." For years there has been a feeling that the law must afford some remedy for the unauthorized circulation of portraits of private persons; and the evil of invasion of privacy by the newspapers, long keenly felt, has been but recently discussed by an able writer. […]"

They continue:

> "Of the desirability – indeed of the necessity – of some such protection, there can, it is believed, be no doubt. The press is overstepping in every direction the obvious bounds of propriety and of decency. Gossip is no longer the resource of the idle and of the vicious, but has become a trade, which is pursued with industry as well as effrontery. To satisfy a prurient taste the details of sexual relations are spread broadcast in the columns of the daily papers. To occupy the indolent, column upon column is filled with idle gossip, which can only be procured by intrusion upon the domestic circle. The intensity and complexity of life, attendant upon advancing civilization, have rendered necessary some retreat from the world, and man, under the refining influence of culture, has become more sensitive to publicity, so that solitude and privacy have become more essential to the individual; but modern enterprise and invention have, through invasions upon his privacy, subjected him to mental pain and distress, far greater than could be inflicted by mere bodily injury."[13]

It is of great interest to point out the similarities – as well as the differences – in the special needs that made lawmen of a time so different than ours to define and protect the notion of Privacy. It is rather obvious that Privacy Management is not a new challenge for our society, created through the establishment of the online world – on the contrary, it has been something that legislators have been concerned about for more than a century.
Nevertheless, claiming that offline and online Privacy Management share a lot in common would be an overstatement. The needs for Privacy may not have changed a lot, but the rules of the online world make the field of applicability rather demanding.
From a legislative point of view, privacy in online Social Networks poses unique challenges, far more specific and complicated than those posed by online privacy in general. This is mostly because users provide the greatest bulk of their information on these networks on their own initiatives (which thus, can be treated as their own consent). However, traditional privacy laws are based on "informed consent" and protect users against unfair or disproportional data collection and application thereof by the websites, and would therefore be rather ineffective in today's new arena. Nevertheless, it is this brave new world's challenges that we are now facing, and the question that arises, is how to do so most effectively.

1.2.1 Challenges to Privacy Management

Before begin to analyze the issues we plan to present in the following chapters, we need to clarify the term of Privacy Management, together with some of the challenges that PM is up against in contemporary (and future) society. One should not forget that, within offline social spaces, privacy management is an active part of everyday life, influencing where, when and to whom we decide to reveal private information. In the same way that technology has affected the nature of communication, that technology mediation of social interaction will change the nature of privacy management.

[13] Samuel, D.W. and Louis, D.B. (1890) "The Right to Privacy." Harvard Law Review, 15 Dec 1890

Privacy management is an essential social skill found in cultures around the world[14]. It is a fundamental component of what the sociologist Erving Goffman called impression management, or the presentation of versions of the self to different audiences[15]. As a result of Goffman's seminal works, we know the strategies people employ for offline privacy management, as well as the critical role it plays in maintaining different levels of social connections. Privacy management consists of social, relational, cognitive and perceptual components that a person constantly monitors in real time. These components are input into an individual privacy calculus that controls the development of boundaries and the disclosure of confidences.

While interaction moves more and more into online social spaces, there has been a rising public debate about the inadequacies of online privacy management tools. An abundance of proof can found in the related literature[16,17,18,19]. While there has been discussion in academic research about the complexities of online privacy management[20], there has been little usable research on how to develop reliable, repeatable Measures of SN use[21]. The development of such Measures would be of great assistance in determining the correct variables and their values, as well as answering questions such as: To what extent are participants in online social spaces aware of the privacy management tools available? To what extent do members of social technologies actually use these privacy management tools?

[14] Petronio, S. (2002) "Boundaries of Privacy: Dialectics of Disclosure." State University of New York Press, Albany

[15] Goffman, E. (1959) "The Presentation of Self in Everyday Life." Doubleday & Co., Garden City, NY

[16] Hass, N. (2006) "In your Facebook.com" The New York Times, 8 Jan 2006

[17] Hempel, J. (2005) "The MySpace Generation." Business Week, 12 Dec 2005

[18] Maag, C. (2007) "When the Bullies Turned Faceless." The New York Times, 16 Dec 2007

[19] Read, B. (2006) "Think Before you Share: Students' Online Socializing Can Have Unintended Consequences." The Chronicle of Higher Education, A38, 20 Jan 2006

[20] Barnes, S. B. (2006) "A Privacy Paradox: Social Networking in the United States." First Monday, 11(9)

[21] Acquisti, A. and Gross, R. (2006) "Imagined Communities: Awareness, Information, Sharing and Privacy on the Facebook." 6th Workshop on Privacy Enhancing Technologies, 36-58, Springer, Cambridge, UK

2 Goals and Planned Contribution

In order for a useful construct to be produced by the end of this research, a finite set of goals must be clearly defined. It is obvious by now, that the Privacy Management issues that evoke from the evolution of SN sites need to draw the attention of the scientific community. A scientific approach is thus necessary.

2.1 How Online Privacy Management is changing

Privacy management in SN sites is proven to be a complex issue for both users and administrators. The social concept of Privacy itself has been a "hot potato" for public figures and lawmen alike, and it has been interpreted differently over the centuries based on cultural or personal perspectives[22]. Research has shown privacy to be a multi-dimensional construct[23].

A comparison of a typical SN member's motivation towards privacy management versus the nature of said privacy management in the online world proves them to be in direct conflict. SN sites work hard to create tools that support the ability to express oneself through a profile. This results in more active engagement with the site and its members. However, privacy management tools are designed to share less information with a smaller audience.

A main goal in online self-presentation within SN sites is to create a rich, authentic profile that keeps friends up to date on your activities and presents an interesting personality to potential new friends[24]. Privacy management, in its sense, consists of a collection of settings that either restrict what information is available or restrict the scope of the audience. It does not seem possible to present a rich, authentic digital profile while, at the same time, offering effective privacy management. This is because of the following issues:

- Privacy management works by limiting information, especially that which is potentially sensitive. This results in a profile that looks more like a resume than something that would spark the interest of others.
- Young consumers value honesty and authenticity, and can easily spot insincerity. They have had enough of old-style marketing, and value something that is "real".
- Privacy management works by limiting the potential audience for your profile. Such a strategy may protect privacy, but will also have a negative effect on the opportunity of users developing new relationships – or even rekindling old ones. It is obvious that, such an approach will miss the actual targets.

By this analysis, there is a conflict between the goals of creating an interesting profile and practicing faithful privacy management. However, experience shows that studying certain aspects of human nature – in our case the way SN community member adapt and adopt new technology, and specifically, privacy management – can very well help us devise methods into resolving this conflict.

There exists a fundamental mismatch between online privacy needs and privacy management functionality that can hopefully be explained more clearly later in this study. One can only hope that future applications will acknowledge this mismatch, as this can be the singular first step in overcoming it.

---

[22] Lessig, L. (1998) "The Architecture of Privacy."

[23] Smith, H.J., Milberg, S. and Burke, S. (1996) "Information Privacy: Measuring Individuals' Concerns about Organizational Practices. " MIS Quarterly, 20(2), 167-196

[24] Boyd, D. (2006) "Identity Production in a Networked Culture: Why Youth *heart* MySpace." American Association for the Advancement of Science (http://www.danah.org/papers/AAAS2006.html, accessed on 28 Mar 2011)

2.2 What is this Thesis aimed at

In this chapter, I would like to put into words what led me to choose this specific field of study for my final and most extensive personal work as a master student. Through the years of studying as a computer scientist, I became extensively interested in system analysis, knowledge engineering and e-government issues, but I always found myself to return to the familiar sites of security and human interaction. At the same time, I soon realized that interdisciplinary research concerning major issues of contemporary society needed direct help from the IT community more than ever before.

The aim of my work was first and foremost the acquisition of security awareness. I felt – and still do – that, no matter how much one believes they know their way around online spaces, they still need always be vigilant. The more I learned in the field of security, the more I realized what a dangerously beautiful world the Internet can be. However, vigilance cannot be the solution for a medium so widespread as the Internet has become – and will become in the future.

It has been said that, the only thing that will be remembered from our time someday, will be the Genesis of the Internet. I often come to believe that – I think the only disagreement among possible debaters on that might be the date of that "someday". Whatever the answer to the debate may be though, a major portion of the world's population is already using the Net and significant parts of our lives are being migrated online, whether we wish it so or not.

The amazing profits and benefits that the Internet has to offer the everyday life of the simple man is a part of another discussion and probably does not need to be discussed at all anymore. However, when the future guides us to an online world of Ambient Intelligence, it is our duty as computer scientist to pave the road and do our best, so that this future world will be a safe and decent one, according to our standards of democracy and morality.

Subsequently, my research interests steered the aim of this study search towards the online areas that gathered the largest masses of people and thus, expected more attention from a humanistic point of view. The issues troubling the online Social Networks drew my attention almost immediately, especially from the angle of Privacy and how it was managed.

2.2.1 Contribution of the Thesis

This Thesis aims at raising the IT community's awareness on the issues of Privacy Management in Social Network platforms. Acknowledging the fact that there are many others who claim that role, it actually aims to raise awareness on other communities too – and was written on that note, particularly. It further aims at collecting all the best related literature on a rather well written form, while promoting best practices on the field of Privacy Management until the time it was constructed. Finally, it aims at adding its own contribution to these practices, by presenting a model that could offer combined merits and advantages from already presented best practices without conflicts.

These three axes are of equal significance throughout the Thesis. Privacy is an issue that affects the fields of law, psychology and sociology probably more than the field of information technology; the fact that this seems to have changed in the last decade is only because of the rapid evolution of the Internet. The related work from the side of IT is immense and scattered and definitely not suitably written for the students of the other disciplines. Thus, it became a very important goal that researchers with a lower level of technical knowledge would equally absorb the issues explained and dealt with in these pages.

Finally, through the process of the research, it became clear that the contribution of this Thesis would not be complete, if it did not attempt to produce a result based on the lessons learned. The idea presented in chapter 9 forms the most practical part of the contribution of this work, as it suggests a model designed to optimize the performance of already tested models[25] that have produced promising experimental results[26]. The aim of

[25] Dwyer, C., Hiltz, S.R., Poole, M.S., Gußner, J., Hennig, F., Osswald, S.,

this Thesis is that, through the design of this model, it will contribute to the promotion of sound Privacy Management designs for Social Networking platforms.
In the following three chapters there will be an extensive overview of the literature supporting this work.

Schließlberger, S. and Warth, B. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." Proceedings of the 43rd Hawaii International Conference on System Sciences, Volume: 1, Publisher: IEEE Computer Society

[26] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

3 Studying Social Networks

The next three chapters will mainly serve as a presentation – and reference – of the multiple types of literature that have been used for the completion of this work. This extensive literature review consists of papers published in the latest years, journals and books' chapters, as well as online articles, symposium announcements, newspaper coverage and others.
The material is divided into a general survey of the big fields of our research (Social Networking Sites, Privacy Management in the Online World, Privacy Management in Social Networking Sites) and a more specific survey of the literature that served as basis for this Thesis. A third chapter is then dedicated to the gap that this work is attempting to fill.

3.1 Dominating the latest News

While the present study is being conducted, more than few articles concerning Social Network Sites have made the front pages of Privacy and IT Security feeds and journals. Heavily dominated by the contemporary giant of SN sites today, Facebook, news are as informative as they are alarming. The present chapter shall attempt to paint the picture, as it looks today.
"You should not worry about privacy issues of your Facebook account", CEO of Facebook, Mark Zuckerberg had assured the world, while announcing changes in Facebook's privacy policy. "Users should be at ease and not worry of their private information being shared with a third party."
It was not before long, when in early 2010 Facebook officials were actually forced to admit that they had been wrong[27], and that privacy policy changes allowed private information to be shared with advertisers and other third parties. Attempts to rectify the mistake were prompt; nevertheless many users abandoned the SN site because of loss of confidence.
Facebook's privacy problems however, seem to resemble a centipede with footwear issues. Before autumn of the same year, connections of Facebook with the Rapleaf profiling scandal[28] led to further slandering of SN sites credibility as far as the protection of privacy is concerned. Of course, when profiling practices get involved, privacy management needs rise to a new level of importance, as the risks to the individuals are considerably higher. One can only imagine what extensive profiling a data aggregator might be able to engage into, if in possession of Facebook's immense databases – full of correlations between names, locations, political and religious beliefs, associations etc.
What seems to be the case is that Facebook, riddled with 550.000+ apps made by scores of different developers, lost control over its data – sensitive user identification data that FB claimed to keep absolutely safe. According to announcement by the company itself, apps were discovered that deliberately mined data and sold it to data brokers. However, the issue at hand is not so much how something like this could have happened as much as what ensued.
Facebook announced the issue to the public, together with the company's intentions to take measurements. They made suspended the guilty developers for 6 month and, at the same time, made "deals" with the data brokers, to erase the data from their storage, claiming at the same time to the world that everything will now be back to normal and ensuring them of their new gained safety.

---

[27] Graham, B. (May, 2010) "Facebook CEO admits that they were wrong." (http://www.latestngadgets.com/facebook-ceo-admits-that-they-were-wrong/4770.html, accessed on 28 Mar 2011)
[28] Cringely, R.X. (Oct, 2010) "Online Advertisers are selling you out." (http://www.infoworld.com/t/social-networking/online-advertisers-are-selling-you-out-811, accessed on 28 Mar 2011)

Nevertheless, no matter how important prevention is considered, one can never argue that misfortunes such as this can forever be avoided. It is thus of great importance to understand, that if SN sites are to be integrated in our social daily routine, we need somebody of higher authority than the companies themselves, to safeguard our rights – and among them, our privacy.
As to whether Facebook should be assigned the role of the victim in this story, it is interesting to observe how things evolve less than a few months later, when a blog post by Jeff Bowen appeared on the platform's Developer Blog[29]. The new post provided step-by-step instructions for the outside developers, on how, by adding a new feature on their app or site hosted by FB, they could coerce or lure the user into providing them with their current address and mobile number. "We are now making a user's address and mobile phone number accessible as part of the User Graph object," Bowen wrote. "Because this is sensitive information, we have created the new user address and user mobile phone permissions. These permissions must be explicitly granted to your application by the user via our standard permissions dialogs."

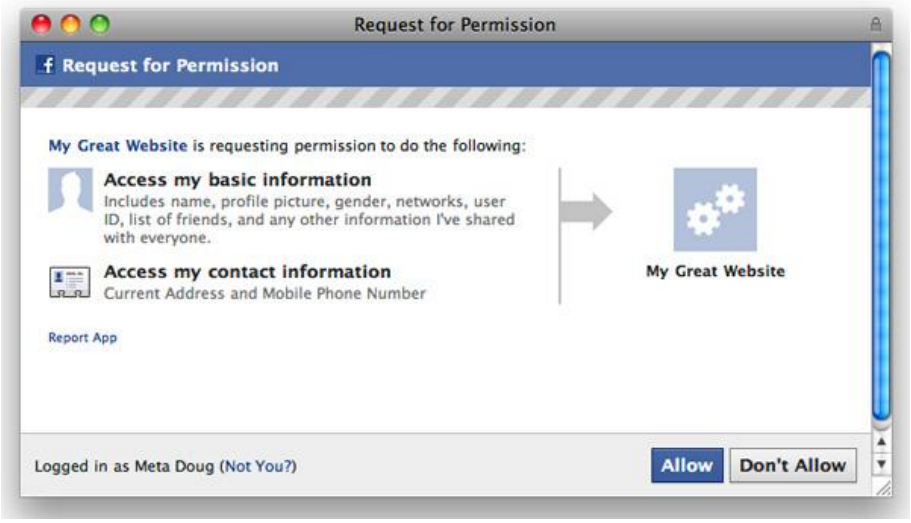


Image 3.1: Facebook Permission Dialog Box

Facebook – through Bowen – admitted that the information is sensitive on the first place. As many prominent SN users have already advocated for, it seems proper that, before even considering implementing such an intrusive feature, Facebook should have made sure that third party applications would not be given such an easy way to scam users out of their addresses and mobile phones.
Apart from general incidents however, one should not underestimate the alarming increase in the number of singular events connected with privacy breaching through SN sites.
Another example (again from Facebook) is the one of a Swiss employee of Nationale Suisse who lost her job as a result of being noticed online in FB by a "friend" while she was taking the day off work because of a migraine[30]. According to the woman, she had only been accessing the net via her iPhone while in bed; when she was noticed by a Nationale Suisse's undercover Facebook operative who, after adding herself in employees' friends lists, had the task to monitor their online activity.
Alone, this incident certainly sounds somewhat paranoid; sadly though, it is accompanied by many of its kind. Companies like Nationale Suisse claim they act "by the book" and

[29] Bowen, J. (Jan, 2011) "User Address and Mobile Phone Number." Facebook Developer Blog (http://developers.facebook.com/blog/post/446/, accessed on 28 Mar 2011)
[30] Haines, L. (Apr, 2009) "Swiss Woman Rolled over Facebook." The Register (http://www.theregister.co.uk/2009/04/27/swiss_rolled/, accessed on 28 Mar 2011)

maybe they are, yet what should concern us is that more than often SN sites are not being used for their original purpose, but rather manipulated for the purposes of profiling and controlling of their users. And it is thoughts like this one that ought to keep us aware and alert of the importance of privacy and privacy management, as well as the general values that stand behind them and need to remain safeguarded.
News similar to those presented to you in the previous pages make an interesting study, especially when paired with announcements such as the one made by Doug Beaver (already mentioned in a previous chapter in respect with Facebook hitting the milestone of 10 billion uploaded photos). Facebook, not only has become the prime photo-trafficking portal in the world, it also constitutes a standing proof, of how the purpose and functionality of the Social Network in general, has changed through the years.
All this, however, would not be of such importance, were it not for the level of connectivity we are brought into through these networks. A picture of what this looks like, can be seen below:

Image 3.2: The World, drawn entirely through Facebook connections

Paul Butler, intern on Facebook's data infrastructure engineering team, uploaded this picture in the company's blog on December 2010[31]. He writes:

> "Visualizing data is like photography. Instead of starting with a blank canvas, you manipulate the lens used to present the data from a certain angle.
>
> When the data is the social graph of 500 million people, there are a lot of lenses through which you can view it. One that piqued my curiosity was the locality of friendship. I was interested in seeing how geography and political borders affected where people lived relative to their friends. I wanted a visualization that would show which cities had a lot of friendships between them. I began by taking a sample of about ten million pairs of friends from Apache Hive, our data warehouse. I combined that data with each user's current city and summed the number of friends between each pair of cities. Then I merged the data with the longitude and latitude of each city.
>
> At that point, I began exploring it in R, an open-source statistics environment. As a sanity check, I plotted points at some of the latitude and longitude coordinates. To my relief, what I saw was roughly an outline of the world. Next I erased the dots and plotted lines between the points. After a few minutes of

[31] Butler, P. (Dec, 2010) "Visualizing Friendships" Facebook Engineering Blog, (http://www.facebook.com/notes/facebook-engineering/visualizing-friendships/469716398919, accessed on 15 May 2011)

> rendering, a big white blob appeared in the center of the map. Some of the outer edges of the blob vaguely resembled the continents, but it was clear that I had too much data to get interesting results just by drawing lines. I thought that making the lines semi-transparent would do the trick, but I quickly realized that my graphing environment couldn't handle enough shades of color for it to work the way I wanted.
>
> Instead I found a way to simulate the effect I wanted. I defined weights for each pair of cities as a function of the Euclidean distance between them and the number of friends between them. Then I plotted lines between the pairs by weight, so that pairs of cities with the most friendships between them were drawn on top of the others. I used a color ramp from black to blue to white, with each line's color depending on its weight. I also transformed some of the lines to wrap around the image, rather than spanning more than halfway around the world. After a few minutes of rendering, the new plot appeared, and I was a bit taken aback by what I saw. The blob had turned into a surprisingly detailed map of the world. Not only were continents visible, certain international borders were apparent as well. What really struck me though was, knowing that the lines didn't represent coasts or rivers or political borders, but real human relationships. Each line might represent a friendship made while travelling, a family member abroad, or an old college friend pulled away by the various forces of life.
>
> Later I replaced the lines with great circle arcs, which are the shortest routes between two points on the Earth. Because the Earth is a sphere, these are often not straight lines on the projection.
>
> When I shared the image with others within Facebook, it resonated with many people. It's not just a pretty picture; it's a reaffirmation of the impact we have in connecting people, even across oceans and borders."

