GROUNDING PRIVACY-BY-DESIGN FOR INFORMATION SYSTEMS

Shan Chen, Innovation and Enterprise Research Laboratory, Centre for Quantum Computation and Intelligent Systems, University of Technology, Sydney, NSW Australia, shan.uts@gmail.com

Mary-Anne Williams, Innovation and Enterprise Research Laboratory, Centre for Quantum Computation and Intelligent Systems, University of Technology, Sydney, NSW Australia, mary-anne.williams@uts.edu.au

Abstract

The Privacy-by-Design approach has gained an increasing acceptance for privacy management in the privacy community. However, there is still a research gap in methodologies for implementing this approach and a need to develop frameworks and systems to support Privacy-by-Design practice. In an attempt to bridge this gap, this paper uncovers hidden issues of the Privacy-by-Design approach as a means to derive privacy requirements for implementing information systems with privacy embedded by design.

Keywords: Privacy, Privacy-by-Design, Privacy Requirements, Privacy Management, Privacy Trade-Off.

1 INTRODUCTION

The global economy has embedded individuals and organizations into complex social networks. To continue playing a crucial role in this economy information systems are being developed in an open and social platforms. These operational platforms present information systems with a new set of privacy challenges. Many existing systems have failed to provide robust privacy management functionality. Examples can be found from many privacy infringement cases arising from Facebook (www.facebook.com), a social networking service provider; and Google (www.google.com), a search-based service provider.

The lack of robust privacy methodologies has led to an increasing acceptance of Privacy-by-Design (Cavoukian 2010), an approach that argues building privacy into technologies as a default (Privacy Generations 2010). However, there is still a research gap in methodologies that can be used to implement this approach during the development of platforms for privacy practice in information systems, due to:

- The Privacy-by-Design approach does not provide details in terms of how each of its seven principles can be implemented – i.e., implementation requirements and conceptual grounds to be adapted - in a dynamic environment.
- The Privacy-by-Design approach does not provide clear and detailed guidelines to address fundamental privacy issues at an ontological level.

In an attempt to bridge this gap, this paper studies the seven principles of the Privacy-by-Design from the perspective of information system (IS) development – in the dimensions of *implementation requirements*, *conceptual grounds* and *IS requirements*. The findings of this study can serve as a basis to develop a robust set of privacy requirements upon which information systems can be designed, developed and deployed with privacy embedded by design.

The rest of this paper is mainly dedicated to elaborating the Privacy-by-Design principles in the three dimensions mentioned above. The paper ends with a discussion and outlook for future work.

2 PRINCIPLES OF PRIVACY-BY-DESIGN: CONCEPTUAL GROUNDING AND REQUIREMENTS

The objectives of the Privacy-by-Design are to protect privacy by embedding it into new technologies and business practices from the beginning. There are seven foundational principles of Privacy-by-Design: 1. *Proactive* not Reactive; 2. Privacy as the *Default Setting*; 3. Privacy *Embedded* into Design; 4. Fully Functionality – *Positive-Sum*, not Zero-Sum; 5. End-to-End Security – *Full life Cycle Protection*; 6. *Visibility and Transparency*; 7. *Respect for User Privacy*.

This section studies each principle in three dimensions namely Implementation Requirements, Conceptual Grounds and their associated IS Requirements. Implementation Requirements are requirements for privacy designers to embed privacy according to the Privacy-by-Design philosophy into an operational environment (like an organizational environment, an information system) in which privacy practice is required. Conceptual Grounds capture implicit concepts that underline the principle and that are identified in the Implementation Requirements. Concepts are captured and identified at an ontological level. IS Requirements elaborate requirements to fulfill each principle within an IS context.

2.1 Proactive

This principle concerns preventing privacy infractions from occurring. It sets privacy design as an action of before-the-fact, not after.

