Association for Information Systems
AIS Electronic Library (AISeL)

PACIS 2011 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

9 July 2011

Grounding Data Purpose And Data Usage For Better Privacy Requirements Development: An Information System Perspective

Shan Chen
University of Technology, shan.chen@uts.edu.au

Mary-Anne Williams
University of Technology, Mary-Anne.Williams@uts.edu.au

ISBN: [978-1-86435-644-1]; Full paper

Recommended Citation
Chen, Shan and Williams, Mary-Anne, "Grounding Data Purpose And Data Usage For Better Privacy Requirements Development: An Information System Perspective" (2011). *PACIS 2011 Proceedings*. 43.
http://aisel.aisnet.org/pacis2011/43

This material is brought to you by the Pacific Asia Conference on Information Systems (PACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in PACIS 2011 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

GROUNDING DATA PURPOSE AND DATA USAGE FOR BETTER PRIVACY REQUIREMENTS: AN INFORMATION SYSTEM PERSPECTIVE

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

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

Data purpose is a central concept to modeling privacy requirements for information systems. Existing purpose-centric approaches for privacy protection have mainly focused on access control. The problem of ensuring the consistency between data purpose and data usage has been under-addressed. Given the lack of practical purpose-centric solutions, we argue that a grounded understanding of the underlying concepts of data purpose and usage is fundamental to modeling privacy requirements. In recognition of an existing "privacy rights" framework, this paper develops an ontological grounding of data purpose and usage that can be used to understand their implications on fundamental privacy rights for modeling privacy requirements for information systems.

Keywords: Privacy, Privacy Rights, Privacy Requirements, Data Purpose, Data Usage.

1 INTRODUCTION

In today's service dominated global economy privacy is valuable. The rapid development of online services has presented a trend in developing the Internet into a network of services to facilitate and optimize our everyday life and business, e.g., email service, chat service, employment service, banking service, travel service, social networking service, etc. Social networking websites are one of the fastest growing online service providers in recent years. These websites provide a platform to facilitate communication, exchanging and sharing of information among users and third parities. Supported by the Internet infrastructure, socialization on social networking websites is free of location and time limitations. To facilitate socialization and gain reputation in the competitive market, these networking websites offer innovative services to users. However, while these services provide attractive features they also introduce high-level privacy infringement risks to their users due to the richness of the information exchanged and shared, and the high degree of user interconnectivity.

The problem of potential privacy infringements has two key aspects:

1. User's awareness: the user lacks awareness of information use control. In other words, risk of privacy infringements can occur due to the user not being made aware to monitor and control their personal information.
2. Service functionality: the user lacks ability of information use control because the service platform does not provide adequate functionalities for fulfilling privacy requirements of information use. In other words, risk of privacy infringements can occur due to the user not being able to monitor and control their information.

From an operational perspective, awareness and functionality reflect the problem of "who can do what" – i.e., i) the user's awareness of "who can do what to information about me"; and ii) functionality of the service that allows users to monitor and control "who can do what" to them.

Given that privacy is personal-dependent, privacy status of the information under consideration is justified by the user self. In other words, the rationality of a privacy infringement claim needs to be justified based on the user's requirements. From an information system's perspective, the operational problem is a system requirement problem, i.e., the service platform provides functionality to allow users to manipulate information. To develop functionalities to satisfy users' desires, it is essential to understand users' intensions about using their information – a problem referred to as "data purpose", which leads to "use limitation" of the data to achieve privacy protection of the data – two privacy protection principles identified by the Organization for Economic Cooperation and Development (OECD) (OECD 2009), namely Purpose Specification Principle (PSP) and Use Limitation Principle (ULP).

A number of purpose-based approaches for privacy protection have been proposed (Byun et al. 2005; Staden and Olivier 2007). However, there is still insufficient support for users to control their personal information – which has been demonstrated in practical contexts, typically, by privacy breaches continually reported in mainstream media. For example, Moses (2009) and France-Presse (2007). This phenomenon highlights the need to review and enrich the semantics of *data purpose* as one of the fundamental privacy requirements. In information systems, the ability to accommodate rich semantics for effectively controlling information use requires effective representations of data. Motivate by such needs, this paper studies two relevant OECD Privacy Principles – i.e., the PSP and the ULP – with a focus on data purpose specification for implementing mechanism for users to control data usage in information systems. We conduct the study within an existing framework built on three fundamental privacy rights namely *choice*, *consent* and *control* (Williams 2009; Chen and Williams 2010a, 2010b).

In the next section we present a motivating example, identify the problem and propose a path to the solution. Section 3 interprets the meeting point of the principles and the fundamental rights at focus.

Section 4 studies basic properties of the principle. Section 5 grounds privacy requirements based on the findings. Finally, Section 6 discusses findings and identifies future work.

