Reusable Knowledge for Achieving Privacy: A Canadian Health Information Technologies Perspective

Ilca Webster, Vera Ivanova and Luiz Marcio Cysneiros
Department of Mathematics and Statistics - Information Technology Program
York University
lwebster, vivanova, cysneiro @mathstat.yorku.ca

Abstract

Privacy is a fundamental aspect when dealing with Personal Information. Privacy requirements are those that capture privacy goals and its associated measures for a system under development. In order to ensure privacy we must identify these elements. However, there are many challenges in their identification. For example, privacy requirements may be difficult to quantify and precisely specify. There is a need for systematic approaches for reasoning, modeling and analyzing privacy from the early stages of the software development. Furthermore, it is necessary to develop a usable ontology or classification of measurable aspects of privacy that can be used to aid in the specification of privacy requirements. These ontologies should be represented in a way that facilitates their use as guidelines for the requirements elicitation process. This work builds on a review of privacy legislation to develop a catalog of aspects of privacy that can be considered during requirements gathering. This catalogue is used to guide the requirements engineer through alternatives for achieving privacy. The approach uses the i framework to model privacy as a special type of goal. We show how privacy can be modelled through different viewpoints with different alternatives for its operationalization. An example in the health care domain is used to illustrate our work.*

1. Introduction

Personal data privacy problems in health-related information systems grow in complexity as new powerful and sophisticated technologies emerge. The wide use of virtual private networks, clustered network storage, web services oriented data centers, and intelligent patient monitoring devices challenges the privacy of health-related information during the different phases of the health information systems evolution. These problems are usually dealt with in the design phase [1], during implementation, and in regular use. They are also significant during every substantial system modification, particularly those generated by mandatory privacy protection assessments [2] or by the introduction of new legislative acts [3].

A portion of the cost of health information systems must, therefore, be spent on privacy protection. Neglecting or not properly addressing non-functional requirements (NFRs), such as privacy, can increase costs and compromise the success of the systems [4], [5] [6] [7], [8]. Dealing with privacy requirements in early stages of software development facilitates future design decisions and leads to the selection of appropriate mechanisms for providing adequate privacy.

Privacy requirements can be considered as requirements that capture the privacy goals and associated measures for a system under development. Privacy goals may include a range of system aspects, especially those related to confidentiality and availability of information. Like other NFRs, privacy can be rarely said to be satisfied by software. It may be said that we frequently satisfy privacy within acceptable limits. Simon [9] coined the term "satisfice" referring to these situations. In this sense, we must refine privacy goals into sub-goals until all the necessary actions and information are represented at the leaves levels of the graphs. These actions and information are called operationalizations.

The complexity of relationships among humans leads to different viewpoints on what notions of privacy apply [10]. Besides, privacy cannot be viewed as an isolated NFR. Security, for example, plays a very important role when implementing mechanisms to support privacy. These aspects make it difficult for a software engineer to easily reason about different ways to achieve privacy. Current studies can be classified into those presenting mechanisms to address privacy, such as [11] or those helping to identify privacy requirements, such as [12]. These types of studies do not show alternatives that can impact other NFRs.

This work presents a reusable knowledge base showing possible alternatives to operationalize privacy requirements, indicating some of them that could impact other NFRs. The knowledge base is expressed in the form of a catalogue to store privacy requirements for health-related information. It was built using the i* framework [13] and is based on The Personal Information Protection and Electronic Documents Act (PIPEDA) [14] and on The Personal Health Information Protection Act (PHIPA) [3]. The catalogue is used to produce a systematic approach to guide requirements engineers through alternatives for achieving privacy. Each alternative is modelled to allow its own comparison with other alternatives and evaluations of its impacts on other NFRs, facilitating the work of requirements engineers. Note that although based in Canadians standards we believe the majority of these standards could also be applied to many other domains and mostly in the health care domain.