The last sentence is very important. One only need to consider that it did not took but a fraction of Facebook's connections' data to create this pretty picture that shocked all these people. Consequently, it is the ugly and brutal truth that it might have been an innocent social network who happened to collect all the necessary information to manage such an enormous feat of intelligence, yet the purpose and functionality of an institution that collects and handles such amounts of personal data cannot be classified under the simple concept of the SN community anymore. And if it does, we should at least revise our attitude towards it.

4 Foundations that we can build on

These next pages are devoted to those who, before us, have offered their time and energy on those areas and issues closest to the ones that drew our attention. If fact, some of the names mentioned in this chapter, played an inspiring role for us while studying Security and Privacy Issues in academia.

Any decent study on Social Networking sites is bound to start with the history of the area, where expertise meets the names of Boyd and Ellison[32], and continue with a survey of demographics. Through the study of predictive demographics, Hargittai proved that ethnicity and income levels among other factors could affect ones choice of Social Network[33]. On a similar note, the research of Gilbert et al. showed among other things, that rural MySpace users had fewer ties in their networks, than urban users[34].

Every scientist, however, is drawn quickly to the big instances of the concept of his study; a process leading countless of them to Facebook. Gross and Acquisti studied the use of Facebook at Carnegie Mellon University[35] and proved that users at that time were totally unaware about privacy issues in SN in general – and in FB in particular. From another angle, Golder et al. showed how Facebook was becoming increasingly prevalent among its college-aged members between the years 2004 and 2006, through the examination of intra-network messaging and "poking"[36].

Lampe, Ellison and Steinfield, an especially prolific team in the area, showed that Facebook users mainly search for people they already have an offline relationship with. What is even more important, in terms or Privacy Management, is that their expected audience is comprised of peers rather than non-peer members of their networks (professors, administrators) or people outside their networks (law enforcement, employers)[37]. This observation is of major importance, when it comes to privacy awareness of SN members. In other parts of their extensive work, Lampe et al. showed that users who displayed addresses or photos on their profiles were associated with more articulated relationships within the SN[38]. This led them to use Donath's signaling framework[39] to analyze Facebook's profile elements, distinguishing signals that identified assorted types of users. Ellison et al. continued their social experiments with Facebook and studied the connection between the users and their social capital (benefits received from individuals in one's network) in a sample of college students[40]. What they found

---

[32] Boyd, D.M. and Ellison, N. (2007) "Social network sites: Definition, history, and scholarship." Journal of Computer Mediated Communication, 13 (1), Article 11

[33] Hargittai, E. (2007) "Whose Space? Differences Among Users and Non-users of Social Network Sites." Journal of Computer Mediated Communication, 13 (1), Article 14

[34] Gilbert, E., Karahalios, K. and Sandvig, C. (2008) "The Network in the Garden: An Empirical Analysis of Social Media in Rural Life." ACM Conference on Human Factors in Computing Systems (CHI), Florence, Italy, 1603-1612

[35] Gross, R. and Acquisti, A. (2005) "Information Revelation and Privacy in Online Social Networks." Workshop on Privacy in the Electronic Society, Alexandria, VA, ACM Press

[36] Golder, S., Wilkinson, D. and Huberman, B.A. (2007) "Rhythms of Social Interaction: Messaging within a Massive Online Network." 3rd International Conference on Communities and Technologies (CT2007), East Lansing, MI, Springer

[37] Lampe, C., Ellison, N. and Steinfield, C. (2006) "A Face(book) in the Crowd: Social Searching VS. Social Browsing." ACM Special Interest Group on Computer-Supported Cooperative Work, ACM Press, Banff, Canada

[38] Lampe, C., Ellison, N. and Steinfield, C. (2007) "Profile Elements as Signals in an Online Social Network." ACM Conference on Human Factors in Computing Systems (CHI), San Jose, CA

[39] Donath, J.S. (2007) "Signals in Social Supernets." Journal of Computer Mediated Communication, 13 (1), 12

[40] Ellison, N., Steinfield, C. and Lampe, C. (2007) "The Benefits of Facebook "Friends:"

was that certain types of Facebook use were associated with higher levels of social capital perhaps because the site allowed users to maintain broader sets of weak ties in their social networks.

Last but not least, the work of DiMicco and Millen approaches the subject from an interesting point of view. Their work focuses on how Facebook users, while migrating from the college environment to corporate constructs, employed various strategies/attitudes in respect to their Facebook profiles – from erasing all information and making new profiles that suited their new context, to double-profiling, to doing absolutely nothing[41]. Evidence for the significance of their field of study can be witnessed in reports produced from simple questionnaires; such as those used for the work of Lampe et al. One user report reads:

> "I've had a lot of people just say, or adults say people are using Facebook now as another tool for interviewing and stuff like that, so I wouldn't want a picture of me on Facebook to hinder me from getting a job."

While another states:

> "I've heard rumors – many people have told me that employers and people – admission committees look at your Facebook profiles and see what you put in them. And any pictures of me at a party, I've untagged myself in. I don't really want to convey a message of – which I'm not a big partier at all – but I just don't want somebody getting the wrong impression."

4.1 Changes in the Use and Perception of Facebook

Our study owes its inspiration – among others – in the work of Lampe, Ellison and Steinfield, and their attempt to identify the process of change within the online social communities. Lampe et al. drew detailed reports and came up with useful results after consecutive years of empirical studies, focused mainly on the giant of the SN sites today, Facebook[42]. Instead of approaching the area of SN through another field, their work focuses directly in dealing with the several crucial questions that arise from the issues that have been presented in the previous pages.

"How has reported use of Facebook to interact with other members changed over time?" "How has the perception of audience on Facebook changed over time?" "How have the attitudes of users towards Facebook changed over time?" Lampe et al. realized, that in order to explore Facebook participation over time one has to examine the types of uses people report they engage in. Through their extensive observations in 2006, they found that Facebook users were, in general, articulating their existing offline networks, rather than creating new relationships online[43]. While they continued they surveys, Joinson[44] showed that people had heterogeneous patterns of use for different features of Facebook. Consequently, Lampe et al. became interested in how people describe their use of Facebook to make connections: whether they are searching for people online to form a

---

Social Capital and College Students' Use of Online Social Network Sites." Journal of Computer Mediated Communication, 12 (4), Article 1

[41] DiMicco, J.M. and Millen, D.R. (2007) "Identity Management: Multiple Presentations of Self in Facebook." Conference on Supporting Group Work, Sanibel Island, FL, ACM Press, 383-386

[42] Lampe, C., Ellison, B.N. and Steinfield, C. (2008) "Changes in Use and Perception of Facebook." Proceedings of the '08 ACM Conference on Computer Supported Cooperative Work (CSCW), ACM, NY

[43] Lampe, C., Ellison, N. and Steinfield, C. (2006) "A Face(book) in the Crowd: Social Searching VS. Social Browsing." ACM Special Interest Group on Computer-Supported Cooperative Work, ACM Press, Banff, Canada

[44] Joinson, A.N. (2008) "Looking at, looking up or keeping up with People?: Motives and Use of Facebook." ACM Conference on Human Factors in Computing Systems (CHI), Florence, Italy, 1027-1036

relationship with, or claim to be articulating their offline networks in an online environment.
Another question Lampe et al. focused on was whether this trend changed as time went by. Additionally, they focused on whether any of the observed changes where because populations were altering their behavior or because new members entering the SN had different behavioral patterns. All these data could form new norms on how new users entering a site like Facebook might engage in different behaviors than veteran users, and, eventually, how the very use and perception of Facebook might itself alter. It became obvious, that the addition or removal of features within the SN affected the user experience[45]. Over the time period reported in this study, Facebook had added many new features and some of these features were designed to affect social patterns on the site.

4.1.1 The Perception of Audience in Facebook

Nevertheless, it was the concept of "Perception of Audience" that took most of Lampe et al.'s interest throughout their work. What, simply put, is defined to be the user's notion of readers and/or listeners of his activity on the SN community has been a central theme for CSCW research in the past. The constrained information channels that restrict knowing your audience have led to innovations in making audience visible[46] and research on the possible benefits of "lurkers"[47]. As we have already mentioned, according to Lampe et al.[48], users who were asked who they thought had seen their Facebook profile, reported in general that their "perception of audience" was comprised of peers, and was much less likely to include non-peers.
Since that time, two changes have occurred which might influence users' perceptions of audience. First, in 2006 Facebook introduced a significant change to the interface of the site: a "News Feed" which tracked changes to Friends' profiles and aggregated them in one, highly visible place. This window into peers' activities may have made users more aware of the visibility of their own online activities, thus prompting changes in perceptions of audience (and, perhaps, privacy settings). Additionally, a number of popular press stories focused attention on Facebook use, as did University responses (such as guidance about online self-presentational strategies) to Facebook use by students. These changes in context could affect how Facebook users perceive their audience. Changes in perception of audience may affect how users behave within the site. If they see their audience as more public, they may disclose less about themselves or become more dissatisfied with their use of the site.
Table 4.2 suggests that the students of Michigan State University (where Lampe et al. conducted most of their surveys) are changing their Perception of Audience over time – although not always in the most obvious ways. The X^2 column shows the degree of statistical change between each year – a higher number indicates a bigger statistical difference.

[45] Lampe, C., Ellison, B.N. and Steinfield, C. (2008) "Changes in Use and Perception of Facebook." Proceedings of the '08 ACM Conference on Computer Supported Cooperative Work (CSCW), ACM, NY
[46] Erickson, T. and Kellogg, W.A. (2002) "Social Translucence: Designing Systems that Support Social Processes." ACM Human-Computer Interaction in the New Millennium, NY, 325-345
[47] Nonnecke, B., Preece, J. and Andrews, D. (2004) "What Lurkers and Posters think of each other." 37th Hawaii International Conference on System Sciences, IEEE, HI
[48] Lampe, C., Ellison, N. and Steinfield, C. (2006) "A Face(book) in the Crowd: Social Searching VS. Social Browsing." ACM Special Interest Group on Computer-Supported Cooperative Work, ACM Press, Banff, Canada

	2006	2007	2008	X^2
My high school friends	90%	86%	94%	25.31***
Friends other than HS friends	84%	81%	87%	5.92*
People in my classes	84%	78%	83%	5.15
Someone I met at a party or social event	73%	70%	72%	0.88
Total strangers from MSU	74%	57%	55%	28.73***
Family members	49%	54%	70%	39.58***
Total strangers from other campuses	35%	30%	28%	3.98
Total strangers who aren't affiliated with any college or school	14%	22%	24%	10.97**
My MSU professors	12%	15%	15%	1.56
Law enforcement	6%	7%	6%	0.52
Future employers	N/A	13%	18%	53.90[3]

Table 4.1: Responses to the question: "Since you have created your profile, who do you think has looked at it?" over three consecutive years' surveys[49]

This table shows very interesting changes. In 2007 and 2008 people were asked whether they felt future employers had viewed their profiles. The percentage that answered in the affirmative increased significantly between 2007 and 2008, though stayed relatively low as a whole (13% and 18%, respectively). Concerning the statement "Facebook is a student-only space" respondents in 2007 had a mean score of 3.11 with standard deviation of 1.27, when in 2008 the mean response was 2.83 with a standard deviation of 1.18. Agreement went down significantly between those two periods ($t=3.14$, $p<.01$), indicating there was a change in perception about the overall audience of the site. However, even in 2008 the mean response is relatively high, given the increasing population of non-students on Facebook, and the announcements about changes in membership in the media. Last but not least, in the interviews, respondents discussed the fact that employers might be looking at their profile and the source of this impression, which came from a variety of sources including peers, potential employers and university officials.

While studying Lampe et al.'s work, it is very illuminating to point out some select reports given from individual FB users. One reads:

> "[Over time my use of Facebook has] probably increased. The features were -- when I first started, it was all about, you know, friending people, finding out who was on Facebook because it was kind of a big deal, you know? But now, I

[49] One star (*) indicates $p<0.05$, two stars (**) indicate $p<0.01$ and three stars (***) indicate $p<0.001$

kind of use it to see what's going on with my friends rather than just friending people. I don't look to expand my friend base. I know I'm not going out there searching people I'm not friends with. I use it now for photos a lot and that wasn't a part of Facebook when I first joined."

Another, however, explains:

"I don't use it as much, and especially -- I know, when I first joined, it was like a year old, or something, and the simplicity of it was nice, but now it is getting way too involved and complex, and it is just hard for me to move around [and] do stuff. So, I don't do a whole lot on it anymore."

For others, the increased amount of members made the SN overwhelming:

"When there were less people, when I first joined... I would actually read the profiles, because it wouldn't take so long and to keep up on what everyone was doing. But now that, you know, pretty much everyone adds you, it's just, it's gotten a little bit overwhelming."

Finally, there were reports of users giving up, after realizing the superficiality that often characterizes online relationships:

"I guess when I first started; I thought it was like cool to have more friends at MSU. Like, oh, yes, I have so and so amount of friends at MSU. And now, it's just like I don't care enough, because now I've been here like three years or whatever. And, I just want to be friends with the people that I'm actually friends with."

4.1.2 The Attitude of the Users is changing

As time went by, the users' attitude toward Facebook has been changing steadily[50]. The list of elements that triggered this constant change of attitude towards the site includes minor, as well as major developments, such as the radical growth of member population, and the innovative features offered by the platform in the years that came. Between 2006 and 2007 several changes occurred in how the respondents in each sample viewed Facebook. According to Lampe et al. all Measures of positive attitude towards Facebook increased significantly.

Table 4.2 depicts the means and standard deviations on several Measures regarding users' attitude towards the Social Network of Facebook. Randomly sampled participants in the survey were asked to report the degree to which they agreed with a series of statements; their responses recorded with the help of the Likert scale ordinal ratings, where higher numbers indicated more agreement. Any significant difference between the years is determined with the help of independent samples ANOVA tests; a higher number in the column "F" of Table 4.3 denotes a larger difference. In addition, a Tukey's post-hoc test was conducted to determine whether there were statistically significant differences between individual years, allowing us to compare 2006 data against both 2007 and 2008.

In the last two lines of the Table, self-reports of mean time spent per day on Facebook and number of Facebook friends are included.

[50] Lampe, C., Ellison, B.N. and Steinfield, C. (2008) "Changes in Use and Perception of Facebook." Proceedings of the '08 ACM Conference on Computer Supported Cooperative Work (CSCW), ACM, NY

Year of survey		2006		2007		2008	
I use Facebook to	F	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Find people to date	2.34	1.48	0.870	1.53	0.865	1.62	0.952
Meet new people	2.18	1.97	1.027	2.13	1.154	2.14	1.116
Check out someone I met socially	6.92	3.99	1.053	4.17[1]	0.851	3.94[1]	0.887
Learn more about other people in my classes	4.21	3.26	1.204	3.49[1]	1.114	3.32	1.068
Learn more about other people living near me	0.63	2.86	1.218	2.97	1.248	2.95	1.149
To keep in touch with old friends	3.86	4.42	0.861	4.58[1]	0.684	4.50	0.671
Number of Facebook Friends	37.51	201	114	308[1]	215	333[2]	227
Minutes per day on Facebook	22.77	28	36	83[1]	152	82[2]	117

Table 4.2: Responses to the question "I use Facebook to…" rated on a Likert scale for likeliness (higher values correspond to higher likelihood to engage to the activity)[51]

I should be noticed that Facebook was a SN that focused on providing social information about peers (and others in ones extended social circle). Between 2007 and 2008, changes were not as significant as in the period before. Nevertheless, Facebook appeared to have become well integrated into its members' daily routines since 2006 and 2007. Still, once participants were integrated into the site, these gains were not replicated the following year. The News Feed, which was launched in the fall of 2006, was probably a major factor explaining these changes, as it encouraged the users to have short sessions with the site, through which they could quickly review the recent activities of their friends and peers.

Table 4.3 depicts the means and standard deviations on several Measures regarding users' attitude towards the Social Network of Facebook. Randomly sampled participants in the survey were asked to report the degree to which they agreed with a series of statements; their responses recorded with the help of the Likert scale ordinal ratings, where higher numbers indicated more agreement. Any significant difference between the years is determined with the help of independent samples ANOVA tests; a higher number in the column "F" of Table 4.3 denotes a larger difference between the years. Finally, a Tukey's post-hoc test was conducted to determine statistically significant differences between the individual years.

Year of survey		2006		2007		2008	
	F	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Facebook is part of my everyday activity	36.12	3.12	1.26	3.75[1]	1.11	3.85[2]	1.12
Facebook has become part of my daily routine	35.82	2.96	1.32	3.70[1]	1.16	3.66[2]	1.19
I am proud to tell people I am on Facebook	2.90	3.24	0.89	3.40[1]	0.87	3.34	0.85
Facebook is just a fad	12.15	3.14	1.03	2.96	1.09	2.75[1,2]	1.00
I would be sorry if Facebook shut down	5.21	3.45	1.14	3.69[1]	1.19	3.72[2]	1.34
I use Facebook to get useful information	78.51	2.55	1.10	3.39[1]	1.02	3.54[2]	1.00
I use Facebook to find out about things going on at MSU	56.59	2.59	1.08	3.34[1]	1.18	3.51[2]	1.10
My Facebook use has caused me problems	22.51	1.67	0.89	2.14[1]	1.10	2.20[2]	1.12
I spend time on Facebook when I should be doing other things	9.44	3.16	1.15	3.52[1]	1.23	3.54[2]	1.18

Table 4.3: Ratings of Attitudes towards Facebook

[51] A "1" superscript in the table values indicates a significant difference ($p<0.05$) with the corresponding value of the year before. A "2" superscript indicates a significant difference between the values of 2006 and 2008.

The items “Facebook is part of my everyday activity” and “Facebook has become part of my daily routine” probe how regularly respondents view the site, and in all three survey periods they largely agreed with the statement regarding “everyday activity.” Respondents also indicated high agreement with two Measures asking about the “usefulness” of Facebook, operationalized by the questions “I use Facebook to get useful information” and “I use Facebook to find out about things going on at MSU.” While agreement with the statement that “My Facebook use has caused me problems” has grown over the different samples, all responses remain low, with the average response being to “somewhat disagree” with the statement. When asked about whether anything negative had happened to them as a result of their Facebook use, interview respondents described fights with romantic partners, spending too much time on the site, or becoming preoccupied with one’s profile and online self-presentation. Many had heard stories from friends, professors, or others about Facebook users losing jobs or opportunities due to questionable content on their profile. However, these stories did not amount to personal experience and the general atmosphere towards the site remained positive.