2 PROBLEM DOMAIN

Social referral, in which one connects to another by a third party referral, is an important social channel to bridge social entities. In social networking systems (SNSs), such referral is implemented using recommendation techniques. We refer to this type of referrals as social recommendations (SRs).

The problem domain of this paper is scoped in SRs, due to: i) social networks provide a rich social infrastructure for information sharing; and ii) SRs in these networks are rich resource for investigating information privacy issues to derive privacy requirements for information systems' development. This section presents a SR scenario, analyzes privacy issues arose from the scenario and proposes a path to the solution.

2.1 A Social Recommendation Example

Mary received a "People You May Know" (PYMK) list from her social network on MySN site, allowing her to send a message to people on the list. She was surprised to find many school friends she had lost contact with on the list. However, she also felt compromised because she saw those who share the same type of profession as her were on the list. She was disappointed not being able to keep her professional information disjoint from her personal social network and to keep her away from those professionals she did not want to network with, because she believed if she saw others' information they could also see hers. Her concern was confirmed by a *friending* request (i.e., requests to be added as a friend) from her manager the next day. Mary joined another social network on RealSN which allowed her to choose who could know her existence in the network (i.e., she could choose who she would not be recommended to) and she believed that she could now stay away from her manager. However, she did not know one of her best RealSN friends, Phoebe, was her manager's daughter who shared online experience with her father. As a result Mary received a *friending* request from her manager on RealSN.

2.2 The Problem

Mary lost her privacy on the PYMK list because she did not have chances to choose:

1. whether to be listed on the PYMK list presented to *who* - e.g., she wanted to be listed and recommended to those who share the same interests with her except her manager;
2. *who* can or cannot see her on the list if she wanted to be listed - e.g., Phoebe can see her but her manager cannot; and
3. *what* those who can see her on the list can *do* about her - e.g., Phoebe cannot pass her information onto her manager without her permission.

These problems constitute a problem of "who can do what to information about me" (WCDW) – a problem of fundamental rights, i.e., Mary did not have sufficient rights to do what she expects. Relevant works (Williams 2009; Chen and Williams 2010a, 2010b) have modelled these rights into three categories namely *choice*, *consent* and *control*. Fundamental to the privacy problem is identified as the right to choose the kind of consent to enact control: WCDW (Chen and Williams 2010b). In other words, one should be granted sufficient rights in choosing the "what" and the "who" to consent such that one can control own information. This "right" model is referred to as 3CR model (Chen and Williams 2010b).

Within the 3CR model, "consent" is a central right. One's right to grant permissions for others to use their information dominates the information's privacy status. Without permissions information cannot be used. In other words, information privacy control is upon permissions. The personal-dependency property of privacy justifies the appropriateness of permissions on user's intensions. Therefore, an

understanding of data purpose is crucial for permission decisions. On the other hand, construction of appropriate permissions requires sufficient and relevant options – i.e., the right to choose what to consent as the ability to construct permissions. Both MySN and RealSN fail to provide functionality for Mary to exercise her rights of choice, consent and control.

In summary, the privacy problem of collaborative information systems is the requirements of three fundamental rights (i.e., the 3CR) w.r.t. the problem of WCDW and "purpose" as the central concept of requirements.

2.3 A Path to Solution: Privacy Requirements as The First Step

The first step to implement information systems is to develop system requirements. To realise privacy protection in information systems privacy requirements need to be developed. Methodologies and guidelines for privacy requirements development are required. The set of principles for privacy protection identified by the OECD is a good candidate because they represent as far as possible a global consensus. The PSP and the ULP are relevant principles; however, they do not provide detailed guidelines for requirements development to address the operational problem namely "who can do what to me" (WCDW). In an attempt to bridge this gap, in this paper we study the semantics of these two principles, and enrich them to interpret the 3CR model.

3 INTEGRATING PSP&ULP INTO 3CR: THE MEETING POINT

As defined by the OECD, the PSP restricts use of the collected data to the purposes of collection, and the ULP enforces data to be used for specific purposes for which consent has been given. How can these two principles be semantically accommodated in the 3CR model? In other words, how can the PSP and the ULP interpret three fundamental rights as privacy requirements for information systems in which users can exercise these rights to preserve their privacy? Semantically,

- The PSP concerns the *consistency* of the purpose of data usage and the purpose for which they were collected. This concern of consistency captures the notion of choice and consent, allowing the principle to be interpreted as users' rights *from choice to consent.*
- The ULP concerns the *consistency* between data usage and the purpose of usage. This concern of consistency captures the notion of consent and control, allowing the principle to be interpreted as user's rights *from consent to control.*

This integration can be demonstrated in the context of social recommendations, where both principles play a dominating role. Typically,

- The PSP restricts relationships because information can only be collected for the purpose of sending social recommendations. This raises important questions regarding the nature of consenting to recommendations. For example, in consenting to a recommendation does the user also consent to spin-off recommendations that might include business propositions like purchasing a product.
- The ULP enforces a social relationship can only exist for the purpose of social interactions. For example, a religious relationship is for religious interactions and not for trading interactions.