This paper is organized as follows: section 2 is an introduction to the privacy protection of personal health information in Ontario and in Canada, section 3 presents a privacy catalogue, section 4 shows the case study we carried out to evaluate the use of the reusable knowledge base and section 5 concludes this work.

2. The need for and importance of privacy protection in health information systems in Ontario and Canada

The legislatures of Canada and Ontario are leaders in terms of enacted legislation for protecting privacy in information systems in general, and in health information systems in particular. The Personal Information Protection and Electronic Documents Act (PIPEDA), effective since January 2004 [14], covers all commercial activities; for health care systems, it addresses their inter-provincial and international aspects. The Personal Health Information Protection Act (PHIPA), effective since November 2004 [3], was enacted specifically for health care systems in Ontario.

Various researchers have demonstrated that Canadian citizens value the privacy of their Personal Health Information (PHI) and are concerned with its protection in the current systems of e-services and document delivery. They are often afraid to authorize transmission of personal information over the Internet. However the nature of PHI exchange requires many transfers from a doctor's office to other specialists, to medical labs, to pharmacies, to hospitals, to insurance providers, etc. In an emergency, there can be a need for immediate PHI disclosure. PHI is also needed for new drugs and treatments, and research and development activities.

Objectively, the privacy protection risks include, but are not limited to, data transfer monitoring, user profile and password matching, identity theft, sale of personal data for profit, unauthorized access, use of PHI in illegal donor activities, etc. These are, in brief, the most significant reasons for enacting of both PIPEDA and PHIPA. The PIPEDA and PHIPA acts codify the privacy protection legislation requirements. However, the implementation of these requirements in a complex information system is far from trivial and will affect its entire lifecycle.

Figure 1 provides an intentional description of the social structure subjected to PIPEDA and PHIPA acts in terms of dependency relationships among them. This model is a base to explore broader implications of privacy requirements for health-related information. It shows intentional dependency relationships among strategic actors and their rationales.