Implementation Requirements:

1. The principle requires privacy designers to understand *what is* privacy and *what is not*, in order to determine what can be qualified as *privacy infractions*. To prevent an event from occurring one needs to understand the *triggers* of the event. Under this principle, triggers of privacy infractions must be identified.
2. A privacy infraction includes a privacy loss. However, a privacy loss situation does not necessary lead to a privacy infraction case. Both situations are important to privacy management. This principle does not recognize the need to prevent privacy loss from occurring. The absence of preventing a privacy loss from occurring in this principle and the important role that privacy loss situations play in privacy management further the need to define privacy in terms of "what is" and "what is not". A clear boundary needs to be defined within context (in the scope of this paper, this refers to IS applications), when applicable. Such a definition is based on the identification of events - what is and what is not.
3. Being proactive to prevent unwanted events from occurring also implies that, in case of such an unwanted event occurring, managing actions to minimize negative privacy implications and procedures adapted for future preventions from unwanted event occurrences are needed for robust privacy management. Since the status of the information is changed after an unwanted event occurs, stakeholders' expectations and goals regarding the information can be changed - friendly reminders to the stakeholders for adaptations about their goals must be issued and concerns identified above need to be addressed.

Conceptual Grounds:

I. Privacy indicates desired status. Within the legal and sociological framework, personal privacy concerns one's desired status of information about them (Chen and Williams 2010).

II. Broadly, when information is disclosed or reaches undesired status due to infractions of permissions, obligations or regulations, privacy of the information is *infracted*. When information reaches an undesired status without breaking any agreements or obligations, privacy of the information is *lost*. For example, if one left a bank statement on a shared desk and it is seen by others, then one's privacy with respect to the financial information is lost. When the bank discloses the financial status without one's permission, his/her privacy on this financial status is infracted. Clearly, situations of "loss" or "infraction" must be referred to the information stakeholder's desired status about the information.

III. Privacy events include privacy loss and privacy infraction. Triggers of these events are identified based on the pre-conditions of each event. A subject that makes a change of a pre-condition of an event is a trigger of the event. Such a subject can be a user, an information status, or an infrastructure that supports the existence of the information. A trigger of a privacy event is qualified when it makes a change of an information status to meet a pre-condition of the event and the post-condition of the event is satisfied when the event occurred.

IS Requirements:

A. To prevent a privacy event from occurring, two basic functions are required:
    a. monitor of the event trigger generation, and
    b. notice of event triggers to privacy stakeholders.

   Associated functionalities required to fulfill secure notice delivery involved:
    a. Delivery channel

       Ways of sending notice to the stakeholders must be secure in terms of only stakeholders being able to receive the notice, to ensure relevant information not being leaked to the wrong hands or unnecessary information not being disseminated.
    b. Layered notice

Various levels of details notice to receivers constructed according to factors like timeframes a notice is sent and expected to be received and the receiver's privacy right to and goals of the information (Chen and Williams 2010).

B. Security of delivery channel can have social implications. Affecting factors include the notice, intended receivers' inter-connections and their networks (i.e., privacy stakeholders vs. non-stakeholders with respect to access and permission to resue the information) need to be managed.

2.2 Default Setting

This principle gives preconditions that provide for the protection of a person's privacy so that if an individual does not do anything, then, his/her privacy will remain intact. The principle also sets a post-condition under which a need of protecting privacy in new times - meaning no action is required on the part of the individual to protect their privacy.

Implementation Requirements:

1. This principle assumes privacy practice is undertaken in a static environment, reflecting in:
   1) the pre- and post-condition of privacy intactness of an individual, and
   2) the individual's privacy requirements will never be changed.
2. The information associated with the part of the individual's privacy under consideration is:
   1) isolated from any other information about the individual, or can be related to the individual;
   2) never used elsewhere; and
   3) never shared with others – i.e., there is only one stakeholder of the information.
3. The privacy rights of the individual whose privacy under consideration surpasses any others' rights.
4. Privacy does not concern the associated information's status with respect to the stakeholder's expectations, instead, it refers to as individual's action.

Conceptual Grounds:

I. The assumption of a static environment rejects information's social feature – which indicates that every existence of information has some social purpose (Chen and Williams 2012; Zhang and Benjamin 2007) - for the stakeholder to access the world or for others to communicate with the stakeholder (Chen and Williams 2012). Information's social feature is particular important in the current information intensive era where individuals and organizations are embedded in complex social networks. Recent emerging online social networks have dramatically increased the complexity of social structures by increasing the ease of making connections online and integrating online relationships with those in the offline world. This phenomenon rejects isolation of the information under consideration from its stakeholder's other information, as well as the stakeholder's privacy surpassing others.

II. On consideration of information's social feature, the dynamics of the information stakeholder's operational environment in which the information status can be changed, must be considered, in particularly with respect to:
   i. communications and connectivity between information stakeholders, and their own networks; and
   ii. the user's new expectations and goals at new times.

