Privacy by Design: Does it matter for Social Networks?

Mohammad Badiul Islam, Computer Science Discipline, Faculty of Science and Technology, Queensland University of Technology and NICTA (National ICT Australia), Queensland Research Lab, Brisbane, Australia, mb.islam@qut.edu.au.

Adpro Dr. Renato Iannella, Semantic Identity, Brisbane, Australia, ri@semanticidentity.com.

Abstract. Privacy *is an important component of freedom and plays a key role in protecting fundamental human rights. It* is becoming increasingly difficult to ignore the fact that without appropriate levels of privacy, a person's rights are diminished. Users want to protect their privacy - particularly in "privacy invasive" areas such as social networks. However, Social Network users seldom know how protect their own privacy through online mechanisms. What is required is an emerging concept that provides users legitimate control over their own personal information, whilst preserving and maintaining the advantages of engaging with online services such as Social Networks.

This paper reviews "Privacy by Design (PbD)" and shows how it applies to diverse privacy areas. Such an approach will move towards mitigating many of the privacy issues in online information systems and can be a potential pathway for protecting user's personal information. The research has posed many questions in need of further investigation for different open source distributed Social Networks. Findings from this research will lead to a novel distributed architecture that provides more transparent and accountable privacy for the users of online information systems.

Keywords: Privacy by Design, Social Networks, Privacy, Access Control, Mobile Social Networks, distributed Social Networks, Open Source Social Networks, Diaspora.

1 Introduction

Privacy is an important component of the freedom of a person and plays a key role in protecting fundamental human rights. It is becoming increasingly difficult to ignore the fact that without appropriate levels of privacy, a person's freedom can be diminished. Failing to protect anyone's privacy personal information affects everyone: friends, family, co-workers, relatives and so on. Any person has the right to share, disclose, access, rectify, delete, and block their own personal information unless there are legitimate reasons provided by the law [1]. However, privacy does not mean simply hiding information; it is the legitimate control over one's own personal information. Additionally, any person has the ultimate right and freedom to exit from the digital world. Without explicit consent, nobody has the right to access another person's personal information.

Users and consumers are beginning to show anxiety for privacy in different "privacy invasive" areas including Social Networks (SN), Cloud computing, Health records, Geo-location Services, Video Surveillance Cameras, Biometrics, Radio-Frequency Identifiers (RFID), Mash-up applications, Network monitoring and Whole body imaging, etc. Consumers' anxiety arises after experiencing incidents in their own lives that threaten their ultimate freedom. Not only users but also technology experts [2], researchers and industry professionals are expressing anxiety about privacy invasion areas. Unless we act now, privacy may not exist by the year 2020 [3].

Users want privacy but they seldom know "how to specify" and "what to seek" for their own privacy [4]. Embedded privacy-enhancing technologies (PETs) in the design level can be the solution for ensuring privacy from the beginning of a system development. An early publication by Langheinrich [5] describes six principles to ensure privacy during the system design process. Those principles are included as "notice", "choice and consent", "proximity and locality", "anonymity and pseudonymity", "security" and "access and recourse". Langheinrich also expresses deep concern that "something" needs to be done since private information may be stored forever in public places. Additionally, Common Criteria [6, 7] can be engaged to evaluate privacy compliance and how the of security properties of IT products and systems. Here, common criteria of a system can be established to reach a wider audience in Social Networks. However, Blarkom et al. [7] draws our attention to the need to avoid the common criteria framework for privacy audits since privacy obligations require system design processes incorporate privacy standards. Moreover, Blarkom et al. encourages utilizing 11 Fair Information Practices and 9 principles of the Privacy Audit Framework. In addition to this, "Privacy by Design (PbD)" [8], is a concept that can be used to protect Personally Identifiable Information (PII). The PbD concept includes seven principles. A brief discussion about PbD principles can be found in section 2. It has also been recommended to

use a more substantial approach to PbD principles, extending the PET Cavoukian proposed principles she calls PET Plus [8]. System development costs increase substantially in later stages so it is always good if privacy can be incorporate from the design phase of a system. To comply with PbD concepts, this research suggests that seven principles are required to be incorporated into a system at the design level. One of the objectives of this research is to encourage engaging privacy in the system design level since in the later stage of system is extremely difficult to incorporate privacy, whereas privacy functionality can easily be engaged in the initial design stage of the system.