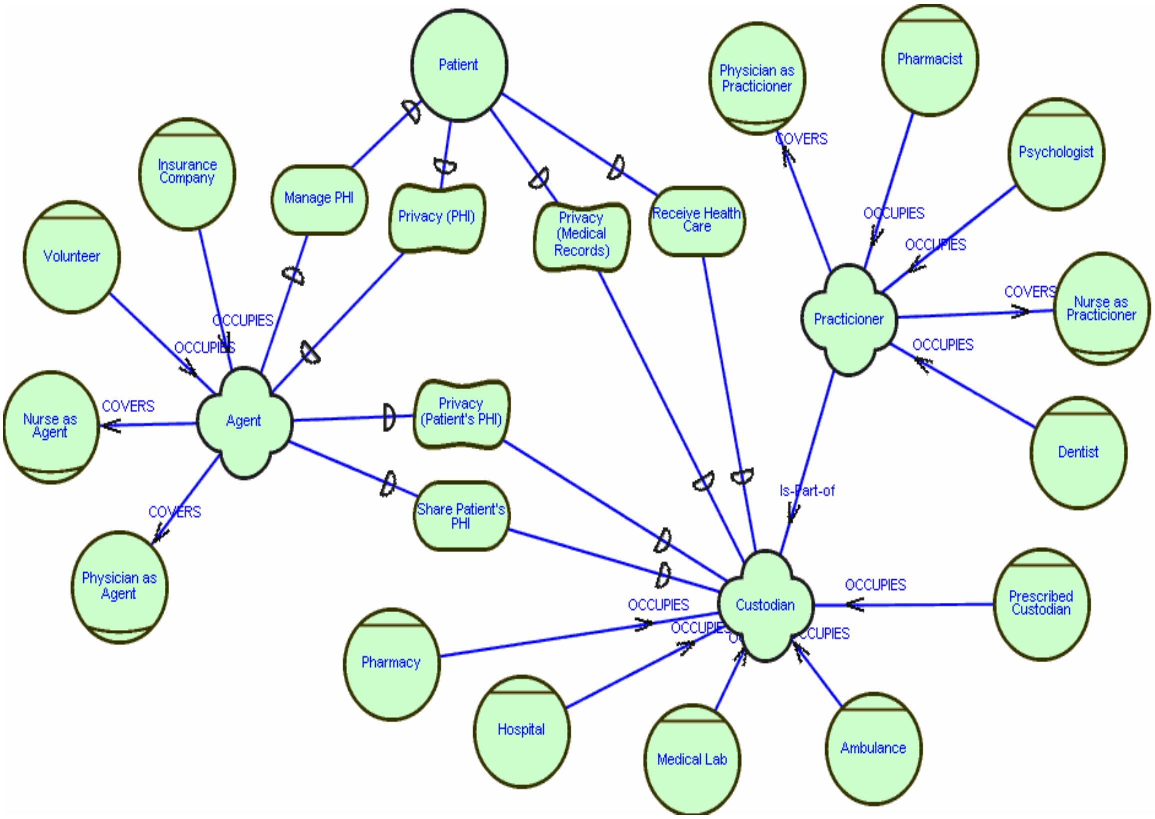

Figure 1 – Dependency Relationships

Actors can be roles, positions and agents. In the context of the i* framework, actors refer to generic entities that have intentionality. To reflect different degrees of concreteness of agency, the concepts of roles, positions and agents are defined as specializations of actors [15]. Actors may be abstract (roles defining responsibilities), concrete (agents – human and non-human individuals or classes with specific capabilities), or other organizational constructs (e.g., positions which package a number of roles together to be assigned to a single concrete agent) [16]. Custodian, Agent and Practitioner represent positions occupied by different agents and covered by different roles. Custodian, Patient, Agent and Practitioner depend on each other to have specific goals satisficed. These goals can be hardgoals or softgoals. For example "receive health care" is a Patient's hardgoal. NFRs are modeled as softgoals to be satisfied from the viewpoint of the various stakeholders. A softgoal is a particular type of goal used in i* to model quality attributes for which there are no a priori clear-cut criteria for satisfaction. Instead, actors may judge that these attributes are sufficiently met ("satisficed") on a case-by-case basis.

Among the actors mentioned above, the Custodian represents the most important role in privacy protection practices. Custodians are responsible for handling, providing, and establishing safe guards for PHI, as well as for obtaining and ensuring the grant of permission to manage patients PHI. Figure 2 depicts the Custodian Privacy Protection Practices.