By this interpretation, two principles intersect at the central *right* "consent". A comprehensive interpretation of three *right*s therefore requires binding two principles coherently to centrally position *consent* in the 3CR model, such that *control* of the operational issue – the WCDW - can be enacted.

4 INTERPRETING PROPERTIES OF PRINCIPLES

Unless properties that constitute the backbone of the principle are understood, coherent binding between principles cannot be achieved. This section studies the key concepts that conceptualize (i.e.,

capture status of) constitutional properties of the principles, namely *data purpose* for the PSP and *data usage* for the ULP.

4.1 Data Purpose

By restricting the use of the collected data to the purposes of collection, the PSP holds properties of *data* and *purpose*. Data has status. Data purpose shows expectations for specific status. It implies permissions for actions on the data object to maintain the expected status. A data object can have multiple purposes, coexist or optional, to indicate expectation of its status. We use the term "validity" to bridge expected and unexpected status in response to satisfaction of data purpose. We say a data status is valid when it meets a set of purposes expectation of the same data object. On this notion of *validity*, a data status is valid only when its status satisfies the expectation of the purpose outcome. Such expectations are "safeguarded" by permissions for actions on the data object. Thus, purpose implies permissions.

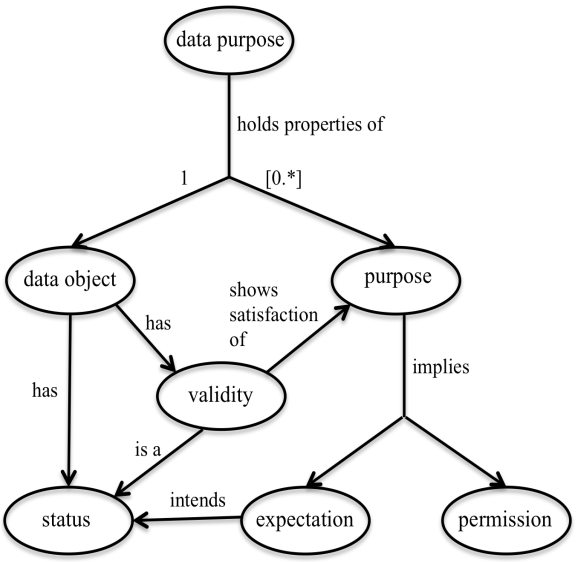

Figure 1. Grounding of PSP

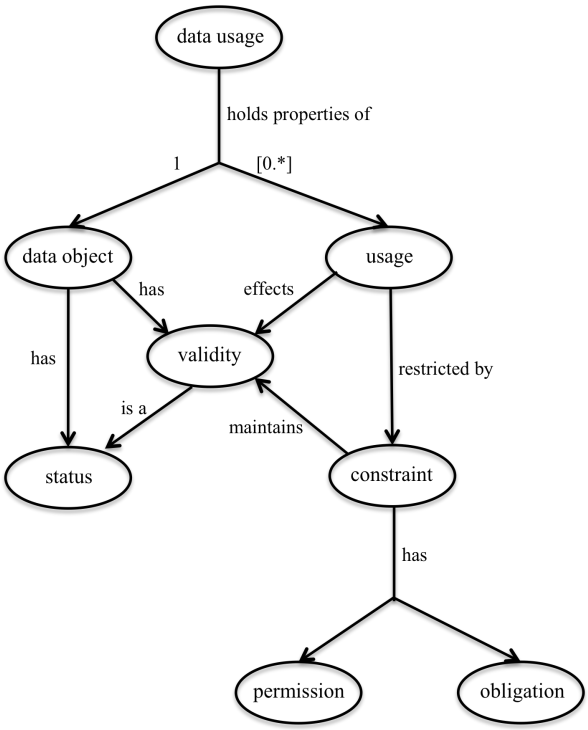

Figure 2. Grounding of ULP

A data object can have a one-to-zero, one-to-one, or one-to-many relation to purpose. One-to-zero means there is no any purpose bound onto the data object. One-to-one refers to a single purpose is expected for the data object. One-to-many signifies more than one purpose is expected for the data object, where these purposes can be optional or coexist to indicate an expectation. Required by the purpose to serve expectations to reach expected status, permissions are multiple. An expectation requires relevant permissions to "safeguard" conditions for an expected achievement. E.g., Mary and Phoebe maintain a connection for communication about travel related matters. Upon agreement, they maintain their relationship for teaming-up i) a culture-exploring trip in China before 2010; ii) an adventure in South Africa before 2010; iii) a driving trip in Tasmania after 2011; and iv) a backpacking trip in Europe before 2010 or after 2011. They intend to build four separated teams and consent to different channels different information.