4.2 The Current Legal Situation in Europe

The next pages will be devoted to a presentation of the current situation of the field of privacy and data protection from a legal point of view. For this purpose will shall rely especially on the work of Quirchmayr et al. on data protection and privacy laws in light of emerging technologies. The issues discussed in the next paragraphs constitute a significant motivational factor for participating in the Security VS Privacy sociopolitical debate. However, this section of the Thesis can also be regarded independently from the rest.

4.2.1 Privacy Protection in Contemporary Society

Privacy protection in today’s world is difficult to argue when an individual’s right to privacy is required to be respected while, at the same time, openly abused by criminals and used to harm the public. When terrorism and organized crime force societies all over the world to cut back on all sorts of human rights, personal privacy seems to be the first victim in the cause of security.

This observation has been increasingly obvious in the new technology introduced in the last decade. From the most sophisticated equipment to the most common everyday devices, efficiency and comfort have come with a frightening price: the amounts of personal data collected and mined are increasing exponentially[52]. While the nightmarish Orwellian scenarios described in several books ring an alarm with the public, Ambient Intelligence has practically already invaded little corners of our lives in ways we hardly notices. Nice gadgets added to our mobile phones and PDA’s have been widely accepted and the users have quickly embraced the fancy applications coming with their new devices.

Location based services, the first widely spreading form of context aware services, are highly helpful, while at the same time revealing a lot of information about the mostly unaware user. Questions such as “Which information about a certain location is the user interested in?” and “What are the typical movements of users at a certain time of day?” will be easy to answer once the user is forced to be online permanently. As long as the paradigm remains that the user is logging on to a system via a device and not a system logging on to a device operated by the user, the control is at least with the user. Ubiquitous or pervasive computing is beginning to change this in a drastic way. Questions such as “Which level of control should the user have in the future?” and “Which level of privacy should the user be granted?” are already starting to dominate the privacy protection discussions.

As comfortable as it is to walk into an area covered by a system and automatically be recognized and provided with the full spectrum of services, this comfort comes at a very

[52] Quirchmayr, G. and Wills, C.C. (2007) “Data Protection and Privacy Laws in the Light of RFID and Emerging Technologies.” Proceedings of TrustBus, 155-164

high price. These services only work if the system has sufficient information about a user, meaning that the more a system "knows" about a person, the better it can tailor the service offered to the user. The extent to which this approach can go wrong can ubiquitously be felt by the pervasiveness of spam.
It is therefore in the interest of consumers, users and technology providers to start an open discussion in order to create an environment of trust in which technology will again be viewed as helping friend, instead of the "tool for surveillance" and "enemy of the people" image it has acquired over the past view years. Data and privacy protection legislation can play a decisive role in achieving this goal.

4.2.2 New Technologies and Legal Privacy Debates

Some of the new systems and technologies that are beginning to be used in either defense and law enforcement environments or in commercial contexts, are bound to cause controversy from a privacy perspective. As necessary as the introduction of this technology may be, the way in which it has been handled has in several cases, provoked an outcry from advocates for privacy. The technologies under scrutiny from privacy groups today, are primarily databases and information systems operated by law enforcement and other governmental agencies for the prevention and investigation of serious crime, location-based and other context aware services aimed at users of mobile equipment, customer cards and the RFID. It was initially not so much the technology itself that provoked the adverse reactions, but the envisaged and in some cases already practiced uncritical and uncontrolled use of person-related information, collected through the application of this technology that has already given some of the technology a very bad name. An envisaged data exchange that has initially been aimed at increasing the security of airline passengers has for example sparked a completely unnecessary conflict between the US and the European Union, finally resulting in the European Parliament taking the EU Commission to court over an alleged breach of data protection legislation[53]. Privacy advocates all over the European Union and safety fears in the US have contributed their share in escalating the situation and damaging the relationship. RFID has led to similarly strong emotional reactions which the discussions accompanying the planned use of the technology by companies in California[54] and by the clothing industry in Europe[55] [boycottbenetton 2003] are frequently being quoted as reference points for the growing fear of consumers.
The recent European proposal to store basic data about phone calls for a length of up to three years, in case this information should be needed for the investigation and prosecution of serious crime, has immediately resulted in very critical reactions from privacy advocates in Europe. In this context, the ability of telecommunications operators to collect an increasing amount of customer-related information that is generated from location-based services and from payments made via the mobile phones becomes problematic. There is no doubt that in the cases of serious organized crime and terrorism, it would be very beneficial to have all this information, but the question arises who else other than law enforcement officers, might be given access to this customer history once the data has been collected.

4.2.3 Efficiency and Security VS Privacy

Countless previously documented attempts to use new technology to circumvent privacy legislation have raised the level of suspicion among customers and employees. The major problem however, is that of the rather careless use of technology whenever it becomes available. This has again been documented by the analysis of WLAN and Bluetooth

[53] Electronic Privacy Information Center, (2004) "EU-US Airline Passenger Data Disclosure", (http://epic.org/privacy/intl/passenger_data.html, accessed on 3 June 2011)
[54] RFID News (March 1, 2004) "Bowen Seeks Balance in RFID Law", RFID Journal, http://www.rfidjournal.com/article/articleview/812/1/1/, accessed on 3 June 2011)
[55] Boycottbenetton Site (2003), (http://www.boycottbenetton.com/, accessed on 3 June 2011)

connections all over Europe. Safe in theory and equipped with technology that can block out an intruder, the equipment usually comes with a standard configuration that is not aimed at security, but at the ease of use. Unaware users installing WLAN access points with standard configurations, turning on Bluetooth enabled mobile phones without checking the status of the Bluetooth connection, all too often find themselves in a situation where they openly invite access to their devices and the connected networks without even realizing the potential dangers they create. In spite of legal regulations[56] requesting that all necessary and financially justifiable measures be taken to keep person-related data safe, unaware users continue to ignore even the elementary basics of data and privacy protection[57].

Movement tracking, combined with increasingly complete consumer behavior profiles, gives companies the possibility to deliver the right product or service at the right time in the right place. As is well known, the position of mobile equipment, typically a mobile phone, being identified by either GPS or location services implemented through provider base stations, can today be determined quite exactly. Future systems will allow the calculation of a position within some centimeters. The core legal question is to what extent this data can be used by applications. The push towards storing more and more information over longer periods and to have it readily available in case it is needed for business evaluations, or for the future prevention and prosecution of crime, is in direct contradiction to the aim of privacy protection, to have only the minimal amount of data stored and to grant access only for predefined applications. The second problem is that the more data we collect about a person, the more sensitive this data becomes, because the increasing amount of available data allows the construction of an increasingly complete subject profile.

The scale and potential implications of identity theft scandals have reached a frightening dimension, which the DSW scandal amply documents:

> "The numbers and the names associated with approximately 1.4 million credit and debit cards used at 108 of our stores primarily during a 90 day period between mid-November 2004 and mid-February 2005 were stolen from DSW ... In addition, checking account information was stolen for around 96,000 checks used to make purchases at these same stores. This included the bank account numbers located on checks that were provided to DSW (the "Magnetic Ink Character Recognition" or "MICR" numbers) and the drivers' license numbers provided when paying by check."

Especially when cases like these emerge, the appropriateness of data protection measures taken by companies handling such large amounts of sensitive financial data needs to be investigated[58].

The current situation can be attributed to a mixture of poor awareness and negligence, both on the part of the system administrator and on the part of the end user side. PIN codes being written on the back of ATM and credit cards in spite of all warnings, completely unprotected WLAN's, and PIN codes on mobile phones being turned off show that many users are at least as careless as some of the worst companies operating the IT systems. Intruders therefore see "phishing" and similar attacks, the intrusion in unprotected or only weakly protected systems and different forms of identity theft, as an easy way to commit crime. With the possibility of organized crime getting involved as well, commercial IT infrastructures might soon become so vulnerable that they become unusable for business purposes.

---

[56] Austrian §14 Bundesgesetz, Datenschutzgesetz 2000 - DSG 2000 (NR: GP XX RV 1613 AB 2028 S. 179, BR: 5992 AB 6034 S. 657

[57] Quirchmayr, G. and Wills, C.C. (2007) "Data Protection and Privacy Laws in the Light of RFID and Emerging Technologies." Proceedings of TrustBus, 155-164

[58] Thiesse, F. (2007) "RFID, Privacy and the Perception of Risk: A strategic Framework." Journal of Strategic Information Systems

Legal frameworks, as well developed as they might be, will therefore have to be complemented with the necessary technological defenses and an according legal obligation to implement them. This legislation partially already exists on a European scale (see Article 17 of the Directive 95/46/EC[59]).

4.2.4 Security and Privacy Legal Framework in Europe

From a legal point of view, everybody claims to want to enforce privacy. Yet, at the same time, everybody seems to be pulling the carpet, by requesting and collecting tons of all kinds of personal data wherever they can be found.

Nevertheless, existing legal frameworks can in some parts of the world cope very well with challenges to privacy. The European Data Protection Directive, which has been in effect since the 1990s has widely been viewed as one of the landmark agreements in privacy protection. As one of the core underlying assumptions of modern privacy protection is that it covers all forms of automated and non-automated processing of data, the change of technology cannot result in the successful circumvention of privacy protection legislation. Debates such as the one on RFID tags in 2004[60] occurring inside the European Union would therefore see European privacy advocates being able to argue on the basis of an already existing and comparatively comprehensive legislation.

The really alarming problem associated with the new technology is its use in cooperation with companies located outside the European Union. Unless covered by international treaties and agreements, such as the Safe Harbor Agreement[61] between the US and the EU, problems will doubtlessly occur as soon as any person-related data is exported outside the EU. This may, if not properly taken care of, become a serious obstacle to free trade, especially whenever customer-related information is to be stored in information systems located outside the European Union.

The fundamental guidelines set out in the European Data Protection Directive are, where necessary, complemented by other European legislation on specialized areas, such as digital signatures, telecommunications, electronic commerce. Privacy legislation is also backed by European Human Rights legislation, which over the years has been embedded in the constitutions of Member States of the European Union.

4.2.5 Summarizing

Consequently, adequate legislation can give consumers and users of the technology the much-needed safety net, which ultimately makes a new technology trustworthy and therefore acceptable. However, the possibility of abuse by criminals will always be there with every new technology. That is why legislation has to be accompanied by the safeguard triplet: trustworthy safety, security mechanisms and organizational arrangements. Only those three combined can prevent the careless and improper use of the new technology, especially in the field of IT where users and consumers will be prepared to widely accept the new technology. With advanced business strategies being highly dependent on information technology, this combination of safeguards becomes essential, not only for the protection of privacy, but also for our economy to be able to successfully continue to develop. Therefore, the avoidance of a public that will aggressively reject new IT as unsafe and insecure, justifies a substantial investment in the development of adequate legislation and in the technologically sound implementation of the fulfillment of requests made by this legislation. Information technology legislation and associated privacy protection technology come at a considerable cost, but not making

---

[59] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data; Official Journal L 281, 23/11/1995 P. 0031 - 0050

[60] RFID News (March 1, 2004) “Bowen Seeks Balance in RFID Law”, RFID Journal, http://www.rfidjournal.com/article/articleview/812/1/1/, accessed on 3 June 2011)

[61] Safe Harbor Website (2005), (http://www.export.gov/safeharbor/, accessed on 3 June 2011)

this investment might lead to short term savings only to cause a very expensive catastrophe later.

5 Bridging the Gap

This chapter aims to construct a bridge between the existing literature and the attempt to propose the new ideas this study aim to promote. The literature that will be reviewed in these pages is part of the literature actually used during the main part of this study, as it constitutes the main foundations and supports of the basic theories explained in later chapters.

Security and Privacy in Social Networks are a significant part of Web 2.0 and constitute crucial research topics[62] of plenty of different disciplines[63]: sociologists, legal experts, computer scientists, economists etc. In this section we overview some of previous work that is most relevant to collaborative privacy management for SNs. Several studies have been conducted to investigate users' privacy attitudes, and possible risks which users face when poorly protecting their personal data[64] in SN sites. Gross et al.[65] provided an interesting analysis of users' privacy attitudes across SN sites. Interestingly, Ellison et al.[66] have highlighted that on-line friendships can result in a higher level of disclosure due to lack of real world contact. According to Ellison et al. there are benefits in social capital as a result of sharing information in a social network that may limit the desirability of extensive privacy controls on content. Following such considerations, our proposed approach does not simply block users' accessibility to shared data, but it ensures that sharing occurs according to all the stakeholders' privacy interests. The need for solutions addressing the problem of information leakage in this context is also reported in Hobgen et al.[67] where an extensive analysis of the more relevant threats that SN site users currently face is reported.

To cope with security and privacy problems, SN sites are currently extending their access control based mechanisms, to improve in flexibility and limit undesired information disclosure. There is a general consensus that a new access control needs to be developed for SN sites[68]. Gollu et al. was the first to make an attempt along this direction[69], where a social-networking based access control scheme suitable for online sharing was presented. They proposed an approach that considered identities as key pairs, and social relationship on the basis of social attestations. Access control lists are employed to define the access lists of users.

---

[62] Gross, R. and Acquisti, A. (2005) "Information Revelation and Privacy in Online Social Networks." Workshop on Privacy in the Electronic Society, Alexandria, VA, ACM Press

[63] Felt, A. and Evans, D. (2008) "Privacy Protection for Social Networking Platforms." Proceedings of Web 2.0 Security and Privacy 2008 (in conjunction with 2008 IEEE Symposium on Security and Privacy)

[64] Rosenblum, D. (2007) "What Anyone Can Know: The Privacy Risks of Social Networking Sites." IEEE Security and Privacy, 5(3), 40-49

[65] Acquisti, A. and Gross, R. (2006) "Imagined Communities: Awareness, Information, Sharing and Privacy on the Facebook." 6th Workshop on Privacy Enhancing Technologies, 36-58, Springer, Cambridge, UK

[66] Ellison, N., Steinfield, C. and Lampe, C. (2007) "The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites." Journal of Computer Mediated Communication, 12 (4), Article 1

[67] Hobgen, G. (2007) "Security Issues and Recommendations for Online Social Networks." ENISA Position Paper N.1

[68] Carminati, B., Ferrari, E. and Perego, A. (2006) "Rule-Based Access Control for Social Networks." Proceedings of the OTM Workshops, 1734–1744

[69] Gollu, K.K., Saroiu, S. and Wolman, A. (2007) "A Social Networking-Based Access Control Scheme for Personal Content." Proceedings of the 21st ACM Symposium on Operating Systems Principles

Carminati et al.[70] have proposed a rule-based access control mechanism for SN sites that is based on enforcement of complex policies expressed as constraints on the type, depth, and trust level of existing relationships. Furthermore, Carminati et al. proposed using certificates for granting relationships' authenticity, and the client-side enforcement of access control according to a rule-based approach. Squicciarini et al.[71] employ privacy policies using a simplified version of the access rules used by Carminati at al. More recently, Carminati at al.[72] have extended their previously proposed model to make access control decisions using a completely decentralized and collaborative approach. However, the method of collaborative privacy management followed in this work does not relate to the privacy of users' relationships. Rather, we follow the approach of Squicciarini and focus on collaborative approaches for privacy protection of users' shared content.

Recently, Gates[73] has described relationship based access control as one of the new security paradigms that addresses the requirements of the Web 2.0. Hart et al.[74] proposed a content-based access control model, which makes use of relationship information available in SN sites for denoting authorized subjects. However, those works do not address collaborative privacy issues.

Another interesting work related to ours is HomeViews[75], an integrated system for content sharing supporting a lightweight access control mechanism. HomeViews facilitates ad hoc, peer-to-peer sharing of data between unmanaged home computers. Sharing and protection are accomplished without centralized management or coordination of any kind. This contribution, although very interesting, is designed around a very different environment, and it con- siders sharing of content without taking into account multi users privacy implications.

Mannan et al.[76] proposed an interesting approach for privacy-enabled web content sharing. Mannan et al. leveraged the existing "circle of trust" in popular Instant Messaging (IM) networks, to propose a scheme called IM-based Privacy-Enhanced Content Sharing (IMPECS) for personal web content sharing. This approach is consistent with our ideas of sharing of privacy controls, and presents an interesting implementation design. On the other hand, IMPECS is a single-user centered solution: that is, only one user is involved in the decision of whether to share his/her content within his/her trust circle.

Finally, with regards to game theoretic approaches, related work has been done by Varian[77], who conducted an analysis of system reliability within a public goods game-theoretical framework. Varian focused on two-player games with heterogeneous effort costs and benefits from reliability. He also added an inquiry into the role of taxes and fines, and differences between simultaneous and sequential moves. Grossklags et al. in

---

[70] Carminati, B., Ferrari, E. and Perego, A. (2006) "Rule-Based Access Control for Social Networks." Proceedings of the OTM Workshops, 1734–1744

[71] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

[72] Carminati, B. and Ferrari, E. (2008) "Privacy-Aware Collaborative Access Control in Web-Based Social Networks." Proceedings of DBSec, 81–96

[73] Gates, C. (2007) "Access Control Requirements for Web 2.0 Security and Privacy." In IEEE Web 2.0 Privacy and Security Workshop, Oakland, CA

[74] Hart, M., Johnson, R. and Stent, A. (2007) "More Content – Less Control: Access Control in the Web 2.0." IEEE Web 2.0 Privacy and Security Workshop

[75] Geambasu, R., Balazinska, M. Gribble, S.D. and Levy, H.M. (2007) "HomeViews: Peer-to-Peer Middleware for Personal Data Sharing Applications." Proceedings of SIGMOD Conference, 235–246

[76] Mannan, M. and Van Oorschot, P.C. (2008) "Privacy-Enhanced Sharing of Personal Content on the Web." Proceedings of WWW, 487–496

[77] Varian, H.R. (2004) "System Reliability and Free Riding." Economics of Information Security, Kluwer Academic Publishers, 1–15

generalized[78] and build from public goods literature to model security interactions through three well-known games, introducing a novel game (weakest target, with or without mitigation) for more sophisticated scenarios. Similarly, Squicciarini et al. model the collective privacy problem as a new game, using the results from game security economics. The adoption of their carefully selected technique ensures the design of a N-player game, in which truthfulness and correctness are the winning strategies.

The Clarke-Tax algorithm[79] has been recognized as an important social decision protocol. The approach has been applied to address problems of different nature[80]. The Clarke-Groves mechanism has already been introduced into artificial intelligence, using it to explore multi-agent planning, where at each step, instead of negotiating over the next joint action, each agent votes for the next preferred action in the group plan and individual preferences are aggregating using a voting procedure. Recently, Wang et al.[81] proposed an interesting secure version of the Clarke-Tax voting protocol. Following the security requirements identified by Wang et al., Suicciarini et al. actually implement a system, which guarantees full protection of users' privacy and universal verifiability. However, Wang's solution heavily relies on cryptographic primitives, and encryption techniques, implying a level of sophistication of users, which may not be appropriate in Web 2.0 settings.

The Clarke-Tax mechanism is appealing for several reasons. First, it is well suited to our domain, in that it proposes a simple voting scheme, where users express their opinions about a common good (i.e., the shared data item). Second, the Clarke-Tax has proven to have important desirable properties: it is not manipulable by individuals, it promotes truthfulness among users, and finally it is simple. Under the Clarke-Tax, users are required to indicate their privacy preference, along with their perceived importance of the expressed preference. Simplicity is a fundamental requirement in the design of solutions for this type of problems, where users most likely have limited knowledge on how to protect their privacy through more sophisticated approaches.