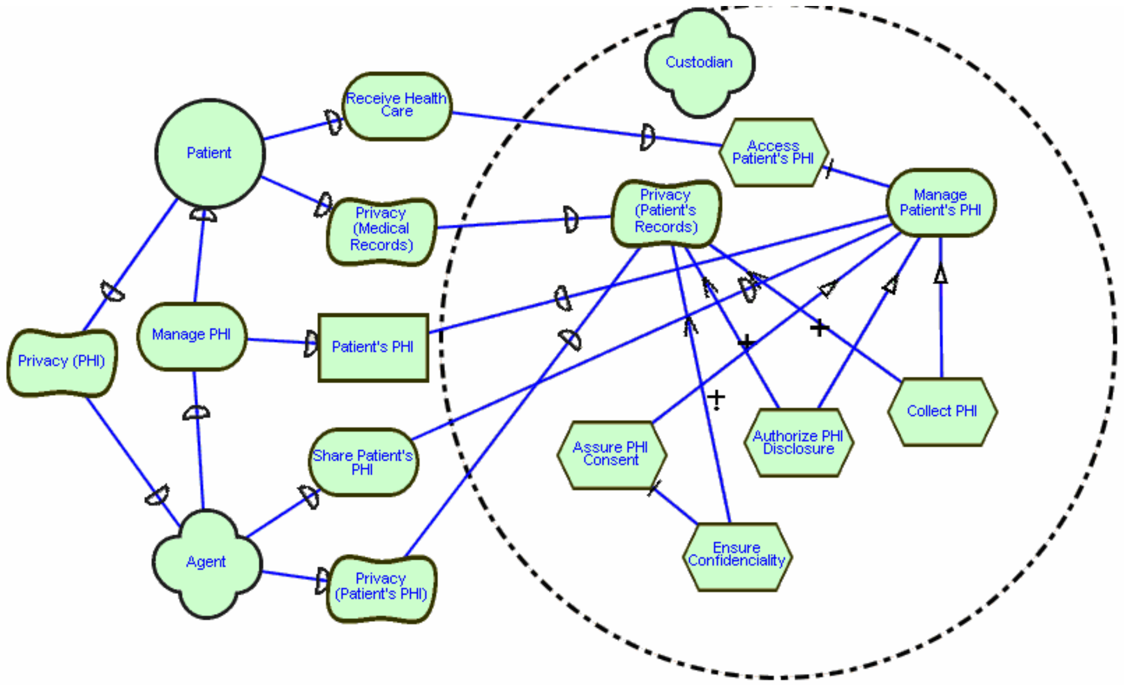

Figure 2 – Custodian Privacy Protection Practices.

3. A Privacy Catalogue

There are a number of studies presenting mechanisms for enforcing privacy [11] [12] and helping identifying privacy requirements [17]. In this work we will be addressing aspects of privacy related to PHI. A catalogue is used to capture knowledge about achieving privacy in many different situations. The knowledge comes from the Personal Information Protection and Electronic Documents Act (PIPEDA) [14] and on the Personal Health Information Protection Act (PHIPA) [3]. Having this catalogue available, we can reuse its knowledge as well as add new knowledge to it.

In this catalogue, privacy is interpreted by refining it into subgoals and subsubgoals and eventually linking them to implementable mechanisms. Various subgoals and mechanisms may contribute to privacy in varying degrees. Each stakeholders' interpretation of privacy may lead to different goal refinements and mechanisms. The various interpretations of privacy can be collected and organized into a catalogue for reference during requirements elicitation, analysis and design.

The knowledge in this catalogue is represented in a primarily hierarchical structure that facilitates representing the organization's knowledge from the highest level goals to achieve privacy. The catalogue allows representing different ways of achieving a goal. This facilitates choosing the one best suited to the problem being analyzed. Besides the operationalizations for privacy, possible correlations to other, maybe conflicting requirements are presented. This allows showing that one specific solution might achieve privacy and contribute positively or negatively to other requirements, for example security.

The catalogue was built using ***i**** [13] constructs including: softgoals, goals, tasks, and beliefs. The softgoal concept is used in ***i**** to express non-functional requirements. NFRs frequently interact with each other in complex ways. Qualitative reasoning can be carried out using contribution links among softgoals. The semantics of the links are based on the satisficing concept [9]. The most common contribution types are Help/Hurt (positive/negative but not sufficient to meet the parental goal), Some+/Some- (positive/negative of unknown degree), whereas Make/Brake indicates positive/negative of sufficient degree. Although these distinctions are coarse grained, they are enough to help us decide whether we need further refinement and search for more specific softgoals and operationalizations or not. Contribution links enable NFRs decompositions up to a point where the operationalizations for a NFR have been reached (i.e., the goals are no longer "soft"). Operationalizations can be viewed as functional requirements that have arisen from the need to meet NFRs. This can explain why people frequently face doubts about a requirement being functional or non-functional.

Operationalizations are typically specified as tasks, each indicating a particular way of doing something. All the subcomponents of a task (refined using the task decomposition link (†) must be carried out. If there is more than one way to accomplish something, then the state of affairs to be achieved is represented as a goal with means-end links (↑) linking to the alternatives.

Contribution links are the core of design decisions. By reasoning about how different operationalizations would contribute to satisfice a softgoal, it is possible to decide the best alternative to pursue. Based on the semantics of the contribution links [13], decision values are propagated from an offspring to its parents and allow visualizing impacts from adopting one alternative over another. A prototype tool (OME3 Tool [18]) has been developed to support automatically propagation of contributions and allow designer interventions in case of conflicts or undecided situations.