In summary, the PSP has properties namely data and purpose, where the former concerns validity and status, the later concerns expectations and permissions.

4.2 Data Usage

By enforcing data to be used in compliance with specific purpose, the ULP holds the properties of *data* and *usage*. As mentioned, data has status. Data usage can be restricted by certain constraints, which can be permissions under which use of data can be undertaken. Constraints can also include obligations indicated by commitments or responsibilities associated with permission grants.

Usage restricted by constraints can be interpreted by "purpose of use", i.e., the purpose of the data usage. Usage effects data validity through change of status.

In summary, the ULP has properties namely data and usage, where data is the object concerned by the PSP and usage concerns constraints and purpose.

5 GROUNDING PRIVACY REQUIREMENTS

We argue privacy requirements for collaborative information systems are threefold: i) data purpose and data usage binding, ii) 3CR accommodation, and iii) implication consideration. In this section we present a grounding of requirements from these three aspects (Sections 5.1-5.3). Then, based on the findings we propose a set of Codes of Practice in Section 5.4.

5.1 Data Purpose and Data Usage Binding

Purpose of using a data object (referred to as "use purpose") can be derived from the *usage* property (Figure 2.) of ULP, however, this property does not "self-contain" a power to enforce the required *data purpose* (Figure 1.) and ensure a consistency between data purpose and use purpose. Thus, a binding between PSP and ULP, semantically, is essential to achieve the consistency.

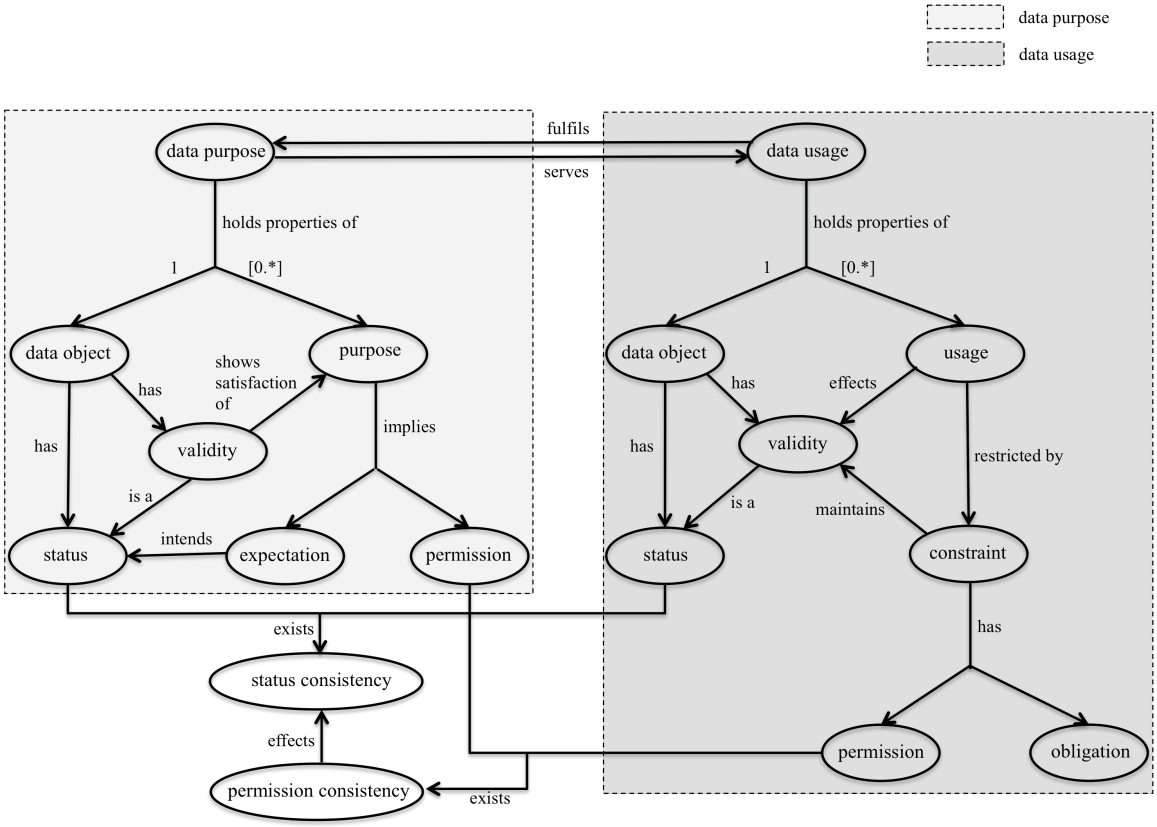

Figure 3. Binding Data Purpose and Data Usage

Figure 1. shows that *purpose* aims at the *status*, whereas it is a "product" of *usage* in Figure 2. A consistency between the *status* in Figures 1&2 can be seen as an indication of a successful binding between two principles. The status, in Figure 1, is "safeguarded" by *permission* of purpose; while in Figure 2 it is by constraints which include permissions and obligations (if any). Thus, there is also a consistency problem between permissions associated with purpose and permissions under which usage can be carried out. Figure 3. depicts the consistency problem between data purpose and data usage. A coherent binding between the PSP and the ULP needs to address the consistency problem in Figure 3.