This paper is organized into four parts. The first part of this paper presents an overview of PbD principles. The second part of this paper reviews diverse approaches of PbD principles in different privacy invasive areas including Social Networks. The third part presents a case study examining the claim by Diaspora that they have declared privacy-aware distributed open source Social Network. The final part discusses different barriers for adopting PbD principles.

This paper is the first study, to date to investigate how privacy can be ensured in different privacy invasive areas including Social Networks through the PbD principles. It seeks to address the enhanced understanding of the PbD principles and how those principles can be utilized in Social Networks.

2 Privacy by Design (PbD) principles

The term "Privacy by Design (PbD) [8] was conceived by Dr. Ann Cavoukian in early 1990. Gradually the author has modified the PbD principles and down to seven key principles (Table 1). PbD principles are not limited to, but include "Information Technology", "Accountable Business Practices" and "Physical Design and Infrastructure" areas. So far, PbD principles remain at the conceptual stage. To comply with the PbD concept and to ensure privacy, a system has to be systematic, predictable and repeatable [9].

Table 1. Privacy by Design principles [8] and analysis

	Principle	Principle Details	Comment
1	Proactive not Reactive; Preventative not Remedial	Privacy by Design comes before-the-fact, not after.	The principle underpinning the mechanism is how the information privacy will be observed and resolved before problems arise.
2	Privacy as the Default	No action is required on the part of the individual to protect their privacy — it is built into the system, by default.	The principle underpinning the rules is how the information will be collected and used with respect to individual privacy.
3	Privacy Embedded into Design	Privacy is integral to the system, without diminishing functionality."	The principle underpinning the mechanism is how to implement the system policies to ensure user privacy.
4	Full Functionality – Positive-Sum, not Zero-Sum	Privacy by Design demonstrating that it is possible to have both such as privacy vs. security	The principle underpinning the methodology is how to create full functionality while protecting individual privacy.
5	End-to-End Lifecycle Protection	Privacy by Design ensures cradle to grave, lifecycle management of information, end-to-end.	The principle underpinning the assessment is how retire information will be pre-processed to ensure individual privacy.
6	Visibility and Transparency	Privacy by Design comes before-the-fact, not after.	The principle underpinning the investigation is how the accountable organization will be open and honest with individual privacy.
7	Respect for User Privacy	No action is required on the part of the individual to protect their privacy — it is built into the system, by default.	The principle underpinning the investigation is how to share, disclose or access, rectify, delete, and block information that is consistent with respect to individual privacy.

PbD principles can be used for adopting Privacy Enhancing Technology (PET) directly at the system design level. Adopting PET will increase the user satisfaction and confidence in using any system. PET can minimize unnecessary disclosure, collection, retention, sharing, trading and unauthorized using of personal information. Additionally, PET can ensure legitimate rights to control their own private information. This will assist in gaining confidence and trust to use the system since the user is capable of controlling own private information. That may eventually lead to an increase in the user reliability of the system.

3 Privacy by Design approaches

PbD applies to diverse privacy invasive areas including Social Networks. It is becoming increasingly difficult to ignore the user privacy and importance of personal information in different privacy invasive areas. One of the most significant current organizational discussions is to protect user information as well as protecting business value in those areas. Recently, researchers also have shown an increased interest in protecting user information in miscellaneous privacy invasive areas utilizing PbD approaches. This section reviews the diverse approaches concerning the effectiveness of using PbD principles.