Figure 3 gives an idea of how this catalogue was built. The general softgoal privacy was decomposed into main concepts used to address privacy requirements: Accountability, Consent, Limiting Collection of Personal Information, Limiting Use, Disclosure and Retention of Personal Information, Accuracy, Safeguards, Openness, Individual Access and Challenging Compliance.

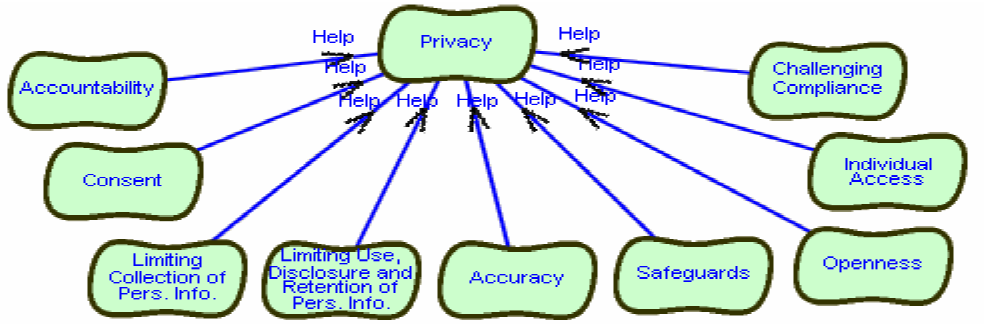

Figure 3 – Initial Approach to Privacy Catalogue.

We further decomposed each of these subgoals into more detailed softgoals. Openness, for example, is refined into Informing Users about Privacy Policies and further refined in Have Clear Policies. Accuracy is refined into Assure Correcteness of Personal Information, Assure Completeness of Personal Information and Assure Currentness of Personal Information. We kept refining each softgoal and moving towards the most concrete and possible mechanism that would operationalize each of them, as explained above. Unfortunately, the resulting graph is too large to fit in here. We will be presenting here selected parts of the catalogue. The entire catalogue can be found in [19]. Note that this catalogue should be considered as being partial. Although it is a comprehensive set of existing knowledge on Privacy Requirements we understand it is not complete. In fact, we encourage contributions from personal experiences or further knowledge that could have been left aside. Figure 4 shows operationalizations for the softgoal Openness.