No other model combines the merits of the Clarke-Tax together with an inference design that exploits folksonomies and automating collective decisions, thus freeing the users from the burden of manually selecting privacy preferences for each picture. We analyze this model in chapter 8.

[78] Grossklags, J., Christin, N. and Chuang, J. "Secure or Insure? A Game-Theoretic Analysis of Information Security Games." Proceedings of the WWW, 209–218

[79] Clarke, E.H. (1972) "Multipart Pricing of Public Goods: An Example," S. Mushkin (ed.), 125-130, Public Prices for Public Products, The Urban Institute, Washington

[80] Ephrati, E. and Rosenschein, J.S. (1991) "Voting and Multi-Agent Consensus." Proceedings of the 9th National conference on Artificial Intelligence, Vol. 1, AAAI Press

[81] Wang, C. and Fung Leung, H. (2004) "A Secure and Private Clarke-Tax Voting Protocol Without Trusted Authorities." Proceedings of 6th International Conference on Electronic Commerce, 556–565, ACM, New York, NY, USA

6 Identifying Core Requirements

In accordance to the literature discussion and the requirement collection that took place in the previous pages, we are now going to start a full-scale presentation of our proposed model for effective Privacy Management in online Social Networking Environments. Our model is based on an aspiring combination of collaborative Privacy Management, as it has already been proposed through instances and game-theoretical approaches using the Clarke-Tax algorithms, with methods of reliable measurement of Appropriation, an idea developed from Adaptive Structuration Theory (for short, AST). These pages are a presentation of the requirements collected for our model, so that the process of choice will be more obvious.

Online PM of our time has grown ever more complex as the disclosure implications of information has to be considered across both time and space[82]. Online data Privacy Management in today's Social Networking sites in currently done via controlling the information access in each data category – and not via controlling each piece of information added to each profile. For example, one can set up restriction for all photos, all videos, all blog entries and so forth. This allows control over access to a category, but not over just one member of it. Since limiting everything in a category is usually overkill, members do not bother. Furthermore, one may forget that they restricted access to their photos, but did nothing about their videos. This may end up in the exposal of an embarrassing video, when the user is thinking that is was visible only to his or her "friends" – and not find out about it, before it is too late and the information has already been accessed by people one did not want to.

6.1 Requirements Identification and Collection

We believe that Collective Privacy Management is an important contribution in the realm of Web 2.0. Nevertheless, even though collaboration and sharing represent the main building blocks of Web 2.0, contemporary social networking sites support privacy decisions mainly as individual processes. Designing a suitable approach to address this problem raises a number of important issues. First, co-ownership in SN platforms should be supported. Second, the approach should promote fairness among users and be lightweight. Moreover, the approach should be practical and promote co-ownership, since users knowingly do not enjoy spending time in protecting their privacy[83].

In case of single-user ownership, the enforcing of a Privacy Policy for a piece of data s is pretty straightforward[84]. The user is responsible for setting a Privacy Policy according to his or her privacy preferences; the Privacy Policy then dictates who has access to data s according to distance and type of relationships between potential viewers and the owner. On the other hand, a shared data object s has multiple owners where each owner might have different and possibly contrasting privacy preferences. Designing an approach that combines different owners' privacy preferences into a unique Privacy Policy is not an easy challenge. It is rather obvious that, in the process of locating the "golden ratio" among overall Privacy Policies, individual preferences will have to be set aside. Nevertheless, when multiple owners share multiple data under multiple Privacy Policies, decisions made for past interactions will be put into the equation, making the model adaptive and flexible.

---

[82] Boyd, D. (2007) "Social Network Sites: Public, Private, or What?" http://kt.flexiblelearning.net.au/tkt2007/edition-13/social-network-sites-public-private-or-what/ (accessed on 28 Mar 2011)

[83] Spiekermann, S., Grossklags, J. and Berendt, B. (2001) "E-privacy in 2nd Generation E-commerce: Privacy Preferences versus Actual Behavior." Proceedings of the 3rd ACM Conference on Electronic Commerce, 38–47, NY

[84] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

In current SN sites, users have little flexibility when specifying privacy policies (also referred to as access rules or privacy settings), and can choose among a limited set of predefined options, such as: friends, friends of friends etc. Additionally, access rights in a SN are limited to few basic privileges, such as read, write and play for media content. It has been proven necessary that each user should be able to enforce locally specified privacy policies over their data posted in his profile. Such privacy policies should be simple statements specifying for each locally owned data item who has access to it, and, in certain cases, which kind of operations can be performed on the data.
Several intuitive approaches prove to be unsuitable, due to the specific constraints that Social Networks present, in contrast to other domains. The option of "selective disclosure of data" is generally not desirable – often not even possible. If, for example, the data item in question were a picture, cropping or blurring would result in a ruined picture, with a decreased to users and owners alike. Similarly, if a document were co-authored, separating contributions to co-authors would likely make it illegible. Note that, cryptographic techniques may theoretically solve the problem of selective data disclosure to entitled viewers. However, these approaches will not compose a unique privacy policy that incorporates the preferences to the different co-owners, and will result in a very unpractical approach, with a very large number of encryption keys for users to manage.
A database-like approach, where different owners could enforce their local "views" would not work either, as this approach may result in privacy violations. Example 6.1 can make this clearer:
Example 6.1: User A may accept only friends to view a specific party picture; yet User B may not care and leave the picture public to any member of the network. Clearly, a User C – who is not a friend of Alice – could easily log into the network and access the picture through User B's profile, thus violating User A's privacy wishes, and despite the fact that the picture itself is not available for User B from User A's profile.

6.2 Requirement Analysis

Based on the considerations above, we identified the following *core requirements* for privacy management:

- *Content Integrity*: The users' data should under no circumstances be altered, or selectively disclosed. In other words, we cannot assume to blur a picture or crop it to remove certain subjects appearing in it. Nor can we alter a document text or data to satisfy conflicting individuals' preferences.
- *Semi-automated*: The access policy construction process should not solely rely on user's manual input for each data, but should leverage users' past decisions and draw from the existing context.
- *Adaptive*: When a new co-owner is added for a data item s, their input should be taken into account, even if the access policy for s has been already set up.
- *Group-Preference*: The algorithm must leverage the individuals' information to develop a collective policy.

In addition to those clearly set requirements, we need to add one more. It is clear that a model designed to serve a system involving such intense human interaction should be integrated with a particularly human-directed concept of measuring results and providing feedback.
At this point, it should be clear, that one should not confuse this requirement with the requirement for "Adaptivity" listed above. We are now talking about the *measuring* of the whole system, in order to provide feedback – no matter how the model may be designed in the end, no matter what its features and other requirements may be. For measuring how users adopt and adapt to a system and, in our case, to a Privacy Policy model, we employed the concept of Adaptive Structuration Theory and the experimental work of Dwyer et al.[85] more of which is analyzed on the next chapter.

[85] Dwyer, C., Hiltz, S.R., Poole, M.S., Gußner, J., Hennig, F., Osswald, S.,

Schließlberger, S. and Warth, B. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." Proceedings of the 43rd Hawaii International Conference on System Sciences, Volume: 1, Publisher: IEEE Computer Society

7 Structuration Theory and its contribution to this Thesis

The Theory of Structuration was proposed in 1984 by Anthony Giddens in an attempt to model dualities, like the agency and the structure (or culture), in social systems and bring a balance between them. The idea behind the model was that it did not focus on the participating actors or the societal totality but "on social practices ordered across space and time"[86]. Those who supported this model were more adept to the idea of the equilibrium[87] and visualized balancing such dichotomies accordingly.
In order to better understand how this theory can be a part of a model that helps coping with Privacy Management issues on the Web, we need to take a more detailed look at its core.

7.1 The Structure and the Agency

It has been a general aim of contemporary sociology to reconcile the two major concepts that have been defined under the terms Structure and Agency. Giddens developed his "Structuration Theory" in his work "The Constitution of Society (1984)". Giddens defines Structures as "Rules" and "Resources", or sometimes sets of transformation relations, neatly compartmentalized as properties of our societal systems. "Rules" formulate how people act on their everyday social environments, whereas "Resources" refer to anything provided via manpower – as opposed to whatever is free from nature (like the air or the wind).
It is ominous in Giddens' work that he attempts to escape from the duality of Structure vs. Agency through arguing the dualities of "social structure" – where social structure is both the medium and the outcome of social action.[88]. Like any other sociological approach, the theory recognizes actors (or agents), who possess both discursive, as well as practical knowledge. What should however prove of interest to the theory's interdisciplinary use would be its deep acknowledgement of the structural fact that "the habitual becomes institutionalized".
The theory lays down more definitions. According to Giddens, structure, together with modality and interaction, constitute social systems in their entirety and full complexity. As already mentioned above, structure consists of rules and resources – constraining and enhancing agents, respectively. The modality of a social system is the means by which its structures are transformed into action. Finally, under the term interaction, any activity between an agent and the social system is to be understood.
This activity is described under the term Agency. Agency is, according to Giddens, human action. Being human means being an agent, he says – although not all agents are necessarily human beings. Agents act based on their knowledge, discursive or practical, and interact with the structures of the social system. Giddens defines this trust that agents show in social structures as "ontological security" – based on the level of this "ontological security", agents' everyday actions possess a certain degree of predictability, thus ensuring social stability. It is perhaps not too soon to mention that, one should not fail to notice the parallels that are automatically drawn between the space of offline and online social structures. In addition to how remarkably useful they have proven for general sociological studies, these theories are quite as applicable in the areas of online social networking sites.
Analyzing his theory further, Giddens describes any form of interaction between an agent and a structure as "structuration". According to the "theory of structuration", all human action (agency) is performed within the context of a pre-existing social system (structure),

[86] Giddens, A. (1986) "Constitution of Society: Outline of the Theory of Structuration." University of California Press; Reprint edition (Jan 1, 1986) ISBN 0-520-05728-7, 2, 281-348
[87] The concept of the equilibrium should not pass unnoticed at this point, as it will reappear more than once in our study, despite the fact that it is not among our keywords.
[88] Jary, J. and Jary, D. (2005) Collins Dictionary of Sociology, 664

defined by a set of norms or laws (rules), which is what differentiate it from other social structures. To which extend, however, human action is externally restricted to predefined rules and norms or enhanced (resources) with the power to affect them through reflexive feedback mechanisms, is obviously debatable. At this point, one should also mention the concept of "reflexivity", an idea that refers to the ability of an agent to change his position inside a social structure. It seems of great importance, throughout Giddens work, how modern "post-traditional" society thrives towards "greater social reflexivity"[89]. Thus, social knowledge, as knowledge of each agent, is considered a crucial factor of his own power to create rules within the structure.

7.2 Adaptive Structuration Theory

The theory of Adaptive Structuration was developed in an attempt to take advantage of those presumable parallels that can be observed between social structures and information technology (for short, IT) constructs. DeSanctis and Poole[90] proposed the main ideas in the 1990s to help explain GSS[91] patterns of use, but soon Adaptive Structuration Theory (for short, AST) found fertile ground in the area of social software analysis[92].

The parallels between the two worlds are drawn especially between the modalities (the means by which a structure is transformed into action), and in the sense that structuration is done under pretty much the same concept. The situations in our offline social structures do not differ very much from those in our online social realities, and since the agents are more or less the same, the gaps are not difficult to fill.

Thus, based on Giddens' Structuration Theory, AST aims (again) in identifying the social "structures" as "rules" and "resources" (though this time provided by technology) that constrain and enhance human activity, respectively.

7.2.1 The Concept of Appropriation

Within the framework of Adaptive Structuration Theory, one finds the concept of Appropriation. It is this very concept that has proven to be both relevant and helpful in the field of Privacy Management.

In the work of DeSanctis and Poole, Appropriation is defined as the ongoing processes and methods by which people adopt and adapt to technology. Through these processes and methods, AST manages to analyze the means used for the easier and more effective adaptation and adaptation of technology by the users. Since AST proposes a non-deterministic view of technology use that provides room for social, cognitive, and technical factors[93], it serves as a good foundation for building Measures of online privacy management.

Through the careful study of appropriations, one may draw results on the intricate workings of any given resource, or in this case, any technical detail – how it operates, how it gives results. At the same time, DeSanctis et al. argue that appropriations are not to be mandated by technological designs. Instead, they are to be formed by how human action (or agency) transforms technological structure to its utilitarian purposes, resulting in varied adoption practices.

From a more sociological point of view, appropriation can be seen as the process through which agents integrate the technology into their everyday actions and tasks. This interaction is comprehended as a very complex sociotechnical procedure, with very

---

[89] Levi, H. (2005) "Reflexivity in Sociology", New Dictionary of the History of Ideas

[90] DeSanctis, G. and Poole, M.S. (1994) "Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory." Organization Science 5(2): 121-147

[91] Group (Decision) Support Systems

[92] Dwyer, C. (2008) "Appropriation of Privacy Management Within Social Networking Sites." Information Systems Department, Ph.D. Dissertation, New Jersey Institute of Technology, NJ

[93] Markus, M.L. and Silver, M.S. (2008) "A Foundation for the Study of IT Effects: A New Look at DeSanctis and Poole's Concepts of Structural Features and Spirit." Journal of the Association for Information Systems, 9 (10)

positive cognitive characteristics. As mentioned before, none of the technological aspects are considered to be deterministic – however, the part they play on how users interact with the technology should not be regarded as insignificant. In the end, the concept of appropriation implies adaptations in both directions – technological novelties causing changes in human behavior, so much as common practices leading to new technologies[94].
A particularly poignant example to the bidirectional transformative nature of appropriation can be found in cell phone technology. When first introduced to the market, cell phones were expensive devices, almost exclusively purposed for wireless voice transmission. However, as the devices became cheaper and gained popularity among the general public, strategies emerged to reduce costs of usage. A particularly popular strategy was the exchange of messages via intentionally missed calls[95], better known as beeping or buzzing, an idea based entirely on the concept of free caller ID. An equally popular strategy was that of text messaging, widely known as SMS. Text messaging was generally offered at much lower prices than voice transfer in most places, and this quickly changed the main usage of the mobile phone, from a voice transfer device to a text messaging one. This did not let technology unaffected, as the market was soon presented with mobile devices aimed toward the SMS-users, often even equipped with full QWERTY capabilities[96]. On the other hand and from a sociological point of view, one could not miss how deeply the technological advances affected social life and usage of resources – new patterns and consumption demands were introduced and yesterdays novelties were todays everyday needs.

7.3 Appropriation of Privacy Management in Social Networks

According to AST's definition, appropriation constitutes of the ongoing processes and methods by which people adopt and adapt to technology – or, as we have understood by now, by which people adapt technology to their practices. It is now time to become acquainted with some more of AST's terminology.
AST defines the general purpose and value, towards which a new technology has been designed and introduced, as its "spirit". When agents use the technology in compliance with its spirit, this is filed under faithful appropriation – whereas, when the usage is not compliant with the technology's spirit, the appropriation is deemed unfaithful[97]. AST points out that, the more faithful the appropriations[98], the more promising the outcomes.
A further notion proposed in AST is defined under the term of "appropriation move". An appropriation move refers to nothing else than the way in which agents can appropriate technology. What is important here is that, in their extensive work, DeSanctis and Poole have thoroughly classified and categorized the possible different types of the appropriation moves, and further organized them in four main categories, nine types and 31 sub-types.
The four main categories are:

1. Direct Use
2. Relating one Technical Feature to another

[94] DeSanctis, G. and Poole, M.S. (1991) "Understanding the Differences in Collaborative System Use through Appropriation Analysis." 24th Annual Hawaii International Conference on System Sciences, Kauai, HI
[95] Donner, J. (2008) "The Rules of Beeping: Exchanging Messages via Intentional 'Missed Calls' on Mobile Phones." Journal of Computer-Mediated Communication, 13 (1)
[96] Mendonça, D., Jefferson, T. and Harrald, J. (2007) "Collaboration Adhocracies and Mix-and-Match Technologies in Emergency Management." Communications of the ACM, 50 (3), 45-49
[97] DeSanctis, G. and Poole, M.S. (1994) "Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory." Organization Science 5(2): 121-147
[98] Dennis, A., Wixom, B. and Vandenberg, R. (2001) "Understanding Fit and Appropriation Effects in Group Support Systems via Meta-Analysis." MIS Quarterly, 25 (2), 167-193

3. Constraining a Technical Feature
4. Expressing Judgment about a Technical Feature

As expected, these categories are divided accordingly into types and sub-types, forming a tree-like construct that aims to plot out how people directly or indirectly, faithfully or unfaithfully, and in the end under what purpose and to what end the resources provided to them through technology. Analyzing the complete structure of this huge tree-graph will not prove helpful to the goal of our study at hand – still, anyone interested may find it in the works of the authors.

7.3.1 Deriving Scales from Appropriation Moves

The idea to use Appropriation as the cradle for Privacy Management measurement tools meant that the Appropriation Moves Tree would have to be studied and analyzed, so that eventually, each node would be projected to its corresponding concept in the Privacy Management context. Thankfully, for our study, Catherine Dwyer et al. had already covered this important academic work thoroughly, accompanying it with plenty of field research[99]. The next pages are dedicated to those of this team's results that play the most important part on our argumentation.

The team of Dwyer et al. investigated as to whether categories, types or sub-types could be converted into scales and measurement tools for Privacy Management. They derived that, as it often happens in such cases, there were those who were readily adaptable (such as the explicit use appropriation move), and those who were not (such as the composition appropriation move). However, following certain patterns, semantic differential scales were built and then tested.

The scales[100] are described here:

A. The Use Appropriation Move:
   - Category: Direct Use
   - Type: Direct Appropriation
   - Sub-type: Explicit

   The Use appropriation move measures the extent of actual privacy settings usage; as reported per user. Simply put, it depicts how much the members report that they are making use of the privacy settings. Explicit use of the privacy setting is the definition of faithful appropriation. A Measure for a Use appropriation move could be: "I changed my personal privacy settings on [SN-in-Q]."

B. The Familiar Appropriation Move:
   - Category: Direct Use
   - Type: Direct Appropriation
   - Sub-type: Implicit

   The Familiar appropriation move measures the extent of familiarity, affinity and kinship that the members report to feel towards the Privacy Management tools and settings they need to work with. Familiarity is considered an implicit appropriation move because it involves knowledge, but not necessarily actual use of a PM feature. In this case, faithful appropriation is shown through demonstration of knowledge. A familiarity Measure would be: "I could make my account invisible to everyone but me."

C. The Restricted Scope Appropriation Move:
   - Category: Relating one Technical Feature to another

---

[99] Dwyer, C., Hiltz, S.R., Poole, M.S., Gußner, J., Hennig, F., Osswald, S., Schließlberger, S. and Warth, B. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." Proceedings of the 43rd Hawaii International Conference on System Sciences, Volume: 1, Publisher: IEEE Computer Society

[100] The scales were designed as 7-point semantic differential scales, ranging from "Strongly Agree" to "Strongly Disagree".

- Type: Substitution
- Sub-type: Part

The Restricted Scope appropriation move measures the extent to which members of a SN site restrict their contact within the community to those they already know, rather than exploring new relationships and engaging with unknown members. This appropriation move corresponds to taking additional (personal) privacy management measures along with the PM tools provided by the SN site; screening out the members you allow yourself to get in touch with. AST sorts this out as faithful appropriation, because taking steps to protect your privacy is consistent with the "spirit" of the privacy settings. An example of this Measure would be "I don't want to be spammed by strangers in [SN-in-Q]."