   Specifically,
   - with regard to i, above, a relationship's properties (Chen and Williams 2010, 2011) need to be managed; and
   - with regard to ii, above, the stakeholder's rights to the information about themself (Chen and Williams 2010) need to be managed.

III. The notion of privacy intactness is unclear - as per the Proactive, both privacy loss and privacy infraction situations need to be defined.

IS Requirements:

A. System design on the assumption of this principle will not be able to accommodate dynamics of the user environment. To remove this inability, the conceptual grounds identified above must be taken into consideration.

B. Accommodation of dynamics implications is a representational issue, which has been partially addressed by Chen and Williams (2011). This principle prioritizes the scalability issue in representations to stressing accommodations of extended defaults, if privacy is to be built-in with respect to a user's rights and the dynamics of the user environment.

2.3 Embedded

This principle requires embedding privacy into design and architecture of IT systems and business practice. It concerns privacy as integrality to systems and business models without discounting system functionalities or compromising business rules.

Implementation Requirements:

To assure system functionalities and business rules, relevant factors and associated information, information's role in fulfillments of these functionalities and rules, and information status of pre- and post-implementation of these functions and rules must be identified.

Conceptual Grounds:

I. The notion of information must be understood at an ontological level. Information's nature (Chen and Williams 2011) like type, presentation, amount, size, volume, granularity, as well as any alternatives, are critical to fulfill the functionality and business rules; required to be identified.

II. The capacity and tolerance of each function or rule, in terms of information flow when the function or rule is implemented. This is important to information privacy, since a satisfied output information status cannot guarantee the process of information at all times where situation that can incur a privacy loss or infraction will not occur. E.g., information might be processed to a situation in which it is at a reachable position and privacy loss can be incured. When a function or rule has such capacity, the implementation of the function needs to be redesigned to minimize the chance of incurring a privacy loss of the information. Tolerance also refers to alternatives that are relevant to information flow and status.

IS Requirements:

A. As an integrity context for business practices and IS implementations, privacy needs to be constructed at the ontological level (as per requirements of the principles of Proactive and Default Setting). Privacy constructs are modeled on the underlying concepts that are identified by the Proactive and the Default Setting, and their interrelationships; as well as functionalities required to support implementations of the principles, i.e., the requirements inherent from the Proactive and the Default Setting.

B. To fulfill system functionalities and to successfully implement business rules, representations of information to achieve privacy must not conflict with system representations to achieve system functionality and business rules.

2.4 Positive Sum

This principle requires accommodating all legitimate interests and objectives in a positive-sum to avoid unnecessary trade-offs (i.e., minimum trade-offs) - e.g., privacy and security should not discount each other, but mutually add value. This principle in particular concerns trade-offs between privacy

and system functionalities that implement business rules. Executing business rules requires processing information. This process can involve user's personal data. When the process is in digital formats, information systems are employed to provide functionality to processing information to implement business rules. Personalized services require personal data to satisfy user expectations with personalization. A user's willingness to provide data is largely based on their trust of the service provider (i.e., the business/system), where trust is built on the user's relationship to the service provider based on the user's acceptance for the service's usability (e.g., application scope, efficiency, effectiveness and robustness of functionalities) and security protection. Following this line of reasoning, we identify six privacy trade-offs namely, *business rule, system functionality, user expectation, trust*, and *relationship (in two sub-dimensions)*. The implementation requirements of this principle are derived from each trade-off category.

2.4.1 Privacy vs. Business Rule

Implementation Requirements:

To achieve a positive-sum for this trade-off, privacy must be managed at a level where business rules can be executed when required. This means satisfaction of user expectations will not create any barriers to information processes that are to execute business rules. The user shall be given knowledge about how their information is used when business rules are executed to allow them justifying their privacy expectations.

Conceptual Grounds:

Executing business rules requires an operational environment. Such an environment can involve third parties for business reasons and can demand more than necessary information to maintain the environment. If only necessary of information is required from the user is a requirement to maintain the operational environment, the fulfillment of this requirement will serve as a basic requirement to grant users' rights to manage their information.

IS Requirements:

System functionality to implement business rules, as per Section 2.4.2 (below).

2.4.2 Privacy vs. System Functionality

Implementation Requirements:

To achieve a positive-sum for this trade-off, privacy must be managed at a level where system functionality can fully function as expected. This means system constructs capturing privacy properties will not create any barriers to information processes that are to achieve the functionality. The user shall be given knowledge about how their information is used when functionalities are performed to allow them justifying their privacy requirements.