Radio Frequency Identification (RFID) within the health sector is an area that requires privacy measures [8]. Cavoukian suggested that RFID should be involved in the health sector without any linkage to personal identifies to avoid privacy threats early. Williams and Weber-Jahnke [10] provided three solutions to prevent privacy breaches in Healthcare Social Networks. They included automated queries to detect fake user accounts and developing improved business processes to detect credentialed users and prevent users from locating hidden network information. Conducting Privacy Impact Assessment (PIA) is one of the early approaches for preventing privacy threats [11]. However, PIA should be repeated after a period of time (half yearly or yearly) to update the impact from the previous PIA. Hence, preventive but not remedial privacy features must be an aim for the SN. Not only providing the proactive features but also raising awareness among users can assist to improve privacy before incidents occur.

Privacy should be built-in into the system to protect a user's private information. The service/platform provider should ensure privacy by default. Creating a User-Centric Identity Management Infrastructure can be another approach for default privacy. If there are any changes in the system the user has to approve the update and receive feedback. Then they might change the privacy settings [12] and relax their privacy. One of the approaches to ensure default privacy can be generalizing the information after a period of time. Williams and Weber-Jahnke [10] suggested that by automatically generalizing the accessible information to an inactive connection. Although built-in default privacy might be appreciated, it is unlikely but in reality, the default privacy settings can be a source of different privacy breaches such as leaking information to an untrustworthy third party. Therefore, the user should be well informed about the status of the default privacy settings.

Privacy embedded in design is a key concern for implementing privacy. One of the privacy-invasive sectors - Biometric Encryption - can utilize PbD principles to provide privacy and ensure full functionality [13]. Another privacy invasive area is Video Surveillance that may ensure public safety with respect to governance and law-abiding citizens privacy [8]. One approach can be to publish general information on a website to inform the citizens about the public video surveillance locations and reasons. Williams and Weber-Jahnke [10] suggested two mechanisms that can be incorporated in the system design level. Hence, Privacy embedded in the design should be the key approach for Social Network. Additional PET solutions should be considered for embedding privacy in the system design level for SNs.

The future of privacy-preserving SN applications is expected to be a win-win scenario that will not destroy the business model for service providers [14] as well as protect the users privacy. There is a myth that one goal is achieved at the expense of other goal which may not true for many cases especially in the health sector which is not always true. Obviously, a positive-sum paradigm is achievable in the system design [8]. In addition to this, Encrypting User ID [15] can be one solution without revealing private information to a third party since personal information becomes the asset for many enterprises [16]. Additionally, private identifier such as Credentica/Microsoft Private Digital Identity [17] can be a possible solution for minimal disclosure. However, mobile SN, personal information has to be accessed with special care since real-time spread can cause an immediate disruptive affect to user's social life.

Hence, Privacy Enhanced Technology (PET) and Transparency-enhancing technology (TET) are required to be incorporated in the initial design level to ensure user privacy [18]. However, the service providers should take necessary steps to decrease the user's burden as well as respect the user's privacy [10]. User centric identity management should be solution future privacy-aware Social Network.

4 Case study: Diaspora

Diaspora [19] claims to be a privacy-aware, personally-controlled and distributed open source Social Network to replace centralized social networks since these have failed to protect the user's privacy. Diaspora states its aim is to protect user information in a philosophy of "secure as much as you can, but no more". Diaspora also claims to make private sharing easy and simple without increasing the user's burden. The Diaspora architecture (Fig. 1) includes a Server (Pod) to host seeds (i.e., user accounts) and claims that the seed is owned by the user which then can be used to aggregate other profiles, tweets or social data.

We have reviewed Diaspora to analyze whether it follows the PbD principles. The privacy-aware Diaspora Social Network has been investigated in terms of how it follows the PbD principles. Diaspora system has selected randomly but for some consequential reason for evaluation. First of all, Diaspora system grabs the attention by the media [20] and technologists [21, 22]. Some technologist claimed that Diaspora might receive attention by the user for privacy issue and the user might change their might to use the well recognized Social Network [23]. Therefore, Diaspora can be a possible test case for evaluating PbD principles.