5.2 3CR Accommodation

Three factors namely *expressivity*, *extensibility* and *adaptability* have been identified as a fundamental problem to accommodate the 3CR in information systems. In addition, rights of consent and control also concern factors of *fidelity* and *adequacy*, respectively (Chen and Williams 2010a). We refer to these factors as the *3CR representation criteria*. How the PSP and the ULP can be interpreted such that detailed guidelines can be established to develop comprehensive 3CR requirements for better privacy management? In Section 3 we identified the relation between the principles and the 3CR namely *PSP as from choice to consent* and *ULP as from consent to control*. In the following we look at fundamental dimensions of each property identified in Section 4 and reason about a set of fine-grained guidelines for exercising the 3CR with the PSP and the ULP.

The PSP on Choice&Consent

The PSP interprets *choice* and *consent* by two core concepts, namely *data object* and *data purpose*, to provide users sufficient and meaningful choice options for constructing their consent.

- Data Object

  The nature of a data object is described by size, volume, amount, granularity and structure. Of these elements, the structure plays a dominant role in having choice to set consent on the object's size, volume, amount and granularity. Specifically, this concerns

  - Having choice to construct consent for any components of a composite data object – a reflection on the *adaptability* criterion.
  - Having choice to construct consent on the level of detail in hierarchies, when hierarchical structures are involved – a reflection on the *expressivity*, *extensibility* and *adaptability* criteria. Scenarios concerned include:
    - The data object or a component of the data object is a hierarchical object (i.e., a hierarchy of objects is considered as a whole for the data object or for the component). E.g., relationship information as a data object, e.g., there is already a set of hierarchical relationships and the relationship of interest can be located in the hierarchy, or the relationship is established at a higher abstraction level of agreements, e.g., trading.
    - The data object can be linked or mapped to a hierarchical object (assume mapping mechanism is available). E.g., a location related concept "Ultimo_Australia_2007" can be mapped to a set of location concepts at different levels of detail - i.e., "Ultimo->Sydney->NSW->Australia".

- Data Purpose

  Data purpose aims at data objects. With its backbone at *permissions*, data purpose concerns fulfilment of expectation to maintain the data object at a desired status. Specifically, this concerns

  - Having choice to provide consent to any of the conditions that remains the validity of a data object – a reflection on the *expressivity* and *adaptability* criteria.

- Having choice to construct complex composite purposes with different conditions on a data object – a reflection on the *expressivity* and *adaptability* criteria. E.g., the validity of a data object D at time t_1, t_2, and t_3 can be in the scenarios below:
  - at time t_1, D can only be used for purpose p_1, and at time t_2 used for p_1 and p_2 and not p_3 where p_1 includes p_3;
  - at time t_1, D can only be used for either purpose p_1 or purpose p_2 and not for both together, and at time t_2 must be used for p_2 and p_3 where p_1 includes p_3; or
  - D can only be used for both purpose p_1 and purpose p_2, or after time t_1 used for either purpose p_1 or purpose p_2 and not for both together.
- Having choice to construct purpose at different abstractions.

E.g.,

Data object: The relationship between Mary and Phoebe.

Purpose: Mary and Phoebe to maintain a relationship for communication on travel related topics, with four specific purposes:
i). a culture-exploring trip in China before 2010
ii). an adventure in South Africa before 2010
iii). a driving trip in Tasmania after 2011
iv). a backpacking trip in Europe before 2010 or after 2011

Expectation: i). communication for travel relationship development.
ii). both Mary and Phoebe's professional space and non-specific-trip-related personal space will not be touched by their communications. "Touched" means any information about the subject of interest is affected.
iii). four teams will be built at different period of times.

Permission: i). communications must be related to a specific trip.
ii). relationship can be used for a specific trip referral.

Figure 4. An Example

The ULP on Consent&Control

The ULP interprets *consent* and *control* by two core concepts, namely *usage* and *constraint*, to provide users sufficient power to control personal information based on the consent they provided – i.e., use constraints to restrict usages. As mentioned,

- *Consent* concerns the criterion of *fidelity* – on which that is satisfied, data purpose can be derived by data usage and permissions implied by the derived purpose will be consistent with permissions under which usage can be carried out. In other words, the criterion of *fidelity* requires usage to be valid *only* under the permissions associated with purpose for the same data object.
- *Control* concerns the criterion of *adequacy* – on which that is satisfied by reflections on *constraints* that require permissions and obligations. Permissions may not be adequate to restrict usages to achieve expected results. For example, communications/referrals under permissions in Figure 4. do not guarantee a result that a professional space is untouched. Phoebe shared information about their trip to China with her father, asking for his advice because he lived in China during 2007 and 2008. Although she mentioned Mary only when she talked about the trip, the way she described Mary "told" her father that Mary was his employee. If Phoebe committed to the obligation in Figure 4., she might have avoided an identifiable description about Mary when she talked to her father.