Conceptual Grounds:

I. Processing information to achieve functionality means sufficient information is required. With privacy concerns, this means the information collected must not beyond what is necessary to fulfill associated business rules.

II. Sufficiency and necessity of information concern information nature, storage and use. Information nature refers to properties of type, structure, amount, volume, size, granularity. Storage refers to properties of device type, location, period of retention and device maintenance. Use refers to access to and operations on the information. Necessity of information in these dimensions concerns information status with respect to the information stakeholder's privacy.

III. Affecting factors include permission options available for grant and uses of the information. Necessity of information applied to these factors requires minimum permissions available for

grant to minimum number of actors - this means the stakeholder of the information will have necessary choice to manage the information about them.

IS Requirements:

A. Information collected to fulfill functionality must satisfy a Minimum Principle, which requires minimum information (with respect to its nature) and its collection, use, retention, permissions and number of actors granted permissions. Minimum Principle enforces maximum obligations onto actors with respect to information use and retention.

B. A Functionality Instruction detailing information processed by each function including inputs and outputs, with respect to the information nature, storage and retention must be accessible to the user.

C. Users are provided with sufficient choice options to implement the Minimum Principle with respect to their privacy expectations.

2.4.3 User's Expectation: On Information vs. On Personalization

Implementation Requirements:

To obtain a positive-sum from this trade-off, user's expectations about the personal information, personalized service to receive, and the interplay between these two dimensions need to be identified and defined. Consistency between different types of expectations needs to be maintained at all times. Conflict detection methodologies, as well as notification and suggestions deliver to the user upon a conflict is detected are required.

Conceptual Grounds:

I. User's expectation on personal information can conflict with his/her expectations about personalized service to receive, due to lack of awareness with respect to how the information will be utilized for configuring a service with obligations tailored to him/her. The trade-off in this regard is understood as "purpose of the information and expectation of information status" vs. "desires to use the information".

II. Information status concerns the information nature and accessible status by others.

IS Requirements:

A. Users are to be given options to specify their expectations on the information status. Approaches developed by Chen and Williams (2010, 2011) can be used for developing options with regard to the information nature and accessible status, respectively.

B. Alerts are to be delivered when there are conflicts with respect to users' expectations in this regard.

C. The use of a third-party domain is highly uncertain to users. Many social networks have established "comprehensive" privacy policies. For example, Facebook (2013) "may integrate third party features …to provide you with better services…To learn more about the information they collect or receive, review their privacy policies." Plink (2012) states "Our Web site includes Social Media Features, such as the Facebook Like button. …Social Media Features are either hosted by a third party or hosted directly on our Site. Your interactions with these Features are governed by the privacy policy of the company providing it." From these examples it can be seen that social networks tend to direct their responsibilities to the user's "self-regulation" on the ignorance of their connections to the third-parties in distributing users' data. When considering a user's privacy protection, measuring the use of third-party domains is crucial; however, it has not been well considered and established in the research community. Therefore, in the case that a third-party is involved, monitoring business rules with third party polices with respect to the user's information under consideration and their expectations on the information must be functioned, and alerts to the user must be delivered, when appropriate.

D. Collection of a user's data requirements:
    a. only collect for a well-defined purpose,
    b. only collect relevant data for the purpose defined,
    c. only keep data as long as it is necessary for the purpose, and
    d. always use real time data collection and storage, when possible.

2.4.4 Privacy vs. Trust

Implementation Requirements:

Trust indicates a relationship that contributes to the connection between two parties. Requirements in this cluster is two-fold:

1. For trust between the user and service providers

   Privacy is managed by sharing information that is selected based on the user's trust towards the service provider. This trust relationship largely relies on the user's awareness about how the service provider will use his/her information. Transparency and visibility about user data usage at the service provider's side are the basic requirements in this dimension.

2. For trust between users

   Privacy is managed by the strength of the connection with respect to information sharing and exchanging. Transparency and visibility between users are important to develop trust relationships, however, at a cost of privacy of the users involved. Robust functionality to help users to gain required level of trust with minimal privacy loss with respect to user's goals on the information under consideration and ensure a privacy loss will not lead to privacy infractions is required. Negotiation functionality can advance the implementation.