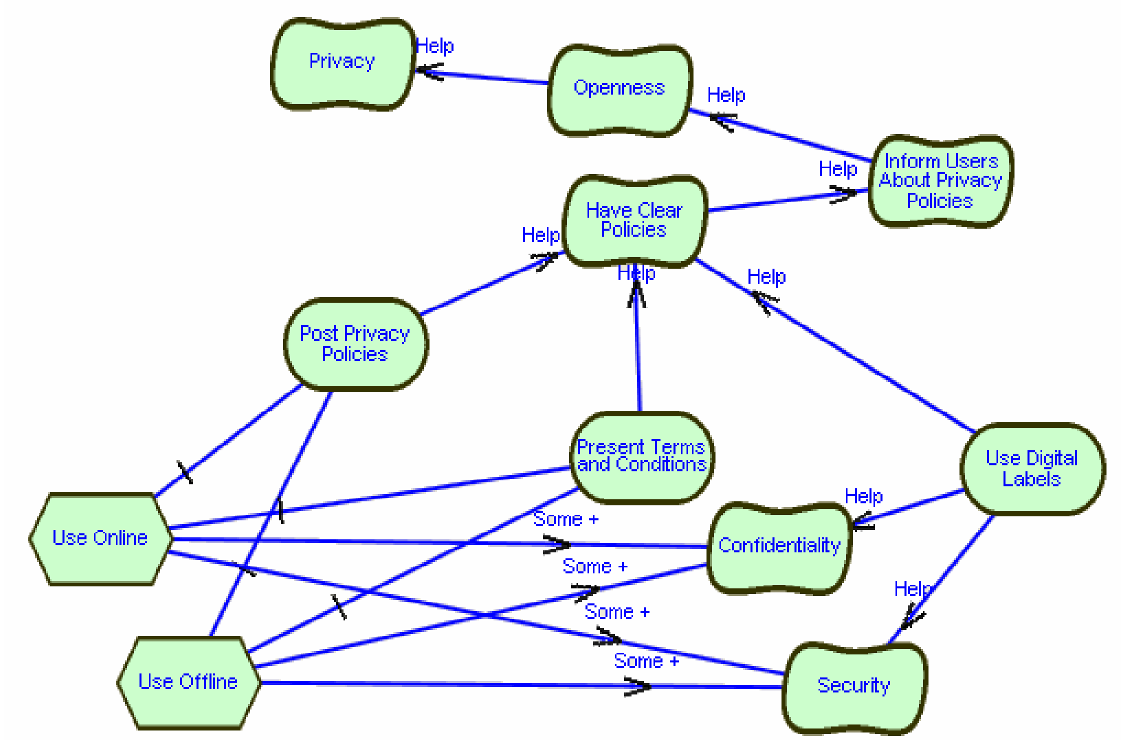

Figure 4 – Satisficing Openness Softgoal.

A way to achieve Openness is to Inform Users About Privacy Policies. We then established the goal Have Clear Policies would help in the operationalization of this softgoal. The softgoal Have Clear Policies has three goals: Post Privacy Policies, Present Terms and Conditions and Use Digital Labels. Post Privacy Policies and Present Terms and Conditions can be operationalized online or offline. Because these operationalizations cause direct impacts on the softgoals Confidentiality and Security, we analyzed these impacts. Some users might be confident enough to post privacy policies online. Others will prefer to use offline methods. When posting privacy policies online or offline one must consider the security aspect of the information available.

Another way to use the catalogue is to check if some specific operationalization would hurt any other NFR. Suppose the requirements engineer believes it is necessary to Use Digital Labels. In this case, the catalogue could be used to check how pursuing this approach would help or hurt other NFRs. Although sometimes this correlation may be clear, there are times when it is not clear at all. Having an available collection of knowledge on this matter could be of great help for requirements engineering. The next section shows an example from our case study to better illustrate the use of the privacy catalogue.

4. Using the Catalogue

To evaluate the use of the catalogue we focused on the Smart Systems for Health Agency (SSHA) and the University Health Network (UHN), which are web based initiatives related to PHI processing in Ontario, Canada. The PHI processing system has the objective of managing patients Personal Health Information (PHI). This includes storing, transmitting, controlling the access and the disposal of this type of information. It implements several measures as means to guarantee the accomplishment of these tasks. These involve the use of Public Key Infrastructure (PKI), Secure Sockets Layer (SSL), encryption, access audit trail, browse monitoring, protecting from denial of services, authenticate access and secure disposal of information.

In this work we show how the use of the catalogue can help to address privacy requirements. The scenario described above was modelled with the ***i**** framework[13]. Then we applied a systematic process to model privacy using the catalogue as a guide. Due to space limitation the process for doing this is not detailed. It will be the subject of future work. On figure 5 we only show how the catalogue can guide the evaluation of the solutions adopted by the PHI Processing System.

Using the Privacy subgoal Limiting Use, Disclosure and Retention of PHI and its subgoals it is possible to reason about the contributions of each of the solutions implemented by the PHI Processing System towards privacy. For example, Secure Disposal helps to satisfice the privacy subgoal Enforce the Disposal of PHI no Longer Needed. The two tasks that are part of this subgoal, respectively Use True Deletion Software and Use Physical Disposal and Destroy are marked with a checkmark which means they satisfice the subgoal Enforce the Disposal of PHI no Longer Needed.