  Adequacy is difficult to measure. Even if Phoebe was aware of her obligation, how could she know what to avoid in order not to "touch" Mary's professional space - if she did not know sufficient information about Mary's profession (which Mary considered as her privacy)? We attempt an implication-driven understanding to approach this problem in the next sub-section.

5.3 Implications

Positioning *data purpose* in the centre, we learn the groundings of data purpose associated with data usage from a semantic perspective on a conceptual implication-based flow. Figure 5. graphically describes the flow of implications starting from the central concept "purpose". Semantically, a data purpose is intended for a data object, which has a status and stakeholders who established the purpose. There are two types of purposes for a data object: purpose for creating a data object and purpose for maintaining the data object's life upon its creation, namely *creation purpose* and *existence purpose*, respectively.

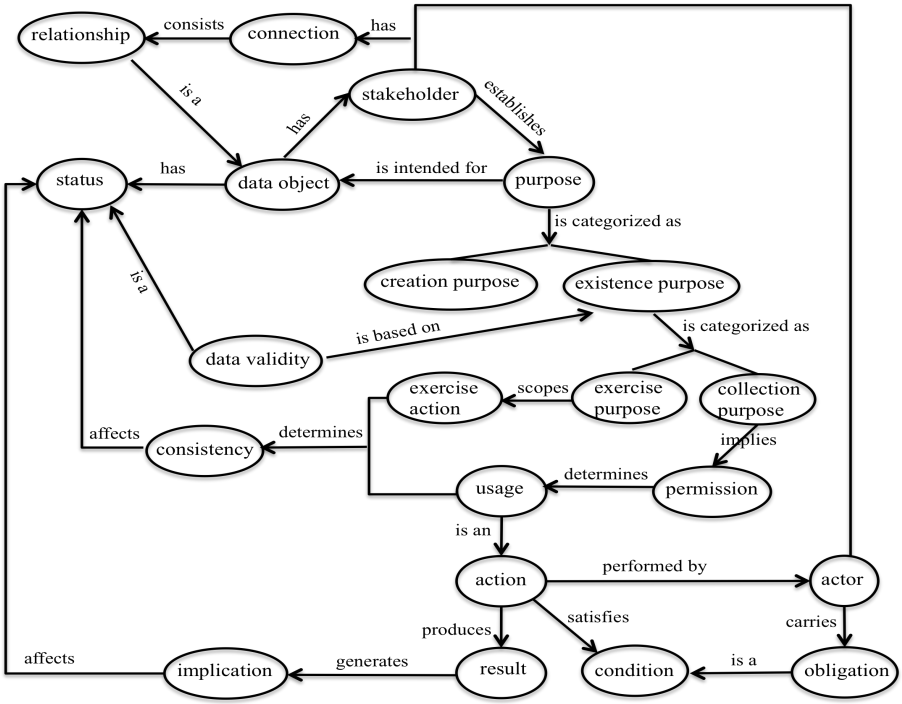

Figure 5. Ontological Grounding of Data Purpose and Usage

Data objects exist for use – to distinguish uses of data for purpose other than existence, we refer to this type of uses as actions of exercising data. Collection is generally required before an exercise action can be carried out. Therefore, the existence purpose can be divided into *exercise purpose* and *collection purpose* categories. A satisfaction of collection purpose is a prerequisite for satisfaction of exercise purpose. However, it cannot guarantee a satisfaction of exercise purpose from the perspective of a collection. A data object allows to be collected means, to a certain extent, it is allowed to be used. In other words, a data object having collection purpose attached implies it can be used under permissions associated with the purpose. Permissions determine the extent of usages, which are actions that require actors to carry obligations (if any) and will produce outputs upon completion. Usage determined from permissions implied by collection purpose may not be consistent with the exercise action intended by the exercise purpose, if the collection purpose was not clearly associated with the exercise purpose.

Categories

- *Permission consistency implication*
  Since action results can have implications on the data status that existed before the action was carried out, implications can be arose from the consistency between i) the permission of usage that can be derived from the purpose of collection, and ii) the permission of exercise that could be created without or incorrectly associating with the purpose of collection. Since usage can result in

new data status, inconsistency between them can hinder an achievement of an expected data status. We refer to implications due in this cluster as *permission consistency implication*.