Conceptual Grounds:

I. Without privacy one has little confidence to trust the service received. A trust relationship, from agent A towards agent B, will in turn impact agent A's privacy decision about his/her information sharing and exchanging with agent B. The trust relationship in this aspect (i.e., a user's trust on using the service), largely relies on the user's knowledge about the service provider with respect to the way they make use of the information, the consequences of the information use (when applicable); and other users' (if any) ability enabled by the functionality, to use his/her information.

II. Often, third parties introduce uncertainties to users in terms of information retention and making use of the information. Users' trust relationship to the service provider can be influenced by the way the service provider handing third party's business rules and polices with respect to use of user information including storage, process and retention. Transparency/visibility and robustness are the keys to gain users' trust and increase the trust level. Transparency/visibility means well-established and well-defined polices and business rules, in terms of comprehensive, expressiveness and easy to understand and ways to access. Robustness means commitment to provide empowering user-friendly options and features – typically, choice, consent and control options for users to decide, declare and control who can do what to them (Chen and Williams 2010).

IS Requirements:

A. Presentation of polices and business rules are comprehensive, easy to understand and user friendly.

B. Empowering options and features are provided to one user without discounting another's privacy.

C. Third party polices and rules are monitored against business rules and user's privacy expectations.

2.4.5 Relationship: Simplex vs. Multiplex

Implementation Requirements:

Requirements in this cluster are to maintain an operational environment to manage separations that are elaborated in the Conceptual Grounds below.

Conceptual Grounds:

I. Privacy is necessary to be understood in a social context. "Without society there would be no need for privacy" (Moore 1984). One needs privacy, to create a personal space in the society. Privacy is needed when one connects to some others. As socially created needs (Moore 1984), privacy enables separation and separations enable different relationships to be established (Rachels 1985) – without privacy we will not have the necessary dignity to build relationships with others in the society (Chen and Williams 2010).

Under the separation theory (Rachels 1985), one can share more than one separation with another; thus, one can have more than one relationship connecting to another – i.e., one's relationship to another can be multiplex, in terms of reasons of the connection. Examples that can be found from everyday life are: be friends, of the same family/community, work for the same organization, and join the same group, etc. Such reasons, are often termed as "relationship type" (Chen and Williams 2010, 2011).

II. Privacy concerns over a multiplex relationship are two-fold:
    i. one's preference to maintain each type; and
    ii. the interrelationships between all existing types.

With regard to i, concerns include "weight" of each type – e.g., each type is of equal importance, one or more certain types are more important than the rest.

With regard to ii, concerns include overlap, disjoint and dependency – e.g., each separation is disjoint from all other separations and thus relationships exist in the separations are disjoint; two or more separations are overlapping or have dependent relations to some extent and relationships in these separations are inherently overlapping or dependent.

III. Trade-offs among the existence of the relationship types involve trust, security, and usability. Privacy trade-offs for a multiplex relationship is two-layered:
    i. On each type (in each separation), trade-offs are the sum of trust, security and usability trade-offs.
    ii. On the relationship integrity, trade-offs are the sum of the trade-off on all types, on the basis of minimum costs to the prioritized type(s).

To illustrate, consider person *A* and person *B* are on a multiplex relationship: After reading B's article about the North Pole, A contacted B to share some of his experience traveling to Finland. They then become friends. Sometime later they decided to co-author a book about the North Pole. As their relationship and interest about the subject developed, they partnering a consultation company for North Pole escape. Decomposing this relationship, we can see:

i. *A* and *B* were friends
ii. *A* was a reader of *B*'s North Pole article
iii. *A* and *B* were co-authors of a book called "7 Days at the North Pole"
iv. *A* and *B* were partners of the consultation company "The North Pole Escape"

The relationship between *A* and *B* was "friend, author-reader, co-author, partner". At this current stage, *A* considered "friend" as the most important relationship connecting him to *B*, and prioritized their partnership over the other two relationships (i.e., reader-author, co-author); whereas *B* considered the partnership and the co-authorship as an extension to friendship that he prioritized.