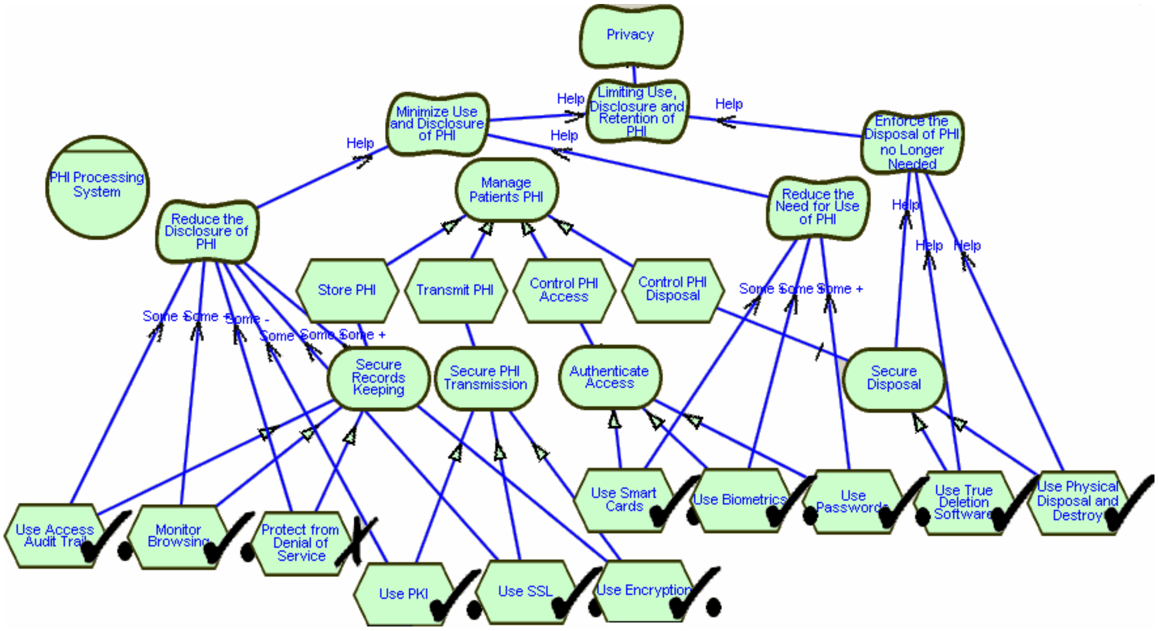

Figure 5 – Reasoning on Privacy

Minimize Use and Disclosure of PHI is the other subgoal of Limiting Use, Disclosure and Retention of PHI. In the Privacy Catalogue, this subgoal is further subdivided in Reduce the Disclosure of PHI and Reduce the Need for Use of PHI. Analyzing other solutions implemented in the PHI Processing System, we found Use Access Audit Trail and Monitor Browsing contribute, to some extent, to satisfice the subgoal Reduce the Disclosure of PHI. However, Protect from Denial of Service does not contribute to satisfice this same subgoal. This is because this task only protects the website from attacks that could prevent it from providing services. Tasks that positively contribute to satisfy a softgoal are labeled with a checkmark followed by a dot. Tasks that do not contribute to satisfice a softgoal at all are marked with an X.

Further analyzing the contribution of the implemented solutions to privacy we found Secure Records Keeping can, to some extent, positively contribute to satisfice this subgoal. Therefore, the tasks that represent a means of operationalizing the goal Secure Records Keeping contribute, to some extent, to satisfice the subgoal Reduce the Disclosure of PHI.

Finally, figure 5 shows Control PHI Access contributes, to some extent, to satisfy the subgoal Reduce the Need for Use of PHI. In the PHI Processing System the PHI access control is done via Authentication and involves the Use of Passwords, Use of Smart Cards and Use of Biometrics. This means these tasks contribute positively, to some extent, to satisfice the subgoal Reduce the Need for Use of PHI.