- *Connection implication*
  Implications can be introduced by permissions due to the connection between actors of the usage and stakeholders of the data object. Since connections between entities can consist of complex relationships, the network of actors and stakeholders has an impact on the data status via action performance. We refer to implications due in this cluster as *connection implication*.

Dominations and Constitutions

Central to the connection implication is the problem of relationship, which can also affect permissions' effect because actor's position in the network has potential to introduce implications through information propagation, positively or negatively, from an information usage perspective. This reasoning uncovers "relationship" as a hidden property to the PSP and the ULP. It also indicates relationship as an important component to constitute permissions for binding two principles to approach the *adequacy* problem in *control*. Based on this understanding, we can constitute permission as:

<*usage*, *actor*, *effectiveTimePeriod*>

Actors as social entities are difficult to be defined by identities due to their public availabilities (i.e., their public openness). They can, however, be reached conceptually by their connections to the stakeholder.

One connects to another directly or indirectly. The connection between an actor and a stakeholder indicates the connectivity between them. Constitutions of the connectivity therefore can be utilized for data purpose decision and realization such that negative implication can be minimized. Grounding the notion of *relationship by connectivity* (RBC) (Chen and Williams 2010b), we learn:

- *Connection end*: A connection end is one of the two connected entities under consideration.
  E.g., in the Figure 5., the actor and the stakeholder are two connection ends.
- *Connecting entity*: A connecting entity is one of the entities between two connection ends are referred to as *connecting entities.*
- *Connection degree* (Chen and Williams 2010b): A connection degree indicates the distance between a connecting entity and a connection end, reflecting in the number of entities (including the connecting entity under consideration) away from the connection end.
  E.g., if *A* connects to *B*, and *B* connects to *C*, then *A* is said to be 2 degrees away from *C*, i.e., the connection degree between *A* and *C* is 2.

In practice, multiple paths connecting two entities often exist. In such cases, the connection degree of two entities is the length of the shortest path between them, because it reflects an entity's ability to connect to another economically. However, when considering privacy, the length of a connection path is not a dominating factor in assessing a relationship; rather, the types of relationships involved in the path have more impact on one's privacy status. Such complexity can be described by RBC - in the above example, if *A* and *B* are colleagues, *B* and *C* are friends, and *A* and *C* do not know each other, then the RBC from *A* to *C* is described as:

- a path: "a work connection at degree 1 followed by a friend connection at degree 2"; or,
- a structured format: {("work",1),("friend",2)}.

In Mary and Phoebe's scenario, the notion of RBC can be used to better construct permissions and obligations to achieve better information privacy preservation. For example, Mary can give permission to those who do not have connection to her professional space – e.g., using "not {("work",1)}" to indicate permissions not giving to those belong to this RBC cluster. It might be ambiguous if Phoebe does not know anything about Mary's professional information. Mary therefore needs to give away some privacy – however, she does not need to tell Phoebe all the details. She can require Phoebe to

commit to " 'work' is restricted to medicine-related areas only". By this specification, Phoebe knows that she cannot tell her father about Mary's work related information since he works for a pharmaceutical company.

On the other hand, social entities involved in a relationship can be at different levels of abstraction. For example, a relationship can be individual-to-individual, individual-to-group where the individual can belong to the group, or group-to-group where one group can belong to another group or they can overlap or be completely disjoint. In describing a relationship to her company as "medicine-related", Mary is able to avoid giving unnecessary information away in order to construct her consents. This shows a way to flexibly specify level of details of a relationship in the RBC.

In summary, implications are due to permission consistency and connection's connectivity, which are "chained up" by relationship complexity. These implications highlight our interpretation of "data purpose" supports the central privacy problem: "who can do what to information about me?" As can be seen from Figure 5, the actor ("who") performs the action/usage ("do") to the data object ("what") of the stakeholder. Having sufficient choice, consent and control rights to express desires ("purpose") and restrict personal ("stakeholder's") information ("data object") used by others ("actor") is the path to better information privacy management.

5.4 Codes of Practice

Based on the findings above, we define the following Codes of Practice (CoP) as a set of guidelines for purpose-centric requirements for developing 3CR information systems.

Data Object (DO)

[DO:COMPOSITION]
For a decomposable data object, choice options for each component and mechanisms for constructing consents to the chosen options must be provided.

[DO:ABSTRACTION]
For a data object with hierarchical structure involved, or can be linked or mapped[1] to some nodes of at least one hierarchy, choice options for each node of the hierarchy and mechanisms for constructing consents to the chosen options must be provided.

Data Purpose (DP)

[DP:EXISTENCE]
The existence purpose, if any, for the data object must be specified with the type of purpose (i.e., exercise or collection). Choice options and mechanisms to specify the purpose must be provided.

[DP: EXISTENCE_ COMPONENT]
For a decomposable data object, existence purpose for each component, if any, must be specified. Choice options to specify the purposes must be provided.