With regard to friendship, *A* and *B* both agreed to respect to each other's privacy unless one was keen to share with another. *A* offered some of his pictures to the book; however, due to the copyright issue, they could not use these pictures to promote their company. *A* did not want to let *B*

know one of his trips to Finland with his ex-girlfriend *C*. However, *B*'s friend *D*, recommended *C*'s photos taken in the North Pole to him to be used as advertisements. When A was aware of *D*'s recommendation, *A* gave up some privacy on his story about this trip and shared some information about his former relationship to *C* with *B*, to persuaded *B* not to consider *C*'s offer. *A*'s decision was made on his understanding about *B*'s perception on partnership and friendship as equal importance, and created a trade-off between his privacy to *B* and *B*'s trust on their relationship. For *B*, this trust was for the whole relationship; for *A*, this trust was to increase the security of the partnership and to strengthen the friendship.

This scenario represents a potential relationship that can exist in systems aiming for socialization support - in fact, in many existing social systems like social network systems there are built-in functions supporting various degrees of multiplex relationships.

IS Requirements:

IS requirements in this cluster is mainly a representational problem:

A. Relationship representation must be able to accommodate the multiplex property to allow users to hold multiplex relationships:

relationship(holder_1, holder_2, set_of_types_in_existence)

B. References need to be accommodated to allow users to exercise trade-offs at two layers:

relation(type_1, type_2, ref)

where, *ref* is a reference for the holders, i.e., *ref(holder_1, holder_2)*.

2.4.6 Relationship: Symmetry vs. Asymmetry

Implementation Requirements:

Requirements in this cluster are to maintain an operational environment to manage separations and directions that are elaborated in the Conceptual Grounds below.

Conceptual Grounds:

I. Relationship can be symmetric and asymmetric. Symmetric relationship connects two entities on one type, regardless directionality. Asymmetric relationship connects two entities on direction dominant types (Chen and Williams 2010, 2011). Relationships like sibling, colleague, classmate, housemate are symmetric, indicating entity A and entity B connecting to each other for the same reason. E.g., A and B are siblings – "A is B's sibling" and "B is A's sibling" are both true. Relationship like friend, parent-child, sister-brother, employer-employee are asymmetric, indicating entity A and entity B connecting to each other on different reasons. E.g., A and B are on a parent-child relationship – if "A is B's parent" is true, then "B is A's parent" cannot be true. "C is D's friend" does not imply "D is C's friend".

II. Asymmetry relationships can be described as symmetric relationships, when certain details can be omitted. Consider a brother-sister relationship: "Tom and Mary are brothers and sisters", i.e., Tom is Mary's brother, and Mary is Tom's sister. This asymmetric relationship can be described as "Mary and Tom are siblings", i.e., "Tom is Mary's sibling" and "Mary is Tom's sibling" are both true. Similarly, a team leader and his team member can be "team leader - team member" as well as "colleagues".

III. Privacy trade-offs with respect to socially created needs reflect in social connections and associated relationships. E.g., C's relationship to D is "friend", i.e., C connects to D as a friend; where D's relationship to C is acquaintance. Privacy concerns over this relationship include:

i. C's privacy to D and C's privacy about his relationship to D
ii. D's privacy to C and D's privacy about her relationship to D

E.g.,

- C's and D's privacy to each other:

As a friend C shares all his contact details with D, while D sees C as an acquaintance and only wants to be in touched via her general-purpose email account.

- C's and D's privacy about their relationship to the other:
  C is keen to let his personal contacts know that D is his friend, while D does not have motivation to introduce C to her contact circle.

Privacy implications from these concerns include:

- If C and D share common friends, then D will not be able to keep his privacy about their relationship - more precisely, the connections between them - to their common friends.
- Regardless issues of common friends, awareness of the other one's privacy to the self will likely have a great impact on one's privacy decision to the other. If C knew D would treat him as an acquaintance only, he might not be willing to share all his contact details with D, but selected information.

These implications concern uncertainty and informativeness, i.e., a trade-off between uncertainty and informativeness on relationships can be reduced to a trade-off between the relationship's symmetric and asymmetric properties. In other words, when one is uncertain about the other's relationship towards him/her, privacy decision justified at a less informative and symmetric level can reduce uncertainties about privacy implications. Consider C and D's case, above, if C is uncertain about D's relationship to him, a decision made on a more certain level – e.g., acquaintance - can help to decrease uncertainty.