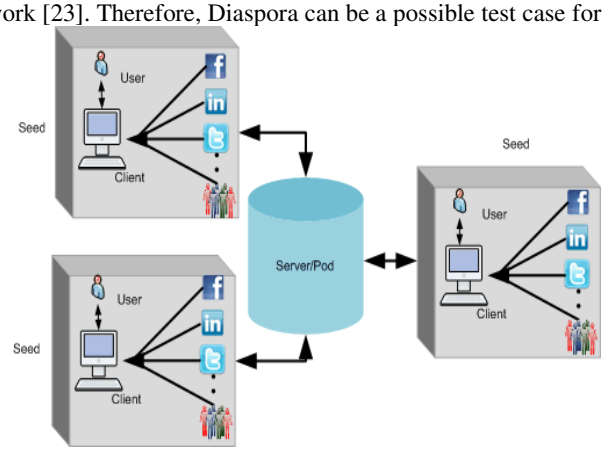

Fig. 1. Diaspora System including User, Client, Seed and Server

Assessment classification

Diaspora privacy features assessment considers based on the privacy scores. If a Diaspora system feature comply with PbD principles then it scores 1 otherwise it scores 0. For example, in Table 2, a feature "Runs on a network of connected servers" complies with PbD principles. Therefore, this feature scores 1. On the other hand, "Not used any privacy model" does not comply with PbD principles. Therefore, this feature scores 0. To make the final assessment, the privacy scores sum up in a principle. Final assessment has done based on this summation. The summation score 0- 'None' or not comply, 1- 'Low'comply, 2- 'Medium' comply, 3-'High' comply with the PbD principles.

Table 2. PbD Principles [8] and Diaspora* Alpha

#	PbD Principles	Diaspora System features	Privacy Score	Diaspora privacy features assessment
1	Proactive not Reactive; Preventative not Remedial	Runs on a network of connected servers	1	3 'High' comply
		Provides flexibility for a user to setup their own server.	1	
		Diaspora provides three different levels of security such as 'None', 'Low', and 'High' for securing the user data	1	
		Diaspora system has not produced a PIA to outline the possible future privacy impacts	0	
2	Privacy as the Default	Engaged GNUPG to ensure privacy in the system	0	2 'Medium' comply
		Not used any privacy model	0	
		Can be better considered as "Security by Default".	1	
		Considers encryption whereas possible though average user might not have any idea what is encryption	1	
3	Privacy Embedded into Design	Diaspora depends upon the third party privacy guard GNUPG instead of designing embedded own architecture	0	0 'None' comply
4	Full Functionality – Positive-Sum, not Zero-Sum	Consider philosophy of "Secure as much as you must, but no more"	1	2 'Medium' comply
		Documentation included little about how the information is utilized in a client and server	1	
5	End-to-End Lifecycle Protection	Not include how to handle end-to-end lifecycle protection	0	0 'None' comply
		Not specified the content deletion policies	0	
		Not specified the content re-distribution policies	0	
6	Visibility and Transparency	Anyone can download the Diaspora's open source code	1	3 'High' comply
		User can setup their own Social Network Third party communication with the server	1	
		User can setup their own Social Network Third party com-	1	

		munication with the client		
7	Respect for User Privacy	Diaspora system contains a model for securing private communications and data between the server, client and user.	1	2 'Medium' comply
		Claimed as trusted system but as with the other distributed system sometime 'trust' becomes more complicated	1	

Overall, Diaspora followed only few of the PbD principles. However, since the privacy aware Diaspora distributed Social Network is still in the early development stages, there are opportunities to address these in the future. Table 2 shows how Diaspora follows the PbD principles from our analysis. For some principles, Diaspora follows them partially whereas in some principles Diaspora follows the PbD principles completely. However, at this stage, the Diaspora system does not truly support full privacy as it primarily substantiates securing personal content utilizing encryption features. The Diaspora system can be better classified as following "Security by Design" principles instead of "Privacy by Design" principles.