D. The Rejection Appropriation Move:
   - Category: Expressing Judgment about a Technical Feature
   - Type: Negation
   - Subtype: Reject

   The Rejection appropriation move measures the extent to which members explicitly dictate that they do not wish to bother with privacy settings. This appropriation move aims to describe the rejection (or negation) of privacy management tools, in general. A simple example of this Measure would be: "I have no idea what my privacy settings on [SN-in-Q] are."
E. The Faithfulness Scale:
   - The Faithfulness Scale includes questions adapted from the Scale to Measure Faithfulness of Appropriation[101]. The original scale was developed for electronic meeting systems. It has been rewritten to refer to privacy management within SN. It includes these Measures:
     i. I probably use the privacy settings for [SN-in-Q] improperly.
     ii. I failed to use the privacy settings of [SN-in-Q] as it should be used.
     iii. I did not use the privacy settings in [SN-in-Q] in the most appropriate fashion.
     iv. The founders of [SN-in-Q] would disagree with how I use the privacy settings.
     v. The original founders of [SN-in-Q] would view my use of the privacy settings as inappropriate.

7.4 Testing PM Appropriations in Actual SN Sites

It has not been part of the current study to do experimental fieldwork. However, our comparative research revealed the results brought to life by the tests and experiments done by the team of Dwyer et al. to prove Appropriation of great significance to the construction of tools and Measures to deal with Privacy Management issues in the online world. This chapter is devoted to the presentation of the experimental results that prove Appropriation to be the "proper tool for the job".

In order to test the Measures described above, several studies collecting data were conducted at three different SNS. The Measures were first tested at Facebook and MySpace. A detailed report on an earlier version of this study can be found in Dwyer et al.[102] The Measures were subsequently translated and tested at StudiVZ, a SN popular in Austria and Germany.

These were the first studies of this kind, and this is why Measures and scales had to be made from scratch. They were used for many years of research, beginning in 2006, and

[101] Chin, W.W., Gopal, A. and Salisbury, W.D. (1997) "Advancing the Theory of Adaptive Structuration: The Development of a Scale to Measure Faithfulness of Appropriation." Information Systems Research, 8 (4), 342

[102] Dwyer, C., Hiltz, S.R. and Passerini, K. (2007) "Trust and Privacy Concern within Social Networking Sites: A Comparison of Facebook and MySpace." in 13rd Americas Conference on Information Systems, Keystone, Colorado

have since been validated and evolved. In the next pages, tests and results from all three studies will be presented, first separately and then in unison.

7.4.1 Testing Appropriation Measures in Facebook and MySpace

The first two attempts on measuring appropriation were done during the summer of 2007, with two customized surveys. One aimed at Facebook, the other at MySpace – the two biggest SN communities on the west side of the Atlantic. There as an attempt to recruit as many subject from MySpace as from Facebook, but it turned out not to be doable – in the end, ratio was 1:3 in favor of Facebook subjects. Details on the surveys can be found here[103]. The tests were repeated during the following winter, based on feedback.

The results of the surveys were able to prove the appropriation Measures against Reliability and Validity checks. Reliability is the extent to which a Measure yields the same result even if administered at different times in different circumstances. It is an indication of a Measure's stability or consistency. Validity of a Measure is a determination of whether the Measure accurately captures what is claimed, and that it is logical to draw conclusions from the results of those Measures (the work of Rosenthal et al.[104] explains this very clearly). Problems with Validity usually take the form of biases or specific events that call into question whether results are meaningful. By using a random sample for this study, the risk of selection bias from using a convenience sample is reduced.

Following the process of "Multivariate Data Analysis"[105], as described in Hair et al., the Measures were examined using Principal Component Analysis. In this step 5 Factors were identified, explaining 66.98% of the variance.

The next step was to further clarify the Factors by creating what is defined as a "Rotated Solution". Following the Factor rotation method of Hair et al., several rotation methods were tested. The rotation method found to return the best results is the Equamax method.

The intermediate results of the Rotated Factor Analysis were very interesting and the immediate feedback led to a reconstruction of some of the constructs under measurement. For example, the Rejection appropriation move was dropped, because its Measures loaded strongly on other constructs, while, the Familiarity move and the Restricted Scope move now have three Measures each. Tables with plenty of intermediate results can be found in the work of Hair et al. In our case, however, intermediate results are just the interesting path to the useful tool that we seek.

7.4.2 Testing Appropriation Measures in StudiVZ (Austria)

StudiVZ is a quite popular European Social Networking site. Despite the fact that it does not have an English interface, it has grown rather popular among European students – especially German speaking ones. However, contrary to some intrigue between StudiVZ and Facebook on grounds of copy-cat[106], important differences exist between the two SN sites when it comes to PM, especially in terms with familiarity.

A good example for this is the following. StudiVZ sets the privacy management options by default at the minimum protection level. This means that anyone, from peers to non-peers, to complete strangers has access to one's profile. On the other side, Facebook lets

---

[103] Dwyer, C., Hiltz, S.R., Poole, M.S., Gußner, J., Hennig, F., Osswald, S., Schließlberger, S. and Warth, B. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." Proceedings of the 43rd Hawaii International Conference on System Sciences, Volume: 1, Publisher: IEEE Computer Society

[104] Rosenthal, R. and Rosnow, R. (1991) "Essentials of Behavioral Research: Methods and Data Analysis." McGraw Hill, New York, 1991

[105] Hair, J., Black, W., Babin, B., Anderson, R. and Tathman, R. (2006) "Multivariate Data Analysis." Prentice Hall, Upper Saddle River, New Jersey, 2006

[106] Allison, K. (Dec, 2008) "Facebook Sues 'Knock off' Site." Financial Times, 2008, (http://www.ft.com/intl/cms/s/0/8cd4ebbe-551f-11dd-ae9c-000077b07658.html -axzz1MYNasgQe, accessed on 17 May 2011)

by default only friends or group members to see such information (of cource, many groups are very large; but still)[107].
The whole idea behind testing the appropriation Measures in StudiVZ was to conduct the survey in a different cultural setting. The survey had already brought some results for Facebook and MySpace and it would be interesting to compare those results with some from another place, if taken with the same tools. The survey was done during the winter months of 2008, on members of StudiVZ SN community. Questions and scales were translated into the German language and a few extra questions were added so that StudiVZ unique Privacy Management settings would be covered by the questionnaire.
The table below displays the final Measures and Factor loadings of the surveys of Dwyer et al.

Measure	**Factor 1 - Familiarity**
Fam3	I am familiar with my privacy settings on [name of SNS]
Reject1	I don't know what my privacy settings are on [name of SNS]
Fam4	When I need to modify my privacy settings for [name of SNS], I am able to do it
	Factor 2 - Actual Use
Use3	I have adapted the privacy settings to control who can view my profile on [name of SNS]
Use1	In order to control who can contact me using [name of SNS] I have adjusted my privacy settings.
Reject3	I don't use the privacy settings to control who can access my profile
	Factor 3 - Restricted Scope
Scope3	I never accept invitations of people I never met before.
Scope4	When I use [name of SNS] I ignore people whom I never heard of and who try to contact me
Scope2	I don't use [name of SNS] to make contact with people whom I've never heard of.
	Factor 4 – Faithfulness
Faith5	The original founders of [name of SNS] would view my use of the privacy settings as inappropriate
Faith4	The founders of [name of SNS] would disagree with how I use the privacy settings
	Factor 5 – Confidence
Fam2	I am confident that I know how to control who is able to see my profile on [name of SNS]
Fam1	I am comfortable with my ability to adjust my privacy settings
All questions were measured as semantic differential measures, from 1 (Strongly Disagree) to 7 (Strongly Agree)	

Table 7.1: Final Measures and their Factor loadings

As one may have imagined, an important difference between these three sites, that also has a lot to do with the members' familiarity with the Privacy Management tools, is that

[107] MySpace default option is to let a new member's profile content visible to anyone on the Internet.

StudiVZ obligates every new member of the community to go through the privacy settings while registering. This is not the case in either Facebook, or MySpace. Moreover, setting the starting options at a rather unacceptable default actually compels the new user to manage his privacy consciously and not dismissively.
Another significant aspect, on which StudiVZ prides to differentiate itself, is that of optional personalized advertisement towards the users. Contrary to the two U.S. platforms, the Austrian site offers users the choice on whether they would be subjects to data mining in order to receive personalized advertisements and special offers. In fact, he user data collection for targeted advertising has been the subject of a public debate in Austria since 2007, and has received a lot of coverage in the press since; therefore almost everyone knows about it and public awareness is high[108]. Finally, StudiVZ does not fail to warn its members of possible implications their privacy setting might cause – a policy that the U.S.-based sites do not care to share.

7.5 Results and Findings of Appropriation Testing

The long process of analysis and testing of appropriation moves performed by Dwyer et al. lead to the 5 Factors and 13 Measures seen on Table 7.1. Verified against strict validity and reliability checks, these Measures form now a clever tool that can be used to distinguish differences between various SN sites.
When compared to the original taxonomy or appropriation moves based on Adaptive Structuration Theory (AST), 4 Factors persisted from the first analysis on online Privacy Management, and are still visible on Table 7.1:

1. Familiarity
2. Actual Use
3. Restricted Scope
4. Faithfulness

The new Factor that was not initially identified in the original taxonomy is called Confidence. The two Measures that make up this factor were originally created for the Familiarity scale, but have loaded on another factor. Comparing this factor with the AST appropriation moves' taxonomy, this factor is related to the category of "Expressing Judgment", Type "Affirmation", sub-type "Agreement".
The fact that 4 Factors were originally extracted from empirical data, while the 5th Factor can also be interpreted using AST, provides evidence that Appropriation Structuration Theory can serve as an excellent theoretical foundation for the study of Social Networking sites. This is a significant observation, because of the importance of providing the study of Social Networking with a theoretical foundation; this is also obvious in the literature[109].
In addition, the taxonomy of appropriation moves within AST is flexible enough to be applied to other examples of social computing. Since the use of social computing platforms is expanding so much, this helps establish AST as an important tool for explaining the complexity of interaction within online social spaces.
After establishing these 13 reliable Measures, the results were compared across the three SN sites. The results of that comparison are shown in Table 7.2. All Measures indicated statistically significant results with the exception of Faith5 and Fam1[110].

[108] Dwyer, C., Hiltz, S.R., Poole, M.S., Gußner, J., Hennig, F., Osswald, S., Schließlberger, S. and Warth, B. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." Proceedings of the 43rd Hawaii International Conference on System Sciences, Volume: 1, Publisher: IEEE Computer Society
[109] Shneiderman, B. (2007) "Web Science: A Provocative Invitation to Computer Science." Communications of the ACM, New York, NY, 50 (6), 25-27
[110] All results presented here are taken directly from the work of Dwyer et al.

Meas.	FB	MY	SVZ	F	Sig.
Fam3	4.71	5.23	5.92	2.982	.051
Reject1	3.00	2.63	2.02	3.292	.038
Fam4	5.49	5.83	6.10	19.299	.000
Use3	4.26	3.76	4.16	28.765	.000
Use1	4.11	3.70	3.08	23.128	.000
Reject3	3.13	4.03	3.85	7.251	.001
Scope3	4.57	4.72	4.39	14.328	.000
Scope4	4.63	4.42	3.90	5.452	.005
Scope2	5.42	4.71	5.53	7.207	.001
Faith5	2.94	2.43	2.52	1.103	.332
Faith4	2.78	2.51	2.35	6.182	.002
Fam2	4.79	5.17	3.74	9.311	.000
Fam1	5.19	5.82	4.75	1.391	.250

Table 7.2: Comparison of Results: Facebook, MySpace, StudiVZ

The results demonstrate the differences between the 3 sites. It is obvious that the widest range of difference appears in Use3, showing Facebook first, followed by StudiVZ and then MySpace. At the same time, Fam4 shows the highest results for StudiVZ and the lowest for Facebook. The interpretation of these results would be that the Facebook tools in question are used more frequently than the StudiVZ ones; however Facebook members do not feel quite familiarized with them. This should not sound contradictory, as commentaries[111] on the Privacy Managements tools of Facebook have often criticized it as complex and confusing[112]

Another interesting result from these findings is the differences between the sites in the Restricted Scope Measures. StudiVZ members have the highest result for Scope 2 and the lowest result for Scope 4. That puts the members of StudiVZ to be the least likely to use the site to contact people they have never heard of (Scope 2), followed by the users of Facebook and then those of MySpace. At the same time, Facebook members are the most likely to ignore contact from strangers (Scope 4), followed by MySpace users and then StudiVZ users. This shows that the nature of the sites influences the initiating of contact versus the acceptance of new members in ones circle in unique social ways.

7.5.1 Relationships between Familiarity and Use Measures

Although relatively few StudiVZ users have adjusted privacy settings in order to control who may contact them (Use1), its users show a much higher result for Use3, which measures whether they have adjusted their privacy settings in general. This seems to be related to two points, on which StudiVZ users are significantly different from the members of the two U.S. based systems.

1. StudiVZ users declare to be way more familiar with the privacy settings of their system (Fam3).
2. StudiVZ users also generally feel highly confident in their ability to modify those privacy settings (Fam4).

[111] Schneier, B. (Sep 2006) "Facebook and Data Control." Sneier-on-Security Blog, (http://www.schneier.com/blog/archives/2006/09/facebook_and_da.html, accessed on May 6, 2007)

[112] Dwyer, C. (2008) "Appropriation of Privacy Management Within Social Networking Sites." Information Systems Department, Ph.D. Dissertation, New Jersey Institute of Technology, NJ

This higher degree of confidence in their Privacy Management Awareness is confirmed by differences in Reject1, that shows how relatively few StudiVZ users do not know the settings for their profiles (Table 7.2)
However, when looking for answers, one should not overlook the cultural reasons that always affect social masses. The Austrian authors of this paper observe cultural differentiations between the US and the spaces of Austria and Germany. American users express less anxiety about their own privacy management settings compared to their European counterparts. Considering the results for Factor 1 (Measures Fam3, Fam4, and Reject1), versus Factor 5 (Measures Fam2 and Fam1), members of StudiVZ express a higher level of familiarity (Fam3, Fam4) with their privacy settings, but express less comfort (Fam1) with their technical abilities. The combination of higher use but lower comfort level for the StudiVZ subjects is an indirect measure of levels of anxiety about privacy management[113].
This last paragraph forms an attempt to summarize the chapter of Structuration Theory in a manner that it will highlight those points that contribute the most to our study.
Despite the fact that Adaptive Structuration Theory was not a concept born under the umbrella of the IT community, it can clearly provide us with tools that we, as designers and developers, have been researching for since a long time ago. The expansion of the Web has reached levels that we cannot deny, making the adaptation of the technology to the human factor the first and foremost axis of innovation.
AST has set its foundations on sociology and psychology and, through IT, created an applicable model than can provide us with useful means of trustworthy and measurable feedback, not only for the Privacy Management Settings of a SN site, but for every other section of a multi-user net platform.

[113] Dwyer, C., Hiltz, S.R., Poole, M.S., Gußner, J., Hennig, F., Osswald, S., Schließlberger, S. and Warth, B. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." Proceedings of the 43rd Hawaii International Conference on System Sciences, Volume: 1, Publisher: IEEE Computer Society

8 Collective Privacy Management

In this chapter, we are going to approach Privacy Management from a totally different point of view. Our goal now (according to the requirements) is to establish how a collaborative approach of usage in Social Networking communities is going to improve the Privacy Management experience for the users – both from a social, as well as from a technological point of view.

In order to get a clear perspective of the concept of "Collective Privacy Management", as it has been conceived in the game-theoretical approach of Squicciarini et al. [114], one needs to start by analyzing the ideas of co-ownership in Social Networking communities. At this point, and for sheer purposes of enabling communication for the discussion to follow, we propose a simple, abstract representation of a Social Network. The intent is not to represent any real system – it just defines its main components, for the purposes of our study[115].

A Social Network is characterized by these main components:

- U: The set of users. A Social Networking community is represented as a collection of users. Each $i \in U$ is uniquely identified.
- RT: The set of relationship types supported by the SN site. It is possible that two users of a SN are connected among each other by relationships of different types.
- ψ: The function that denotes the assignment of a certain relationship between a couple of users. More specifically, $\psi: Rt \rightarrow U \times U \cup \emptyset$. Given a pair of users i, j we denote their relationship as $i\ Rt\ j$, where Rt is a member of the RT set – that is, a relationship name of one of the supported types[116].
- $Profile_i$: The profile of a user i. We represent it as a tuple $Profile_i = (GRelType_1, \dots, GRelType_k, Set)$, where $GRelType_l$ represents the list of users having a relationship Rt_l such that $i\ Rt_l\ i$ where $Rt_l \in RT$.
- S: The set of data posted on i's profile. We denote the profile components of a user i by means of the "dot" notation. For example, i's friends are represented as $Profile_i.Friends$, while the data set S as $Profile_i.Set$.
- D: The set of data types supported by the SN site. Supported content types can be image files, video and music files, plain documents, hypertext etc.

The connection between two users in a SN is also defined through the help of graph terms. Two users i, j are "directly" connected when there is relationship of the SN tying them together, i.e. $(i\ Rt\ j)$. However, users are also indirectly related, and this is equivalent to a path connecting them through the SN graph.

- *Definition 8.1*: Two users 1,n are related if there exists a path of the form: $[(1\ Rt_1\ 2), (2\ Rt_2\ 3), \dots, (n-1\ Rt_m\ n)]$, where each tuple $(i\ Rt_k\ j)$, denotes an existing Rt_m-type relationship between users i and j. If there is more than one path between users (1 and n), the "distance" between them is defined as the shortest path between their nodes in the graph – the path passing through the minimum number of users.

[114] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

[115] In fact, for reasons of simplicity and adaptability, the representation presented here is a slight modification of the one used by Squicciarini et al.

[116] It should be clear, that the same pair of users could be related by different type of Rt. We assume that all the relationships in general are binary, non-transitive and not-hierarchically structured – for they have no reason to be otherwise. Unary relationships are also possible; however not important to our study.

Example 8.1: Let us say, that we have three users who are part of the same Social Network, and users A and B are friends, whereas users B and C are partners. The *distance* between User A and User C (with respect to the relationships of this SN), is 2 – for the shortest path between the two users is through User B: [(A FriendOf B), (B PartnerOf C)]. Faithful to our Requirements, Collective Privacy Management offers SN users locally specified privacy policies over data posted on their Profile[117]. Here, these policies are combined with user-specified distance-based access conditions, in an attempt to create a generic model. Thus, a user's access to other users' data becomes a function of their *distance* within the SN (as the concept has been defined above). The type of this access (read, write, execute etc.) is not of conceptual significance at this moment. Squicciarini et al. define more helpful tools:

- *Definition 8.2*: A Privacy Policy denoted as $PrP\ (i, n)_{RtSet}$ regards all users within distance n from user i, through Relationships belonging in the Set $RtSet$.

According to the above definition, the extreme cases go as follows. The Privacy Policy of a user i who wishes to leave his profile open to the whole SN would be denoted as $PrP\ (i, \infty)$, while the Policy of a user j who wishes to keep his profile accessible to himself only would be denoted as $PrP\ (j, 0)$. In all other cases, when the path between two users j and i is smaller than the *distance* between them through relationships within the $RtSet$, then the Privacy Policy is satisfied. At this point, one should bear in mind that distance-based access control rules are also employed in the study offline Social Networks, as well as in recent access control models proposed for SN sites[118].