[DP: EXISTENCE_ PERMISSION]
For a data object that exists for some purpose, the permissions expected by the purpose must be specified.

[DP:VALIDITY]
For a data object that exists for some purpose, the validity of the data object must be specified against the existence purpose. Choice options for expectations and permissions, and mechanisms for constructing consents to the chosen options must be provided.

[1] This requires mapping mechanisms. However, details about mapping techniques are not in the scope of this paper.

[DP:COMPOSITION]
For a data object that exists for multiple purposes, choice options to specify coexisting purposes and optional purpose, and mechanisms for constructing consents to the chosen options must be provided.

[DP:ABSTRACTION]
For any data purpose with hierarchical structure involved, or can be linked to or mapped to some nodes of at least one hierarchy, choice options for each node of the hierarchy and mechanisms for constructing consents to the chosen options must be provided.

Data Usage (DU)

[DU:PURPOSE]
The purpose for which a data usage intends for the data object must be specified.

[DU:PERMISSION]
The permissions under which a data usage can be carried out must be specified.

[DU:ACTOR]
The type of actor must be specified by their *relationship by connectivity* to the stakeholder of the data object.

Consistency Check (CC)

[CC:PURPOSE]
The codes of [DP:EXISTENCE] and [DU:PURPOSE] of the same data object must be consistent on content and format.

[CC:PERMISSION]
The codes of [DP:PERMISSION] and [DU:PERMISSION] of the same data object must be consistent on content and format.

6 CONCLUSIONS

Information privacy concerns focus on "who can do what to information about me" (WCDW) (Chen and Williams 2010b). Within the existing legal and sociological framework, fundamental privacy rights have been identified as choice, consent and control (Williams 2009; Chen and Williams 2010a, 2010b) and a 3CR model (Chen and Williams 2010a) has thus been proposed as a basic framework for building privacy-friendly information systems.

Data purpose is a central concept to model privacy requirements. How data purpose can be modelled to utilize data usage such that the user can manage their WCDW requirements? Given that existing purpose-centric approaches lack satisfaction for user's privacy rights, this paper studies the concepts of "data purpose" and "data usage" that are central to two relevant OECD Privacy Principles namely Purpose Specification Principle and Use Limitation Principle, aiming for privacy requirements for information systems in which users have the ability to exercise their the fundamental privacy rights within the 3CR model.

This grounding study ontologically interprets data purpose and data usage from an information system's perspective. It not only analyses privacy requirements, but also addresses their representation problems – i.e., the 3CR representation criteria (Section 5.2). As a result, a set of Codes of Practice is proposed as a set of general guidelines for further development of detailed guidelines in domain-specific application areas.

This research has initiated an ontological approach to grounding the fundamental privacy problem for building robust collaborative information systems (CISs). It is ontologically promising; however, still at its infant stage. We aim for a comprehensive grounding study of the privacy problem to gain a fundamental understanding to build an operational environment for privacy management in CISs. In

this paper we focus on the static aspects of data purpose. Future work will include a deep study of dynamic aspects of the proposed approach and formalism of requirements towards robust CISs.

References

Byun, J.-W., Bertino, E. and Li, N. (2005). Purpose Based Access Control of Complex Data for Privacy Protection. SACMAT05, pp 102-110.

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

Chen, S. and Williams, M.-A. (2009). Privacy in social networks: A comparative study. In *PACIS 2009 Proceedings.* Paper 81.

Chen, S., and Williams, M.-A. (2010a). Towards a Comprehensive Requirements Architecture for Privacy-Aware Social Recommender Rystems. In Link, S., and Ghose, A., eds., *Proceedings of the Seventh Asia-Pacific Conference on Conceptual Modelling (APCCM 2010)*, 33–41. Australian Computer Society Inc.

Chen, S., and Williams, M.-A. (2010b). Privacy: An Ontological Problem. In *PACIS 2009 Proceedings.* Paper 134.

France-Presse, A. 2007. "Home trashed in myspace party." Retrieved 08 April, 2009, from http://www.news.com.au/story/0,23599,21549624-2,00.html

Moses, A. 2009. "Social not-working: Facebook snitches cost jobs." Retrieved 08 April, 2009, from http://www.smh.com.au/articles/2009/04/08/1238869963400.html

OECD. (2009). OECD Guidelines on The Protection of Privacy and Transborder Flows of Personal Data. Retrieved 08 April, 2009, from http://www.oecd.org/document/18/0,3343,en 2649 34255 1815186 1 1 1 1,00.html

Staden, W. and Olivier, M.S. (2007). Using Purpose Lattices to Facilitate Customisation of Privacy Agreements. TrustBus, LNCS 4657, 201-209.

Williams, M.-A. (2009). Privacy Management, the Law & Business Strategies: A Case for Privacy Driven Design. In *Proceedings of the 2009 International Conference on Computational Science and Engineering.* pp.60-67.