5 Conclusion and future work

The PbD principles are more conceptual than a technique or framework. To comply with the PbD principles requires focusing on both regulatory and engineering issues [24]. Information and privacy commissioners can help to solve the regulatory and legislation issues and for the technical issues, engineers and researchers should place a strong emphasis on adopting the PbD principles in their information system design practices. However, the PbD concepts are not only limited to compliance or technical issue but also to organizational and managerial issues. Business managers also have a definite role for engaging PbD principles. They should have clear perception of engaging PbD concepts in an organization ecosystem to avoid future issues.

PbD can also be a matter of political choice [25]. Additionally, information system design with PbD principles may need to support different legislation requirements [26] such as Electronic Identification (eID) design. For eID, a balance is required between identity efficiency and protection and it is required to harmonize the understanding between regulators, engineers, business managers and politicians to achieve the ultimate success when implementing PbD concept in information systems.

PbD can be a solution for the future information system privacy and there are several challenges such as management, process and technology that may affect the issue of privacy at the design level of information systems [27]. The reluctance of management engagement, poor attitudes towards privacy and data protection, lack of appropriate privacy languages and uncertain benefits of privacy management, are all factors that impact in privacy support in online information systems.

Typically, it is hard to justify investment in privacy functionality until a severe incident occurs. An organization might use a "privacy policy" to protect themselves from the negative outcomes. The organization may also fail to plan appropriate information systems privacy support due to inadequate risk analysis as well as limited Privacy Impact Assessments (PIA) and, hence, fail to consider the value of personal information of their consumers. External pressure to share personal information with "privacy-friendly" third party can also lead to different privacy-related issues.

The PbD principles indicate that a service provider needs to increase both visibility and transparency of operations. The service provider has to be accountable for any service provided through their information system such as external links or third party services, and mismatch with any regulatory or compliance environment. As the PbD principles can have different data protection legislation requirements [26], then these barriers must be overcome to successfully utilize the PbD principles to protect user information.

This paper has given an overview and review of the increasing use of PbD principles and how those principles are being utilized in different privacy invasive areas. This paper has argued that the PbD principles are the current best instrument to designing protection for user privacy in online information systems. This research has thrown up many questions in need of further investigation for Diaspora and other open source distributed social networks. A further study will review The Distributed Friends and Relations Network [28], GNU Social [29], Lorea [30], NoseRub [31], StatusNet [32], , The Mine! Project [33] on how they address the PbD principles and which one is better support the PbD principles. Such reviews of privacy aware information systems would help us to establish a greater degree of accuracy on the PbD principles approaches and would assist the researcher to suggest or design explicit technical solutions for ensuring privacy in Social Network.

"Privacy by Design" is an emerging and important concept. The current findings add substantially to our understanding of how and why PbD principles can be utilized to guard information system development towards privacy awareness. The current study contributes additional evidence that the "Privacy by Design" concept is does matter for the design and operation of Social Networks to manage the user's privacy more effectively and transparently.