IS requirements:

IS requirements in this cluster is mainly a representational problem:

A. Representation of relationship between two parties needs to be expressed in an asymmetric form when required:
   relationship(holder_from, holder_to, set_of_types_in_existence)
   type_in_existence (type, duration, constraint)

"Duration" is an important property of a relationship type that indicates the reason the relationship exists. With respect to privacy, information shared prior to this duration can affect the information privacy within the duration; in particular for information that can be used to infer privacy associated with the relationship and its stakeholder, duration is important to identifying potential issues for privacy decision made and the value of this property can be used to design solutions to avoid privacy loss. For example, Mary and Tom are both identifiable to Phoebe, via their relationship to Company A - i.e., they are all colleagues working for the same company. However, Phoebe does not know Mary has a brother, and Tom has a sister. This means Mary and Tom's relationship is unknown and not identifiable to Phoebe. Later, from time to time, Phoebe learns: a) Mary went to UniB 2002-2005, and Tom 2005-2008; b) Mary was the only one mentor of subjectA in 2008; c) Tom has a mentor for subjectA in his last year in the university; and d) Tom's mentor was his sister. After learning a-d, Phoebe concludes Mary and Tom are sisters and brothers. If this conclusion is to be avoided, the mentor-mentee relationship can be made unavailable or the duration can be made vague.

To design service allowing user interactions, engagement and transaction, this property is an important affecting factor to user's privacy. The relationship between two entities can be user-to-user (e.g., social networking between users), user-to-object (e.g., user purchasing a product), or user-to-service (e.g., user using a service – e.g., "People you may know" (PYMK) (Chen and Williams 2011). Allowing users to manage this property value's availability to other can reduce chances of privacy loss to the associated relationship, and therefore its stakeholders.

Asymmetry and symmetry relationships can be "hidden" through use of this property. E.g., if Mary did not want Phoebe to know that she was Tom's mentor, she can either not disclose the mentor relationship or utilize the duration (e.g., make it imprecise – e.g., "awhile ago" or "previously". This extends the asymmetry to a wider range. To give another example, if Mary

tells Phoebe that Tom is her friend, and Tom tells Phoebe that Mary is his lifetime friend, then their relationship is asymmetric to Phoebe. This symmetric (recognition between Mary and Tom) to asymmetric (made available to Phoebe) can preserve Mary some privacy on the relationship, and prevent future identification chances.

B. The relationship-holder should be able to exercise trade-off uncertainty vs. informativeness via transfers between symmetric and asymmetric.

Type ((asymmetric_type, asymmetric_constraint),
(symmetric_type, symmetric_constraint),
trigger))
relation(asymmetric_type, symmetric_type, ref)

where, *ref* is a reference for the holders, i.e., *ref(holder_1, holder_2)*

- *constraint* holds requirements for the type to be activated.
- *trigger* holds the conditions to trigger the transfer to occur. Such a trigger holds both directions' transfer – i.e., from asymmetry to symmetry, and vice versa.
- *relation* holds the connections between asymmetric types and symmetric types, e.g., hierarchical structure, and mapping mechanism for locating the appropriate type to transfer to are required.

2.5 Full Lifecycle Protection

This principle concerns privacy with respect to lifecycle management of information. It begins with the first element of information being collected, throughout the entire lifecycle of all information involved, and securely destroys all data at the end of the process.

Implementation Requirements:

This principle sets privacy grounds in data collection, use and retention. To respect privacy as a personal goal in relation to the status of information about oneself, these grounds are rooted at a purpose setting for the information under consideration in the dimensions of data collection, use and retention. To implement this principle, methodologies to manage purpose of information in these three dimensions are required.

Conceptual Grounds:

In a dynamic social environment, one's privacy requirements can evolve. Accordingly, one's purpose setting for his/her information can require to be adjusted. Lifecycle management of information purpose necessary be integrated into the dynamics environment. Termination of information's lifecycle means the information no longer exists in all forms and will not be able to be recreated in any form.

IS Requirements:

A. Accommodation of purposes of user information.
B. Allow users to develop purpose for their personal information to manage their collection, use and retention against their privacy expectations.
C. Purpose development includes management of purpose evolvements.

2.6 Visibility and Transparency

This principle concerns all stakeholders' obligations and objectives for both business practice and technology involved.

Implementation Requirements:

1. Polices and procedures of the service provider and third-parties involved in the information process, as well as responsible contacts must be visible and accessible to all stakeholders.
2. Requirements are inherited from the trade-off concluded in the Conceptual Grounds below.

Conceptual Grounds:

Visibility and transparency place a weight on trust between stakeholders and their rights to manage privacy. Trust in this dimension inherits those in the same dimension of the trade-off "Privacy vs. Trust" (Section 2.4.4) under the Positive Sum.