Example 8.2: Let us say that User A from example 8.1 decides to enforce a Privacy Policy of the type: $PrP\ (i, 2)_{FriendOf}$. User B, being a friend of User A and within 1 hop of him, satisfies the policy. John, on the other hand, does not; that is so because, despite the fact that he may be within 2 hops of User A through the relationships $FriendOf$ and $PartnerOf$, the $RtSet$ of the Privacy Policy only contains the relationship $FriendOf$.

8.1 Data Co-ownership in Social Networking Environments

The tools defined in the proceeding paragraphs are going to show their importance, as we continue studying Collaborative Privacy Management. Squicciarini et al. take advantage of its intrinsic characteristics and introduce the concept of Data Co-ownership in Social Networking Environments. Finally, on this concept, they establish their Clarke-Tax based mechanism; a model that shows promising results for collective Privacy Management in SN sites.

Before we proceed to the algorithmic detail of the mechanism, we shall introduce the notion of collaborative data sharing in SN and discuss the possibility of semi-automated detection of co-ownership of data.

The current situation in Social Networks dictates that data uploaded by SN members on their profiles is considered owned by them[119], as profile owners. A profile owner is accordingly expected to be responsible of managing the Privacy Policy concerning all data on said profile. Nevertheless, logic and experience prove that a profile's contents contain data attached to more persons than the one holding ownership over the profile. For example, several users may appear in a same picture or other media content, such as videos and movies. Documents and other digital works can be co-authored or co-created and belong to multiple individuals. However, if User A uploads a photo in his profile of

[117] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

[118] Carminati, B., Ferrari, E. and Perego, A. (2006) "Rule-Based Access Control for Social Networks." Proceedings of the OTM Workshops, 1734–1744

[119] Note that "ownership" in this discussion is not defined in terms of legislation, but in terms of data and its relationship to the users.

himself and Users B and C, he is in charge of setting the Privacy Policy for that piece of data, regardless of whether User B is happy with that policy or not.
It seems obvious that ownership in Social Networks is a concept that needs to be detached from where – or by whom – a given piece of data gets uploaded. Simple arguments, such as the ones listed above, explain why the idea of co-ownership of data in SN can make a difference in today's situation.
Naturally, the challenge immediately becomes the identification of the co-owners of a given piece of data. This section focuses on a classification of SN users based on their relationship with their data, given by Squicciarini et al. in their work on Collaborative Privacy Management[120].
For the purpose of our discussion, we assume that our piece of data s is a picture or photo image. The idea of collective PM is generic and can be applied to any data type, so this assumption is not restricting our model. Three classes of users are defined: originators, owners and viewers. Users who originally post data s on their profile are classified as originators, while users with access to the data s are classified as viewers. Finally, users who share ownership of the data s with the originator within the network are classified as owners.
The potential owners of a data item posted on a profile are identified through the use of tagging features supported by current SN sites. In general, tagging consists of annotating social content by means of set of freely chosen words[121]. Their semantics can be analyzed by means of similarity tools[122]. In the case of pictures, the specific type of tags widely used in Facebook is employed for the model. These tags, known as id-tags, give the ability for users to add labels over pictures to indicate which users appear in them. Therefore, each id-tag essentially corresponds to a unique user id. Through access to id-tags, one can easily identify the potential owners of a given picture. We formally define potential owners as follows:

- *Definition 8.3*: If s a shared data item posted on user's i profile Profile_i and $TSet$ the set of tags associated with data item s, then the set of *potential owners* of s, $Pot\_Own_s^i$ is defined as the set of users whose id-tags appear in $TSet$.

Simplistic as it may appear, the above definition sets the groundwork for the collective Privacy Management model. For data types other than pictures, the set of potential owners can be identified by using the meta-data associated with the content, or though the originator's initiative. A user j belonging to the set of potential owners is only qualified as an owner of a data item s if the originator $\underline{i}$ agrees to grant him ownership it. Ownership privileges are exclusively granted by the originator to ensure that ownership is managed between users who are in fact not complete strangers, but related by a number of relationships that the originator deems acceptable. This network of admitted owners will be automatically specified by the originator through the application of distance-based policies that would be indicative of the type of relationships and the distance between the users.
Example 8.3: Consider Users A, B and C who are part of TheSN Social Network. Users A and B are friends, while B and C are partners. User A was present at the Christmas party of Users B and C's company. User A took pictures of all three of them and posted them on her profile in TheSN. User C ask originator User A to become an owner of the picture, as he happened to appear in it. User A's Privacy Policy for the Christmas photo album

[120] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY
[121] Wu, X., Zhang, L. and Yu, Y. (2006) "Exploring Social Annotations for the Semantic Web." Proceedings of the 15th International Conference on World Wide Web, New York, NY, 417–426, ACM Press
[122] Miller, G.A. "Wordnet: a Lexical Database for English." Communications of the ACM, New York, NY, Volume 38 Issue 11, 39–41

was $PrP_{\{FriendOf,PartnerOf\}}(Alice, 3)$. Since the distance between User A and User B through relationships $FriendOf$ and $PartnerOf$ is less or equal to 3, John is automatically granted the ownership.
The rules of the model imply that a set of owners, does not only decide whether to post/edit/delete a data item s, but more importantly shares the responsibility of managing access to s, by specifying the data Privacy Policies. This leads to a new status quo in Social Networking sites, not only from a Privacy Management point of view, but from a sociological point of view as well.
At this point, one might argue the obvious risk of originators not sharing ownership with entitled users. Squicciarini et al. have faced this issue and propose an incentive-based mechanism that motivates the sharing of ownership rights. The intrinsic workings of that mechanism are best described in the next section, where the Clarke-Tax based PM algorithm is directly presented.

8.2 A Collective Privacy Management Algorithm

The most intuitive approach to aggregate users' decisions consists of a combination of co-owners iteratively disclosing their preferred settings while explicitly agreeing on the set of viewers each owner proposes to include. During this process, owners should update their preferences after the review other co-owners' preferred settings, and try to reach a common settlement on a shared Privacy Policy.
However, this approach is not very effective, since it requires every single owner to agree to a unique and final set of privacy settings – a literally endless task. Additionally, since users typically access the SN independently, it is also impossible to force synchronization, inevitably introducing unacceptably long decision processes.

A more conservative solution is to construct a privacy policy that allows viewers' rights only to the set of users who satisfy each of the owners' preferences, avoiding the need of the owners explicit consent on the final set of viewers'. However, even this approach is pretty simplistic and fails to leverage the individuals' preferences within the co-owners' group. In addition to the identified drawbacks, it's been shown[123] that, majority and ranking-based approaches, such as the ones described above, have proved to be unfair, as astute individuals may manipulate outcomes at their advantage.
The approach suggested by Squicciarini et al. is based on two main ideas:

1. An algorithm that promotes certain desirable behaviors (i.e., granting ownership where it is due and being truthful towards co-owners about privacy preferences). More specifically, an application of the Clarke-Tax mechanism [124] designed to enforce Collective Privacy Management decisions.
2. A deduction (inference) technique, aiming to save users from having to input the same privacy settings multiple times for similar data. It will be based on the users' previous privacy decisions, and applied whenever certain similarity conditions hold true.

8.2.1 Credit Bargaining in Privacy Contexts

We are now going to describe the basic notions the incentive-based mechanism of Squicciarini et al. designed for SN users to share data and, at the same time, make thoughtful decisions regarding their privacy. First, a credit-based system is introduced. The user earns credits proportional to the amount of data (i.e. pictures, documents) he/she decides to expose to other users, as well as to the number of times he grants co-ownership to potential owners.

[123] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY
[124] Clarke, E.H. (1972) "Multipart Pricing of Public Goods: An Example," S. Mushkin (ed.), 125-130, Public Prices for Public Products, The Urban Institute, Washington

A user i is assigned an initial virtual counter (or "numéraire") $k_i \in \mathbb{R}$ to track his credits upon joining the SN. Mechanisms that credit and debit the counter are defined.

- *Definition 8.4*: For each uploaded data item s, shared among n co-owners, the counter c of originator i gains:

$$c = m_i + (\beta \times m_i) \times n$$

In the above definition, $m_i \in \mathbb{R}$ are the credits assigned to a data item (or type), while $\beta \times m_i, \beta \in [0,1]$ corresponds to the rise of the counters assigned for each user accepted as a co-owner for a data item. Each user accepted as a co-owner for s gains $\alpha \times m_i, \alpha \in [0,1]$. As shown, the more the user shares the ownership of an item, the more he/she gets rewarded. It should be noticed that the user's numéraire (or counter rise) is credited (taxed) according to how significant the user's preferences were in reaching the group's final decision.

Example 8.4: Let us assume that each document uploaded in TheSN offers 100 points of the numéraire, while $\alpha = 0.7$ and $\beta = 0.5$. If User A uploads a document, granting co-ownership to Users B and C, his counter is increased by: $c = 100 + (0.7 \times 100) \times 2 = 240$ points of the numéraire. At the same time, Users B and C increase their counters by $\beta \times m_i = 0.5 \times 100 = 50$ points each.

Through this procedure, the owners may keep the personal decision of whether to upload a data item or not, but they proceed into collective agreements regarding the exposure preferences of their uploaded data to potential viewers. Users will associate a *value* with each data preference, represented by function $v_i(g)$, that depicts the perceived benefit of the user, where he to expose a data item with preferences setting g. For example, a user who is interested in maximum disclosure would assign a high value to Privacy Settings g in order not to limit disclosure and allow more users to view his data.

Naturally, there will be cases when multiple users will have to be involved in a single decision. In those cases, they may select different optimal choices. Therefore, a new function is defined, known as the Social Utility Function[125].

- *Definition 8.5*: The Social Utility Function takes all the individual privacy preference functions as input and produces a certain unique collective output X:

$$F\big(v_i(g), \dots, v_n(g)\big) = \mathrm{X}$$

The fundamental requirement of a decision function is that it should produce an "optimal" in some sense. "Optimality", however, is not the most well defined concept itself. Different kinds of desirable attributes that characterize optimality have been suggested in various decision functions in Game Theory, Economics and Voting Theory. As a rule, these attributes care for the influence of the individual user on the collective outcome, as well as the impact of the outcome on the individual. Some common criteria include Pareto Optimality, Symmetry, Fairness, and Individual Rationality.

In the context of the users of the Social Networking site, measuring global utility is not a obvious process. Pure utility values, such as income and fairness, do not seem to be enough for the task; extra ones might need to be taken into account. It is at this point that, one simple approach, is chosen by Squicciarini et al. to fill the gap.

Rather common in Game Theory due to Nash[126] it often proves very effective to choose the outcome based on the maximization of the collective values (utilities). This approach, as we will see, satisfies three important properties[127]:

---

[125] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

[126] Mas-Colell, A. and Whinston, M.D. (1998) "Microeconomic Theory." Chapter 23, Oxford University Press, 4th Edition

[127] Ephrati, E. and Rosenschein, J.S. (1991) "Voting and Multi-Agent Consensus." Proceedings of the 9th National conference on Artificial Intelligence, Vol. 1, AAAI Press

- It guarantees a relatively fair distribution of the mutually earned utility
- It is simple
- It is non-manipulable

8.2.2 Privacy as a Tax Problem

The idea of the model's mechanism is to combine all the individual wishes into one group preference. The aim of this "aggregation" would be to adequately unify (and subsequently dissolve) the differences among the separate users' preferences. For this to be achieved, each user needs to offer some guidelines to the system, as to how important each preference stands according to his perception of privacy. This works through the proportional association of a value $v_i(g)$ to each preference g with relevance to importance.

It should become clear at this point, how the incentive-based system we described in the previous paragraphs offers a real incentive for users to deal with their privacy. Whether they are introvert of extrovert, or anywhere in the middle of this scale, the system rewards them accordingly and keeps them interested, caring only that they care – raising their awareness.

Given n co-owners of a data item s for which privacy preferences $g \in G$ need to be setup, each co-owner i can essentially opt for the different possible privacy preferences by assigning their value $v_i(g)$ for each $g \in G$. In this paper, we consider the additive social utility, which for a given preference g is the sum of value $v_i(g)$ for all the co-owners, where:

$$F\big(v_1(g), v_2(g), \dots, v_n(g)\big) = \sum_{i=1}^{n} v_i(g)$$

Since synchronization is not possible in SN sites, we let the users express their net values privately. Afterwards, the system calculates the collaborative outcome as a maximization of the collective social value:

$$g^* = arg_{g \in G} max \sum_{i=0}^{n} v_i(g)$$

As one may remember, $v_i(g)$ stands for the increase in user i's counter – or numéraire. The essence behind the above calculation is that we attempt to maximize the sum of the net value of each separate user's increase in their counters, over an item's privacy. The outcome g^* is the setting that maximizes the social utility function (definition 8.5).

From this point on, things are simple. The concept of the Clarke-Tax mechanism dictates that, if an outcome g is adopted, then each user i is required to pay "tax" π_i. Finally, the utility of a choice $c = (g, \pi_1, \pi_2, \dots \pi_n)$ equals the value of a preference g minus its tax raise π_i:

$$u_i(c) = v_i(g) - \pi_i.$$

The model of Squicciarini et al.[128] utilizes the Clarke Tax mechanism by maximizing the social utility function and encouraging truthfulness among the individuals, regardless of other individuals' choices. This algorithm requires each user to state the net value $v_i(g)$ for their preference simultaneously. Unlike the original Clarke Tax mechanism, this formulation does not require a fixed cost to be paid by the n co-owners. We consider therefore the fixed cost to be equal to 0. The tax levied by user i is calculated based on the Clarke-Tax formulation as follows:

$$\pi_i(g^*) = \sum_{i \neq j} v_j \left(arg_{g \in G} max \sum_{i \neq k} v_k(g) \right) - \sum_{i \neq j} v_j(g^*)$$

[128] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

Note that the two parts that compose the tax $\pi_i(g^*)$ of user i are computed over the preferences of all users but user i. The first part calculates the social utility for the outcome produced by computing a collaborative solution without taking user i into account. The second part computes the social utility for the outcome g^* – again excluding user i. Final tax $\pi_i(g^*)$ is defined as the difference between the first and the second part[129].

We have already assumed that each co-owner i can choose privacy preferences based in the path distance between the network graph nodes, which take values from $g \in \{0, n_{RSet}, \infty\}$, denoting owners only privileges with $\{0\}$, n-distant viewed stuff with relations in $\{RSet\}$ and public data with $\{\infty\}$, respectively. In case $n_{Friends}$ is the winning option, the set of final viewers is identified as the conjunction of the *pivotal users* friends' set. That is, $Profile_1.Friends \cap \ldots \cap Profile_n.Friends$.

	$v_i(g)$			
u_i	0	n	∞	$\pi_i(g^*)$
u_1	4	2	0.5	0.5
u_2	0	1	4	0
u_3	0.5	4	1.5	1.5
$\sum_i v_i(g)$	4.5	7*	6	
$\sum_{i \neq 1} v_i(g)$	0.5	5	5.5	
$\sum_{i \neq 2} v_i(g)$	4.5	6	2	
$\sum_{i \neq 3} v_i(g)$	4	3	4.5	

Table 8.1: An Example of the Clarke-Tax Mechanism (note that the outcome $g = \{n\}$ maximizes the social value with a value of 7)

As expected, each user indicates his own respective $v_i(g)$ value for each of the preferences in $g \in \{0, n_{RSet}, \infty\}$. Table 8.1 (from the work of Squicciarini et al.) shows an example including three users, each user i places their values $v_i(g)$ as indicated. The users u_1 and u_3 are the *pivotal users* and get taxed for their contributions to the social value function. User u_2 only contributed $v_2(n) = 1$ which was not pivotal to the decision made, thus user u_2 was not taxed.

8.2.3 Truthfulness and the Importance of Clarke-Tax

What is hugely significant about the Clarke-Tax mechanism is that it ensures users have no incentive to lie about their true intentions. Scuicciarini et al. show why the Clarke-Tax approach maximizes the users' truthfulness by an additional, simpler example. Consider two individuals a and b and a particular picture p: user a feels that the privacy settings on the picture should be private (option $g = 0$), and what he is willing to spend in order to keep the picture private among the owners is $v_a(0) = 20$. User b, on the other hand, wishes to keep the picture public (option $g = \infty$) and is willing to spend $v_b(\infty) = 10$ to do so. Suppose the maximum users a and b are willing to spend is denoted by $\overline{v_a}$ and $\overline{v_b}$ respectively. Also, suppose that the best response for users a and b is denoted by $\widehat{v_a}$ and $\widehat{v_b}$ respectively. The charge mechanism in this case goes like this:

$$\pi_a = \begin{cases} 0, & \widehat{v_a} < \widehat{v_b} \\ \widehat{v_b}, & \widehat{v_a} \geq \widehat{v_b} \end{cases}$$

Essentially, if user a wins he will be charged an amount that is as equal to the loss of the other owner, user b follows a similar formulation. In this case, user a's best response is:

$$\widehat{v_a} = \begin{cases} [0, \widehat{v_b}), & v_a < \widehat{v_b} \\ [max\{0, \widehat{v_b}\}, \overline{v_a}), & v_a \geq \widehat{v_b} \end{cases}$$

[129] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

It is important to notice that $v_a = \widehat{v_a}$ always falls in the range for the best response in both cases. Given that users a and b are both truthful, the wish of user a will prevail, and he will have to pay tax $\pi_a = 10$ in order to see his wish enforced. If user a aims at spending less and declares, falsely, $\widehat{v_a} = 11$, he may still win, but will still have to pay tax $\pi_a = 10$, according to the mechanism. Thus, underestimating the real value that someone holds for privacy is not going to help the result of the taxing process. Similarly, even if b declares less than what he thinks the real value is, since the credit is not going to be reimbursed at him, he is not going to get any advantage through his lie[130]. The simplicity of strategy is highly desirable in the design of solutions for this type of privacy problems, where users most likely are going to make intuitive and simple decisions to address their privacy considerations.

8.3 Inference Logic in Privacy Reasoning

One of the ominous disadvantages of the approach described in the previous section is that it requires manual input for each and every one of the pictures co-owned. Users may have up to hundreds of uploaded data items, and a significant percentage of them are going to be co-owned. As such, asking users to repeat the bidding procedure for each of them will be very cumbersome and is, in the long run, bound to failure. An effective idea to overcome this problem is to utilize inference-based techniques and exploit the results of previous decisions in order to free the users from repeating voting processes numerous times for similar cases.

For example, it is easily verifiable, that most users appear in their pictures with more or less the same small set of other users; typically directly related among each other. Also, the sensitivity of a given picture depends generally upon the context in which the picture has been taken. Building upon observations such as these, the use of tags and similarity analysis has proven to be an excellent inference tool in the determination and suggestion of best privacy policies for pictures shared among owners who have a history of shared content.

As already mentioned at the beginning of this chapter, users often use free text, in order to associate a context or a topic with their content. In the case of pictures, content tags can be added to each picture, or at the album level[131]. For simplicity the focus here is on the case where users add up to one tag each per picture. As such, for a given picture owned by k users, k tags are associated at most. This meta-data is used to conduct a Similarity Analysis with pictures already shared by the same set of users.

A Similarity Analysis constitutes of two major parts. First, a way to define the similarity of tags needs to be established. In order to utilize similarity metrics, we rely on the informal classification system that results from the practice of users' collaborative tagging. This user-generated classification system[132], is referred to as *folksonomy*[133], and is generally defined in terms of a collection of posts, each associated with one or more tags.