6 Reference

1. European Commission, A comprehensive approach on personal data protection in the European Union. 2010: Brussels.
2. Skinner, C.-A. Berners-Lee: Social networks are a 'threat to the web'. 2010 [cited 2011 17 February]; Available from: http://www.pcadvisor.co.uk/news/index.cfm?newsid=3249764.
3. Cavoukian, A. (2010) Landmark Resolution passed to preserve the Future of Privacy.
4. Shapiro, S.S., Privacy by design: moving from art to practice. Communications of the ACM, 2009. **Volume 53**(Issue 6): p. 27-29.
5. Langheinrich, M. Privacy by design—principles of privacy-aware ubiquitous systems. 2001: Springer.
6. Denning, D.E.R., Information warfare and security. Vol. 118. 1999: Addison-Wesley.
7. van Blarkom, G., J. Borking, and J. Olk, Handbook of Privacy and Privacy-Enhancing Technologies. Privacy Incorporated Software Agent (PISA) Consortium, The Hague, 2003.
8. Cavoukian, A., Privacy by Design … Take the Challenge. 2009, Information & Privacy Commissioner of Ontario.
9. Cavoukian, A., S. Taylor, and M.E. Abrams, Privacy by Design: essential for organizational accountability and strong business practices. Identity in the Information Society, 2010. **Volume 3, Number 2**(Privacy by Design: The Next Generation in the Evolution of Privacy): p. 1-9.
10. Williams, J.B. and J.H. Weber-Jahnke. Social networks for health care: Addressing regulatory gaps with privacy-by-design. in Privacy Security and Trust (PST), 2010 Eighth Annual International Conference on. 2010.
11. Cavoukian, A. and P.C. Spencer (2010) Ontario Health Study Assessment Centres A case study for Privacy by Design.
12. Ahern, S., et al., Over-Exposed? Privacy Patterns and Considerations in Online and Mobile Photo Sharing, in Proceedings of the SIGCHI conference on Human factors in computing systems. 2007, ACM: San Jose, California, USA. p. 357-366.
13. Cavoukian, A. and A. Stoianov, Biometric Encryption: A Positive Sum Technology that Achieves Strong Authentication, Security AND Privacy. 2007.
14. Weiss, S., Privacy threat model for data portability in social network applications. International Journal of Information Management, 2009. **29**(4): p. 249-254.
15. Vernal, M. Encrypting Facebook UIDs. 2010 [cited 2011 17 February]; Available from: http://developers.facebook.com/blog/post/419.
16. Cutler, K.-M. Facebook, Google Offer Conflicted Definitions of Data Portability. 2010 [cited 2011 17 February]; Available from: http://www.insidefacebook.com/2010/11/10/facebook-google-openness/.
17. Credentica Inc. Credentica: The U-Prove technology. 2004-2010 [cited 2011 17 February]; Available from: http://www.credentica.com/.
18. Roig, A., Privacy and Social Networks: From Data Protection to Pervasive Computing, in AAAI Spring Symposium Series. 2010, Association for the Advancement of Artificial Intelligence: Palo Alto, California.
19. Diaspora. Diaspora* Alpha. 2010 [cited 2011 17 February]; Available from: https://joindiaspora.com/.
20. The New York Times Company. Four Nerds and a Cry to Arms Against Facebook. 2011 [cited 2011 19 May]; Available from: http://www.nytimes.com/2010/05/12/nyregion/12about.html.
21. Kickstarter Inc. Decentralize the web with Diaspora. 2010; Available from: http://www.kickstarter.com/projects/196017994/diaspora-the-personally-controlled-do-it-all-distr.
22. Diaspora. Questions From Luis Villa. 2010 [cited 2011 19 May]; Available from: http://blog.joindiaspora.com/2010/04/30/a-response-to-mr-villa.html.
23. TechClump. Diaspora!!! Facebook Killer??? 2010 [cited 2011 19 May]; Available from: http://www.techclump.com/diaspora-facebook-killer/.
24. Davies, S. (2010) Why Privacy by Design is the next crucial step for privacy protection.
25. Le Métayer, D., Privacy by Design: A Matter of Choice, in Data Protection in a Profiled World, S. Gutwirth, Y. Poullet, and P. De Hert, Editors. 2010, Springer Netherlands. p. 323-334.
26. Lusoli, W. and R. Compañó, From security versus privacy to identity: an emerging concept for policy design? info, 2010. **Vol. 12**(Iss: 6): p. 80-94.
27. Information Commissioner's Office (2008) Privacy by Design.
28. DFRN Information Center. DFRN-The Distributed Friends and Relations Network. 2010 [cited 2011 17 February]; Available from: http://info.dfrn.org/.
29. Free Software Foundation Inc. GNU social. 2008-2010 [cited 2011 17 February]; Available from: http://foocorp.org/projects/social/.
30. Lorea. Lorea. 2010 [cited 2011 17 February]; Available from: http://lorea.org/.
31. NoseRub. NoseRub. 2010 [cited 2011 17 February]; Available from: http://noserub.com/.
32. StatusNet Inc. StatusNet. 2010 [cited 2011 17 February]; Available from: http://status.net/.

The Mine! Project. The Mine! Project. 2009 [cited 2011 17 February]; Available from: http://themineproject.org/about/.