IS Requirements:

A. Functionality to view and search polices and procedures of the service provider and third-parties, as well as responsible contacts need to be implemented.
B. As per IS Requirements of "Privacy vs. Trust" (Section 2.4.4)

2.7 Respect for User Privacy

This principle concerns user-centricity above all and requires empowering user-friendly options. Within the legal and sociological framework, respect for users in relation to privacy means allowing personal goals to be achieved and facilitating an operational environment to exercise personal information to achieve the goals. Such Implementation Requirements are the sum of all the Implementation Requirements under the other six principles described above. Accordingly, the Conceptual Grounds and IS Requirements of this principle receive a sum of those are in the same dimensions under the other six principles.

3 CONCLUSION AND OUTLOOK

This paper seeks to advance an understanding of the implementation issues of the Privacy-by-Design approach within an information system's context. By uncovering the grounding issues of the seven principles, privacy requirements for information system's design are derived at a conceptual level.

Each principle is studied in three dimensions for uncovering hidden issues and concepts, as a path to derive privacy requirements for implementing information systems with privacy embedded by design. These three dimensions are: Implementation Requirements for privacy designers to design privacy; Conceptual Grounds for uncovering key concepts as a basis for developing building blocks for systems; and IS Requirements for IS designers to design privacy embedding into information systems.

Privacy naturally creates many trade-offs. Therefore, the Positive Sum is elaborated with a priority in this study. This priority aligns with the Privacy-by-Design philosophy that sets the Positive Sum as the ultimate goal (Cavoukian 2010).

Through this study we have found the seven principles overlap each other to some extent, due to the privacy nature or system design and implementation needs. For example, the Visibility and Transparency is embedded into other principles – e.g., the Respect for User Privacy highlights users' rights to privacy requires Visibility and Transparency support; the Positive Sum of privacy and trust can be better achieved via implementing Visibility and Transparency. Another obvious example is the Respect for User Privacy sums other principles; as a result it overlaps with all other six principles. To some extent this finding of the interconnections between the seven principles contradicts with Cavoukian's (2010) proposition. We argue that this is due to the Privacy-by-Design framework lacking a consideration of the fundamental conceptual grounds required to support requirement analysis. Our study has initiated a path to requirement analysis for implementing Privacy-by-Design systems.

Future work will be based on the findings of this study to develop a formal and robust set of privacy requirements and building blocks to provide a grounding support for privacy management in information systems that can be deployed in an open and social environment.

Acknowledgements

The authors thank the anonymous reviewers for their time and constructive comments. The authors thank Prof. Robert Davison for his proofreading and comments on the camera-ready version.

References

Cavoukian, A. 2010. www.privacybydesign.ca/publications.htm. (Accessed on 19 Jan 2013).

Chen, S. and Williams, M.-A. 2010. Privacy: An Ontological Problem. In *PACIS 2010 Proceedings.* Paper 134.

Chen, S. and Williams, M.-A. 2011. Grounding Data Purpose and Data Usage for Better Privacy Requirements: An Information System Perspective. In *PACIS 2011 Proceedings.* Paper 43.

Chen, S. and Williams, M.-A. 2012. Information Makes a Difference for Privacy Design. In *PACIS 2011 Proceedings.* Paper 178.

Facebook. 2013. How do third parties use cookies, pixel tags ("pixels") and other similar technologies on Facebook? http://www.facebook.com/help/159967110798373/. (Accessed on 19 Jan 2013)

Moore, B. 1984. Privacy: Studies in Social and Cultural History. M.E. Sharpe ; Distributed by Pantheon Books.

Plink. 2012. Plink Privacy and Security Policy. https://www.plink.com/index.cfm?fuseaction=main.privacy. (Accessed on 19 Jan 2013)

Privacy Generations. 2010. http://www.justice.gov.il/PrivacyGenerations/. (Accessed on 28 April 2013)

Rachels, J. 1985. Why privacy is important. In *Ethical Issues in the Use of Computers* Wadsworth Publ. Co., Belmont, CA, 194-201.

Zhang, P. and Benjamin**,** R.I. 2007. Understanding Information Related Fields: A Conceptual Framework. *Journal Of The American Society For Information Science And Technology*, 58(13):1934–1947, Wiley Periodicals, Inc.