- *Definition 8.6*: A *folksonomy* is a tuple $F := (U; T; R; Y)$, where U, T and R are finite sets, whose elements are users, tags and resources, respectively. Y

[130] Ephrati, E. and Rosenschein, J.S. (1991) "Voting and Multi-Agent Consensus." Proceedings of the 9th National conference on Artificial Intelligence, Vol. 1, AAAI Press

[131] Content tags are not to be confused with id tagging, which we used to identify pictures' potential owners.

[132] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

[133] Mathes A. (Dec 2004) "Folksonomies – Cooperative Classification and Communication Through Shared Metadata," Graduate School of Library and Information Science, (http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html, accessed June 2, 2011)

is a ternary relation between them, i.e. $Y \subseteq U \times T \times R$. A post is a triple $(u; T_{ur}; r)$ with $u \in U, r \in R$ and $T_{ur} := \{t \in T | (u; t; r) \in Y\}$.

With the use of the folksonomy, the mechanism can compare any two pictures and assign them a similarity value, based on their tags. Tags relatedness may choose to rely on various metrics[134]. In this case, it is based on the frequency of occurrence of tag pairs. Based on this idea, similarity of data (especially pictures) is defined as the overall relatedness among the tags associated with it.

Example 8.5: Let us base this on Example 8.3. User A tags the new picture, Pic_{new}, as "*party*", when User B tags it as "*fun*" and User C as "*night*". If Users A and B already share among each other a couple of other pictured with freely chosen tags, Pic_1 tagged under "*gathering*, *fun*, *game*" and Pic_2 tagged under "*friend*, *beer*, *home*", we can assume that the *similarity value* between Pic_{new} and Pic_1 is bigger than that between Pic_{new} and Pic_2. Since Pic_1 is more similar to Pic_{new}, its privacy policy will be proposed to the three co-owners as the best suggestion for the new picture.

The privacy policy associated with the new item is prompted to all the users in the group of co-owners. If the users agree on the inferred privacy policy, it is used, and the tax pay is the same as the one originally spent for the picture that won the similarity contest. If the users do not agree, or no picture significantly similar to the new picture is found, the auction mechanism described previously is proposed to the users. Until a final decision is taken, a temporary policy, chosen among previously adopted ones is used.

8.4 Experiments and Results on Collaborative PM

In the next paragraphs we shall present a proof-of-concept social application of collaborative PM of shared data, implemented by Squicciarini et al.; it is being referred to as Private Box[135]. Private Box is fully integrated with the Facebook Social Networking site and supports the following features: controlled sharing of pictures; automatic detection of pictures' co-owners based on id-tags; collective privacy policies enforcement over shared pictures based on auctions.

The exact mechanism and technical details of Private Box are not of significance to our presentation and can be found on the work of Squicciarini et al. It is enough to say that it offers full auction functionalities and, at the same time, implements the bulk of the requirements requested by our definitions of collaborative Privacy Management in the previous sections of this chapter. Nevertheless, the inference component of the system is not currently implemented, and its deployment is part of future work.

According to research related to face recognition[136] in online albums there are between 2 to 4 faces per photo[137]. The scalability of the collaborative privacy policies enforcement has been evaluated based on auctions by varying the number of co-owners that appear in a photo under auction from 2 to 12. Figure 3 reports the execution times to perform Clarke-Tax algorithm once all the co-owners have placed a bid, while varying the number of co-owners. In other words, the graph shows the execution time of finding a privacy setting,

---

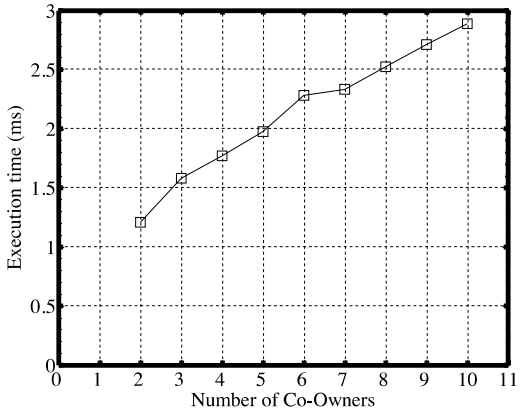
[134] Pirro, G. and Seco, N. (2008) "Design, Implementation and Evaluation of a New Semantic Similarity Metric Combining Features and Intrinsic Information Content." Proceedings of the OTM 2008 Confederated International Conferences, CoopIS, DOA, GADA, IS and ODBASE 2008. Part II on "On the Move to Meaningful Internet Systems", 1271-1288, Springer, Heidelberg

[135] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

[136] Davis M., Smith M., Canny J., Good N., King S. and Janakiraman, R. (2005) "Towards Context-Aware Face Recognition." Proceedings of the 13th Annual ACM International Conference on Multimedia, 483–486, ACM, New York, NY, USA

[137] Naaman, M., Yeh, R.B., Garcia-Molina, H. and Paepcke, A. (2005) "Leveraging Context to Resolve Identity in Photo Albums." Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries, 178–187, ACM, New York, NY, USA

which satisfies each co-owner privacy preference, and of calculating the bid score to be levied to the pivotal users.



Graph 8.2: Repeated Algorithm Execution

The execution time linearly increases with the increase of the number of co-owners because the Clarke-Tax Algorithm has to find the maximum for function $F(.)$ over a greater number of co-owners bid scores. However, the increase is negligible with respect to the number of co-owners. The execution time is so fast that the collaborative enforcement of privacy policies is transparent to the user.

9 The Combined Model

This chapter is devoted to the suggestion of a synthesis. As most readers must have understood by this point, we aim to suggest how the Collaborative Privacy Management model of Squicciarini et al.[138] can be improved through the Appropriation Measures generated through the AST approach of Dwyer et al.[139]

Collaborative Privacy Management offers the world of Social Networking a new philosophy together with a wide range of benefits, not only in compliance with the ordinances of Web 2.0, but also in direct coordination with them. Privacy is about personal information, and information in the online world corresponds directly to data. And data is anything but personal; it is shared. This is exactly what makes privacy such a difficult matter to address in online social communities.

Ownership of data is not like ownership of property. When four people appear together in a picture, the ownership of this item is equally divided among them – they are co-owners. Collaborative Data Management leads directly to Collaborative Privacy Management, and the methods and algorithms suggested and analyzed in chapter 8 have shown optimistic experimental results based on the current situation for contemporary Social Networks.

The model of Squicciarini et al. covers plenty of significant angles and requirements providing a mechanism that promotes truthfulness and, at the same time, enhancing the system with an inference algorithm in order for users to avoid repeating similar procedures again and again. However, it does not boast any efficient way to measure the users' adaptation with the Privacy Management model and facilitate feedback.

9.1 The Main Concept

The main concept of our idea is to improve the collaborative Privacy Management model of Squicciarini et al. by appropriately applying the AST Framework designed by Dwyer et al. to measure user appropriation and adaptation in the privacy settings and provide intelligent feedback.

As described in chapter 7 of the Thesis, every social system can be entirely described through structure, modality and interaction[140]. Structure consists of rules and resources – constraining and enhancing agents (in our case, SN users) respectively. Under the modality of a social system we understand the means by which its structures are transformed into actions. Finally, interaction can be any activity between a user and the social system (or network).

Furthermore, Appropriation is defined as the ongoing processes and methods by which people adopt and adapt to technology (see ch.7). Through the proper application of appropriations, one may draw results on the intricate workings of any technical system – how it operates, how it gives results. In the case of Privacy Management, Appropriation can be used to measure how users adopt the new privacy settings and features and to provide accurate feedback on the se of new privacy policies. According to Markus et al., AST proposes a non-deterministic view of technology use that provides room for social, cognitive and technical factors and thus, serves as an excellent foundation for building measures for online PM. However, the part they play on how users interact with the technology is very important. In the end, the concept of appropriation implies adaptations

---

[138] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

[139] Dwyer, C., Hiltz, S.R., Poole, M.S., Gußner, J., Hennig, F., Osswald, S., Schließlberger, S. and Warth, B. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." Proceedings of the 43rd Hawaii International Conference on System Sciences, Volume: 1, Publisher: IEEE Computer Society

[140] Giddens, A. (1986) "Constitution of Society: Outline of the Theory of Structuration." University of California Press; Reprint edition (Jan 1, 1986) ISBN 0-520-05728-7, 2, 281-348

in both directions – technological novelties causing changes in human behavior, so much as common practices leading to new technologies[141].
It is obvious that a system affected so much by human behavior within social frameworks – such as the collaborative Privacy Management model proposed above – is in direct need of the characteristic assets of AST.

9.2 Design of the Combined Model

We shall base our design on the Collective Privacy Management Algorithm of Squicciarini et al., described on chapter 8. We also plan to suggest a model for future experimental practice, based on the already successful Private Box, whose promising results can be found here[142]. Since Private Box is fully integrated with the Facebook Social Networking site, we suggest using the "Measure derivation method" especially for this SN, as described in section 7.4.1 of this document. We see below the table with the "rotated solution" of the Measures' factor loadings, resulting from the "factor rotation method" of Hair et al.[143] that was done by Dwyer et al. before us.

Meas.	1	2	3	4	5
Use 3	**.866**	-.130	.174	.040	-.087
Use 2	**.827**	-.215	.238	.039	-.043
Rej 3	**-.805**	.214	-.093	-.042	-.024
Use 4	**.708**	-.364	.198	.052	.090
Fth2	-.237	**.799**	-.210	-.001	.167
Fth1	-.155	**.737**	-.371	.128	.107
Fam2	.189	-.083	**.798**	.097	-.030
Fam4	.078	-.283	**.722**	-.103	-.094
Fam1	.153	-.177	**.678**	-.135	-.138
Sco4	-.017	.002	.048	**.865**	.065
Sco3	.036	.127	-.026	**.837**	-.008
Sco2	.075	-.060	-.129	**.776**	-.096
Fth5	.004	.087	-.041	.049	**.887**
Fth4	.006	.093	-.113	-.085	**.868**

Table 9.1: Factor Loadings for Facebook-oriented Study

In the experiment described in chapter 7, the results of the rotated factor analysis led to a reconstruction of some of the core constructs under measurement, analyzed thoroughly in section 7.3.1. The Rejection appropriation move was dropped, because its measures loaded strongly on other constructs. For the Use appropriation move, one initial measure was dropped (Use1) and replaced with Rej3. The Familiarity move and the Restricted Scope move now have three measures each.
In the new model, a factor rotation method with the same five core factors needs to be re-applied, this time on Squicciarini's Private Box design. Factor loadings shall be

[141] DeSanctis, G. and Poole, M.S. (1991) "Understanding the Differences in Collaborative System Use through Appropriation Analysis." 24th Annual Hawaii International Conference on System Sciences, Kauai, HI
[142] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY
[143] Hair, J., Black, W., Babin, B., Anderson, R. and Tathman, R. (2006) "Multivariate Data Analysis." Prentice Hall, Upper Saddle River, New Jersey, 2006

determined and thus, their corresponding final Measures (see table 7.1). These, in turn, shall be applied on the collaboratively operating Privacy Management System of Private Box and offer us an insight, not on how Facebook Privacy Policies are adopted by their users today, but on how they would be adopted, if Privacy Management was collaborative.

9.3 What is to be expected from the Model

The model proposed above is a wasted idea if not embraced and driven to scientific praxis. Consequently, it is my opinion that, no further conclusions can be drawn, before practical experiments ensue. However, I would like to establish what makes me believe in the model's design, as well as what is to be expected from its implementation.

For us the importance of feedback in a system based on human interaction is immense. No matter how efficiently a collaborative PM model (like that of Squicciarini et al.) may operate, the experiments of Dwyer et al. have shown that the user is always the final judge.

The synthesis that we suggest is not only feasible – since the two concepts model different parts of the PM system – it will produce a construct brandishing the merits of both formulae. Items will be auctioned among users through the use of the Clarke-Tax, or simple co-owned via intelligent inference, while at the same time, privacy settings will be appropriated according to properly designed measures, which will be continuously re-evaluated and, thus, keep the system in a self-improvement cycle indefinitely.

It is clear for us, that such a design constitutes a step towards the future. At the same time, it definitely does not reach the limits of the system. The next chapter tries to focus on the limitations of this model; right after it presents the perceived achievements of this Thesis as a whole.

10 Achievements and Limitations

As has already been mentioned in chapter 2 of this Thesis, the three main goals of this work are:

1. To raise awareness regarding the issues of Privacy Management in SN sites.
2. To collect all the related literature in a rather interdisciplinary written form.
3. To contribute an idea to the "best practices" of the field.

As to which extent the first goal will be achieved, is not so much a question, as it is a wish. However, we would like to claim that both the other two goals where achieved through the pages of this study.

By approaching the issue of privacy management from the angles of sociology, private life, law, work/academia and, of course, online social networking, we have managed, not only to make an attempt on the general awareness raise, but also to get a chance to present the best ideas and related literature on the matter from every point of view. We have shown that, privacy management is not a subject that its discipline can hope to address on her own. The people can only adopt the Internet when they trust it, and it is obvious from our work that, progress in that area is only feasible, when sociologists and computer scientists work together with legislators and businessmen.

However, the final days of our work did not present us only with the fulfillment of our goals – far from it. Despite the fact that we chose the best ideas we could find as basis for our suggestion, there were intrinsic defects and limitations within their core that no model so far has managed to overcome.

The Clark-Tax approach is far from perfect in itself. A significant drawback is the de facto assumption that users should be able to compute the value of the different preferences. We assume users can map the value to the number of users able to access the shared data, and this is possible using several social network indicators, such as the set of friends, the set of common friends, and several other network metrics[144]. This is totally unrealistic. The results of the experiments indicate that the users "learn" how to roughly estimate these values and assign these metrics, but how will these results variate when the Squicciarini auction system gets combined with the more sophisticated AST appropriation measurement model[145]? And how can we depend on the appropriation measures, if the users are only measuring stuff "roughly"? At the same time, one needs to keep in mind that the extend, to which human action is externally restricted to predefined rules and norms with the power to affect them through reflexive feedback mechanisms, is obviously still debatable.

These are only the most general of the combined model's limitations. From a detailed point of view, the limitations are many more and diverse, especially since the combined model has not been through an experimental stage yet. This is why we would consider the implementation of said model an interesting subject of future research.

[144] Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY

[145] Dwyer, C., Hiltz, S.R., Poole, M.S., Gußner, J., Hennig, F., Osswald, S., Schließlberger, S. and Warth, B. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." Proceedings of the 43rd Hawaii International Conference on System Sciences, Volume: 1, Publisher: IEEE Computer Society

11 Summary and Conclusions

The aim of this chapter is to collect and summarize all the critical points made during this work, from the literature search to the theory analysis and synthesis. The aim of this procedure is first of all organizational; nevertheless, it shall also procure a list of conclusions, useful to any who needs to assess what – if anything – this study has to offer to the community.

The beginning of this Thesis is formed with a presentation and an extended briefing on the concepts of online Social Networking and online Privacy Management. It continues with an outline of the threats that Privacy in Social Networking sites is facing and proceeds to the identification of the challenges of contemporary online Privacy Management. The area has recently grown of great importance and the literature produced in the last decade is immense. This gave fruit to the idea of attempting a rather interdisciplinary address of the matter, instead of following the strict "computer-science" approach.

This part of our research took much more time than any other, mainly because it was done with a widely inquiring attitude and on a much broader spectrum than the one it actually ended on. It was only after this part of the research, that the subject of the Thesis was actually finalized and its goals were precisely defined.

Our work aims at raising awareness on the alarming issues of Privacy Management in the exponentially growing area of online Social Networks. It aims at pointing out the respective threats and risks in all the communities involved, scientific or otherwise – and we realize that this is no easy task. It promotes the best practices proposed in the field of Privacy Management, and claims to stand on firm literature foundations. Finally, it hopes to make a valid suggestion on how the community could make progress.

Based on lessons learned, core requirements are identified to lay the groundwork for a privacy management model. However, the main goal in defining requirements is not only to have the compass necessary for designing our model, but also to discern which already existent models could help us in the process.

The next step led us to extensive literature research. Soon, previous experimental research was combined with other theoretical concepts into a model that, according to us, offered qualitative advantages to each of the previous models alone.

Chapter 7 of this Thesis is devoted to the concept of Adaptive Structuration Theory, and the Privacy Management model proposed by Dwyer et al., while chapter 8 analyzes the model of Squicciarini et al. and Collective Privacy Management.

The combined model is presented in chapter 9. It is basically a Collaborative Privacy Management model enhanced with Appropriation Measurement methods. It aims to possess the advantages of both previously presented models and offer combined services to users of Social Networking sites.

The last chapter of this work attempts to summarize what was achieved during these months of intellectual turmoil. The lessons learned, the literature discussed, the theories analyzed. The experiments studied, the questionnaires overviewed, the ideas proposed. The thoughts developed and the models constructed. The chapter closes with an identification of the limitations of the suggestion.

12 Appendix

12.1 Zusammenfassung (Deutsch):

Diese Masterarbeit beschäftigt sich mit Privacy Management in Fragen der Online-Welt, die durch die rasche Entwicklung der Soziale-Netzwerke-Technologien entstehen. Die heutige Situation im Bereich der Sozial-Netzwerke-Gemeinden, zusammen mit den jeweiligen Gefahren in Bezug auf Privacy Wahrung werden gründlich untersucht und dargestellt. Spezifische Ziele werden ausgearbeitet, über die Richtung die Bemühungen der wissenschaftlichen Gemeinschaft in naher Zukunft nehmen sollten. Die organische Lücken des Privacy Management Gebietes werden thematisiert und analysiert, da sie die Ursachen für eine Vielzahl von problematischen Situationen bieten, die kontinuierlich die gesellschaftspolitische Frieden unserer Gesellschaften in den letzten paar Jahren erregt haben.

Der Hauptteil dieser Arbeit beginnt mit der Identifizierung derjenigen Punkte, die eine de facto Lösung auf jeden Fall zu erfüllen muss, um die oben genannten Themen zu behandeln. Bewaffnet mit spezifischen Anforderungen, folgt eine Analyse durch die Wahl der theoretischen und praktischen Methoden. Ein Modell, dass den Spiel-Theoretischen Ansatz der kollektiven Privacy Management mit dem Konzept der Aneignung Konstrukt zu kombinieren versucht wird vorgeschlagen. Eine erweiterte Analyse erfolgt, über wie der Clarke-Tax-Mechanismus für das Kollaborative Management, verbindet mit dem Adaptive Strukturationstheorie, in Online-Datenschutz-Management verwendet werden können.

Bevor man zu den Schlussfolgerungen und Reflexionen der Arbeit kommt gibt es ein Kapitel über die Grenzen des vorgeschlagenen Modells. Es ist offensichtlich, dass das Konzept nicht generisch ist. Nichts desto trotz sollte eine deutliche Präsentation warum nicht aus dieser Studie fehlen.

Ein Wunsch, der zu dieser Masterarbeit beigefügt kommt, ist dass es mehr als ein Einführungsschritt der als akademischen Aufstieg eines jungen Mannes dienen wird. Es ist mit Hingabe an die Moralität die mit ihren Zielen zusammen kommt geschrieben, und damit den Wunsch, dass es gelesen wird und hoffentlich auch begrüßt.

FACHGEBIETE: Privacy Management in Soziale Netzwerke
STICHWORTE: Privacy, Soziale Netzwerke, Clarke-Tax Mechanismus, Aneignung Theorie

12.2 Abstract (English):

This MSc Thesis deals with Privacy Management issues in the online world that arise through the rapid evolution of Social Networking technologies. The current situation in the field of Social Networking communities, along with the respective dangers in regard to privacy safeguarding are thoroughly examined and presented. Specific goals are being set, as to the direction the efforts of the scientific community should take in the near future.

An extended survey in the field of Privacy Management on Social Networking Sites is followed by an analysis on the missing parts of this already huge living organism. These organic gaps need to be addressed firmly, as they constitute natural causes for a multitude of problematic situations that have continuously stirred the sociopolitical peace of our societies for the past few years.

The main part of this study starts with the identification of the singular points that any de facto solution is required to fulfill in order to address the issues mentioned above. Armed with specific requirements, an analysis is followed by a choice of theoretical and practical methods. A model that attempts to combine the game-theoretical approach of Collective Privacy Management with the concept of the Appropriation Construct is proposed. An extended analysis, on how the Clarke-Tax mechanism for collaborative management, conjoined with the Adaptive Structuration Theory, can be used in online privacy management, ensues.

Before coming to the conclusions and reflections of our work, a chapter is dedicated to the limitations of the proposed model. It is obvious that the present approach is not generic; however, a more explicit presentation should not be absent from this study.

The wish that comes attached to this MSc Thesis is that it will become more than a launching step to the academic advancement of a young man. It is written with faith and devotion to the morality that goes with its goals and thus, the wish that follows it is that it will be read and, hopefully, appreciated.

SUBJECT AREA: Privacy Management in Social Networking
KEYWORDS: Privacy, Social Networks, Clarke-Tax mechanism, Appropriation Theory

12.3 Curriculum Vitae

Last Name:	Perellis – Konstantinidis
First Name:	Leonidas – Dimitrios
Date of Birth:	Apr 19, 1984
e-mail:	leoperellis@gmail.com

Education:

2011: Spring:
MSc student in the University of Vienna (http://www.univie.ac.at/), in the field of Informatics Management.

- MSc Thesis:
  - Topic: «The Effect of Social Networking on Privacy Management»
  - Advisor: Univ.-Prof. Dipl.-Ing. DDr. Gerald Quirchmayr

2008: September:
BSc graduate (Ptychion) of the National and Kapodistrian University of Athens, in the field of Informatics & Telecommunications (http://www.di.uoa.gr), with a GPA of 7.4/10.

- BSc Thesis:
  - Topic: «Digital Identification and RFID – Issues and Morality»
  - Advisor: Prof. Panagiotis Georgiadis

2002: June:
Graduate of the German School of Athens (DSA) (http://www.dsathen.gr) with graduation grade 17.797/20. Holder of the German Abitur.

Educational Experience:

2010: July:
Participation in the IPICS 2010 Summer School (Intensive Program on Information and Communication Security) that took place in the University of the Aegean, Samos, Greece (http://www.ipics-school.eu/).

2008: March:
Participation in the IPICS 2008 Winter School that took place in the University of Lapland, Rovaniemi, Finland.

2007: Spring Semester:
Participation in the «Erasmus» program, as a student in the University of Vienna (http://www.univie.ac.at/), department of Informatics.

2005-08: For 3 academic years:
Elected representative of the students committee in the Senate of the University for 3 consecutive years, through the independent party of informatics' students (http://www.di.uoa.gr/~fpp).

Research Interests:

- Privacy issues
- Educational Software
- Human Resource Management
- System Analysis, Knowledge Engineering

Significant Projects:

- Research and development of a widely applicable e-voting platform in the premises of the Informatics & Telecommunications Department in NKUA.
- Founding member of the virtual business «Come Along», for the course of «Innovation and Entrepreneurship».
- Development of a time scheduling system for the flights and crew availability of an existing airline (Olympic Airlines) for the course of "Artificial Intelligence".

Languages:

Greek: native
English: proficient – holder of the Cambridge Certificate of Proficiency in English since 2000
German: proficient – holder of the Sprachdiplom of the German Abitur since 2002

Scholarships/Awards:

- Scholarship and Award of the State for achieving 1st position in the ranking of the University's first-year students, after the national examinations.
- Scholarship of the National and Kapodistrian University of Athens for the duration of the Masters Degree – success in examinations, June 2009.

Computer Related Skills:

- Programming: C, C++, HTML, PHP, SQL, Prolog, Assembly etc.
- Operating Systems: Windows, Unix based (i.e. Linux, Mac OSX, Solaris)

Professional Experience:

2008:	Cooperation with the STEPSIS Group. Installations of networks and Siemens «Instabus» systems and electronics in habitats and «smart» houses.
2006:	January – July: Translation of the book «Regards Croises de L'Union Europeenne», by Leonard Messi from English and German to Greek.
2003 – today:	Private lessons in Mathematics, Informatics and Physics for school pupils of various levels.

Other Interests:

- 11 years of classical piano studies, followed by two years of jazz studies. Member of various bands through the years.
- Blogging: http://timetoleave.wordpress.com
- Founding member and editor of the student magazine «Pliktro».
- Traveling, theater, philosophy

Catalog of Images, Tables and Graphs

[1] Image 3.1: Facebook Permission Dialog Box
Source: http://developers.facebook.com/blog/post/446

[2] Image 3.2: The World, drawn entirely through Facebook connections
Source: http://www.facebook.com/notes/facebook-engineering/visualizing-friendships/469716398919

[3] Table 4.1: Responses to the question: "Since you have created your profile, who do you think has looked at it?" over three consecutive years' surveys.
Source: Lampe, C. et al. (2008) "Changes in Use and Perception of Facebook." CSCW Conference, ACM, NY

[4] Table 4.2: Responses to the question "I use Facebook to…" rated on a Likert scale for likeliness (higher values correspond to higher likelihood to engage to the activity).
Source: Lampe, C. et al. (2008) "Changes in Use and Perception of Facebook." CSCW Conference, ACM, NY

[5] Table 4.3: Ratings of Attitudes towards Facebook
Source: Lampe, C. et al. (2008) "Changes in Use and Perception of Facebook." CSCW Conference, ACM, NY

[6] Table 7.1: Final Measures and their Factor loadings
Source: Dwyer, C. et al. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." 43rd HICSS

[7] Table 7.2: Comparison of Results: Facebook, MySpace, StudiVZ
Source: Dwyer, C. et al. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." 43rd HICSS

[8] Table 8.1: An Example of the Clarke-Tax Mechanism
Source: Squicciarini, et al. (2009) "Collective Privacy Management in Social Networks." 18th WWW, ACM, NY

[9] Graph 8.2: Repeated Algorithm Execution
Source: Squicciarini, et al. (2009) "Collective Privacy Management in Social Networks." 18th WWW, ACM, NY

[10] Table 9.1: Factor Loadings for Facebook-oriented Study
Source: Dwyer, C. et al. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." 43rd HICSS

Bibliography

- Acquisti, A. and Gross, R. (2006) "Imagined Communities: Awareness, Information, Sharing and Privacy on the Facebook." 6th Workshop on Privacy Enhancing Technologies, 36-58, Springer, Cambridge, UK
- Barnes, S. B. (2006) "A Privacy Paradox: Social Networking in the United States." First Monday, 11(9)
- Boyd, D. (2004) "Friendster and Publicly Articulated Social Networks." Conference on Human Factors and Computing Systems (CHI 2004), ACM, April 24-29, Vienna
- Boyd, D.M. and Ellison, N. (2007) "Social Network Sites: Definition, History, and Scholarship." Journal of Computer-Mediated Communication, 13(1), Article 11
- Bryant, S., Forte, A. and Bruckman, A. (2005) "Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia." ACM-Group, Sanibel Island
- Carminati, B. and Ferrari, E. (2008) "Privacy-Aware Collaborative Access Control in Web-Based Social Networks." In Proceedings of DBSec, 81–96
- Carminati, B., Ferrari, E. and Perego, A. (2006) "Rule-Based Access Control for Social Networks." Proceedings of the OTM Workshops, 1734–1744
- Chin, W.W., Gopal, A. and Salisbury, W.D. (1997) "Advancing the Theory of Adaptive Structuration: The Development of a Scale to Measure Faithfulness of Appropriation." Information Systems Research, 8 (4), 342
- Churchill, E.F. and Bly, S. (1999) "Virtual Environments at Work: Ongoing use of Muds in the Workplace," WACC'99, ACM Press, 99-108, San Francisco
- Clarke, E.H. (1972) "Multipart Pricing of Public Goods: An Example," S. Mushkin (ed.), 125-130, Public Prices for Public Products, The Urban Institute, Washington
- Davis M., Smith M., Canny J., Good N., King S. and Janakiraman, R. (2005) "Towards Context-Aware Face Recognition." Proceedings of the 13th Annual ACM International Conference on Multimedia, 483–486, ACM, New York, NY, USA
- Dennis, A., Wixom, B. and Vandenberg, R. (2001) "Understanding Fit and Appropriation Effects in Group Support Systems via Meta-Analysis." MIS Quarterly, 25 (2), 167-193

- DeSanctis, G. and Poole, M.S. (1991) "Understanding the Differences in Collaborative System Use through Appropriation Analysis." 24th Annual Hawaii International Conference on System Sciences, Kauai, HI
- DeSanctis, G. and Poole, M.S. (1994) "Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory." Organization Science 5(2), 121-147
- DiMicco, J.M. and Millen, D.R. (2007) "Identity Management: Multiple Presentations of Self in Facebook." Conference on Supporting Group Work, Sanibel Island, FL, ACM Press, 383-386
- Donath, J.S. (2007) "Signals in Social Supernets." Journal of Computer Mediated Communication, 13 (1), 12
- Donner, J. (2008) "The Rules of Beeping: Exchanging Messages via Intentional 'Missed Calls' on Mobile Phones." Journal of Computer-Mediated Communication, 13 (1)
- Dwyer, C. (2008) "Appropriation of Privacy Management Within Social Networking Sites." Information Systems Department, Ph.D. Dissertation, New Jersey Institute of Technology, NJ
- Dwyer, C., Hiltz, S.R. and Passerini, K. (2007) "Trust and Privacy Concern within Social Networking Sites: A Comparison of Facebook and MySpace." in 13rd Americas Conference on Information Systems, Keystone, Colorado
- Dwyer, C., Hiltz, S.R., Poole, M.S., Gußner, J., Hennig, F., Osswald, S., Schließlberger, S. and Warth, B. (2010) "Developing Reliable Measures of Privacy Management within Social Networking Sites." Proceedings of the 43rd Hawaii International Conference on System Sciences, Volume: 1, Publisher: IEEE Computer Society
- Ellison, N., Steinfield, C. and Lampe, C. (2007) "The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites." Journal of Computer Mediated Communication, 12 (4), Article 1
- Ephrati, E. and Rosenschein, J.S. (1991) "Voting and Multi-Agent Consensus." Proceedings of the 9th National conference on Artificial Intelligence, Vol. 1, AAAI Press
- Erickson, T. and Kellogg, W.A. (2002) "Social Translucence: Designing Systems that Support Social Processes." ACM Human-Computer Interaction in the New Millennium, NY, 325-345

- Felt, A. and Evans, D. (2008) “Privacy Protection for Social Networking Platforms.” Proceedings of Web 2.0 Security and Privacy 2008 (in conjunction with 2008 IEEE Symposium on Security and Privacy)
- Gates, C. (2007) “Access Control Requirements for Web 2.0 Security and Privacy.” In IEEE Web 2.0 Privacy and Security Workshop, Oakland, CA
- Geambasu, R., Balazinska, M. Gribble, S.D. and Levy, H.M. (2007) “HomeViews: Peer-to-Peer Middleware for Personal Data Sharing Applications.” Proceedings of SIGMOD Conference, 235–246
- Giddens, A. (1986) “Constitution of Society: Outline of the Theory of Structuration.” University of California Press; Reprint edition (Jan 1, 1986) ISBN 0-520-05728-7, 2, 281-348
- Gilbert, E., Karahalios, K. and Sandvig, C. (2008) “The Network in the Garden: An Empirical Analysis of Social Media in Rural Life.” ACM Conference on Human Factors in Computing Systems (CHI), Florence, Italy, 1603-1612
- Goffman, E. (1959) “The Presentation of Self in Everyday Life.” Doubleday & Co., Garden City, NY
- Golder, S., Wilkinson, D. and Huberman, B.A. (2007) “Rhythms of Social Interaction: Messaging within a Massive Online Network.” 3rd International Conference on Communities and Technologies (CT2007), East Lansing, MI, Springer
- Gross, R. and Acquisti, A. (2005) “Information Revelation and Privacy in Online Social Networks.” Workshop on Privacy in the Electronic Society, Alexandria, VA, ACM Press
- Hair, J., Black, W., Babin, B., Anderson, R. and Tathman, R. (2006) “Multivariate Data Analysis.” Prentice Hall, Upper Saddle River, New Jersey, 2006
- Hargittai, E. (2007) “Whose Space? Differences Among Users and Non-users of Social Network Sites.” Journal of Computer Mediated Communication, 13 (1), Article 14
- Hart, M., Johnson, R. and Stent, A. (2007) “More Content – Less Control: Access Control in the Web 2.0.” IEEE Web 2.0 Privacy and Security Workshop
- Hart, M., Johnson, R. and Stent, A. (2007) “More Content – Less Control: Access Control in the Web 2.0.” IEEE Web 2.0 Privacy and Security Workshop
- Hass, N. (2006) “In your Facebook.com” The New York Times, 8 Jan 2006

- Hempel, J. (2005) “The MySpace Generation.” Business Week, 12 Dec 2005
- Hobgen, G. (2007) “Security Issues and Recommendations for Online Social Networks.” ENISA Position Paper N.1
- Joinson, A.N. (2008) “Looking at, looking up or keeping up with People?: Motives and Use of Facebook.” ACM Conference on Human Factors in Computing Systems (CHI), Florence, Italy, 1027-1036
- Lakhani, K.R. and Hippel, E. (2003) “How Open Software Works: “Free” User-To-User Assistance.” Research Policy, 32(6), 923-943
- Lampe, C., Ellison, N. and Steinfield, C. (2006) “A Face(book) in the Crowd: Social Searching VS. Social Browsing.” ACM Special Interest Group on Computer-Supported Cooperative Work, ACM Press, Banff, Canada
- Lampe, C., Ellison, N. and Steinfield, C. (2007) “Profile Elements as Signals in an Online Social Network.” ACM Conference on Human Factors in Computing Systems (CHI), San Jose, CA
- Lampe, C., Ellison, N. and Steinfield, C. (2008) “Changes in Use and Perception of Facebook.” Proceedings of the ’08 ACM Conference on Computer Supported Cooperative Work (CSCW), ACM, NY
- Lessig, L. (1998) “The Architecture of Privacy.”
- Levi, H. (2005) “Reflexivity in Sociology”, New Dictionary of the History of Ideas
- Maag, C. (2007) “When the Bullies Turned Faceless.” The New York Times, 16 Dec 2007
- Mannan, M. and Van Oorschot, P.C. (2008) “Privacy-Enhanced Sharing of Personal Content on the Web.” Proceedings of WWW, 487–496
- Markus, M.L. and Silver, M.S. (2008) “A Foundation for the Study of IT Effects: A New Look at DeSanctis and Poole’s Concepts of Structural Features and Spirit.” Journal of the Association for Information Systems, 9 (10)
- Mas-Colell, A. and Whinston, M.D. (1998) “Micro-Economic Theory.” Chapter 23, Oxford University Press, 4th Edition
- Mendonça, D., Jefferson, T. and Harrald, J. (2007) “Collaboration Adhocracies and Mix-and-Match Technologies in Emergency Management.” Communications of the ACM, 50 (3), 45-49
- Miller, G.A. “Wordnet: a Lexical Database for English.” Communications of the ACM, New York, NY, Volume 38 Issue 11, 39–41

- Naaman, M., Yeh, R.B., Garcia-Molina, H. and Paepcke, A. (2005) "Leveraging Context to Resolve Identity in Photo Albums." Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries, 178–187, ACM, New York, NY, USA
- Nonnecke, B., Preece, J. and Andrews, D. (2004) "What Lurkers and Posters think of each other." 37th Hawaii International Conference on System Sciences, IEEE, HI
- Petronio, S. (2002) "Boundaries of Privacy: Dialectics of Disclosure." State University of New York Press, Albany
- Pirro, G. and Seco, N. (2008) "Design, Implementation and Evaluation of a New Semantic Similarity Metric Combining Features and Intrinsic Information Content." Proceedings of the OTM 2008 Confederated International Conferences, CoopIS, DOA, GADA, IS and ODBASE 2008. Part II on "On the Move to Meaningful Internet Systems", 1271-1288, Springer, Heidelberg
- Quirchmayr, G. and Wills, C.C. (2007) "Data Protection and Privacy Laws in the Light of RFID and Emerging Technologies." Proceedings of TrustBus, 155-164
- Read, B. (2006) "Think Before you Share: Students' Online Socializing Can Have Unintended Consequences." The Chronicle of Higher Education, A38, 20 Jan 2006
- Rosenblum, D. (2007) "What Anyone Can Know: The Privacy Risks of Social Networking Sites." IEEE Security and Privacy, 5(3), 40-49
- Rosenthal, R. and Rosnow, R. (1991) "Essentials of Behavioral Research: Methods and Data Analysis." McGraw Hill, New York, 1991
- Samuel, D.W. and Louis, D.B. (1890) "The Right to Privacy." Harvard Law Review, 15 Dec 1890
- Shneiderman, B. (2007) "Web Science: A Provocative Invitation to Computer Science." Communications of the ACM, New York, NY, 50 (6), 25-27 Mas-Colell, A. and Whinston M.D. (1998) "Micro-Economic Theory." Chapter 23, Oxford University Press, 4th Edition
- Smith, H.J., Milberg, S. and Burke, S. (1996) "Information Privacy: Measuring Individuals' Concerns about Organizational Practices. " MIS Quarterly, 20(2), 167-196
- Smith, M. "Measures and Maps of Usenet." In Lueg, C. and Fisher, D. eds (2002) "From Usenet to CoWebs: Interacting with Social Information Spaces." Springer Verlag, New York, NY

- Spiekermann, S., Grossklags, J. and Berendt, B. (2001) "E-privacy in 2nd Generation E-commerce: Privacy Preferences versus Actual Behavior." Proceedings of the 3rd ACM Conference on Electronic Commerce, 38–47, NY
- Squicciarini, A.C., Shehab, M. and Paci, F. (2009) "Collective Privacy Management in Social Networks." Proceedings of the 18th International Conference on World Wide Web, ACM New York, NY
- Thiesse, F. (2007) "RFID, Privacy and the Perception of Risk: A strategic Framework." Journal of Strategic Information Systems
- Varian, H.R. (2004) "System Reliability and Free Riding." Economics of Information Security, Kluwer Academic Publishers, 1–15
- Wang, C. and Fung Leung, H. (2004) "A Secure and Private Clarke-Tax Voting Protocol Without Trusted Authorities." Proceedings of 6th International Conference on Electronic Commerce, 556–565, ACM, New York, NY, USA
- Wu, X., Zhang, L. and Yu, Y. (2006) "Exploring Social Annotations for the Semantic Web." Proceedings of the 15th International Conference on World Wide Web, New York, NY, 417–426, ACM Press