5. Conclusion

In this paper we present a reusable catalogue with strategies to satisfice privacy requirements. This catalogue collects knowledge on achieving privacy goals from the literature and enriches it with personal experiences also. The aim of this catalogue is to help requirements engineers to address privacy requirements from the early stages of software development. Besides, it aims to alert requirements engineering for possible conflicts that might arise from the availability of different choices to satisfice privacy. Because this is a very large catalogue, it was only possible to show part of it here (around 30%). However the full catalogue is available at [19].

This catalogue does not intend to be complete. It is designed to have experiences of other requirements engineers added to it. The examples presented here are based on the ***i**** framework but, being goal oriented, they can help requirements engineers who use other goal oriented approaches, such as KAOS [20] and GBRAM [21].

We recognize that the structure used to store and retrieve information from this catalogue can pose some challenges. However, at the present moment we are not aware of other means of representation that would facilitate understanding and use. Furthermore, from the results of our case study we believe the present representation allows the requirements engineer to benefit from the knowledge represented in the catalogue.

Future work will involve investigating different alternatives for storing and retrieving information from catalogues such as the one presented here, in order to facilitate retrieving information at different levels of granularity. It will also try to expand the knowledge base presented here to incorporate the needs from a larger community.

References

[1] Government of Alberta Enterprise Architecture Privacy Overview, IBM Global Services, 2003.

[2] Privacy Impact Assessments – Government of Canada – http://privcom.gc.ca/pia-efvp/index_e.asp.

[3] Personal Health Information Protection Act, Ontario – 2004.

[4] Breitman, K. K, Leite J.C.S.P. e Finkelstein Anthony. The World's Stage: A Survey on Requirements Engineering Using a Real-Life Case Study. Journal of the Brazilian Computer Society No 1 Vol. 6 Jul. 1999 pp:13:37.

[5] Lindstrom, D.R. "*Five Ways to Destroy a Development Project"* IEEE Software, September 1993, pp. 55-58.

[6] Chung, L., Nixon, B. *"Dealing with Non-Functional Requirements: Three Experimental Studies of a Process-Oriented Approa*ch*"* Proc. 17th Int. Con. on Software Eng. Seattle*,* Washington, April pp: 24-28, 1995.

[7] Cysneiros,L.M. and Leite, J.C.S.P. "*Non-Functional Requirements: From Elicitation to Conceptual Model"* IEEE Transactions on Software Engineering – May, 2004 (Vol. 30, No. 5).

[8] Mylopoulos,J. Chung, L., Yu, E. and Nixon, B.*, "Representing and Using Non-functional Requirements: A Process-Oriented Approach",* IEEE Trans. on Software Eng, 18(6), pp:483-497, June 1992.

[9] Simon,H.A. "The Sciences of the Artificial", 3rd edition MIT Press, 1981.

[10] Yu, E. and Cysneiros, L.M.; *"Designing for Privacy in the Presence of Other Requirements",* in Proc. of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems, pp:283-297, july 2002.

[11] "Inventory of instruments and mechanisms contributing to the implementation and enforcement of the OCDE privacy guidelines on global networks" Head of Publications Services, OECD, 2 rue-André-Pascal, 75775 Paris Cedex 16, France.

[12] Antón, A.I. and Earp., J.B. "A taxonomy for Web Site Privacy Requirements" NCSU Technical Report TR-2001-14, 18 December 2001.

[13] Yu, E. "Towards Modelling and Reasoning Support for Early-Phase Requirements Engineering" *in Proc. of the 3rd IEEE Int. Symp. on Requirements Engineering*, pp:226-235, 1997.

[14] Personal Information Protection and Electronic Documents Act, Canada, 2004.

[15] Yu, Eric; "Modelling Strategic Relationships for Processing Engineering", Ph.D. Thesis, University of Toronto, 1994.

[16] Cysneiros, L. M. and Yu, E.; "Requirements Engineering for Large-Scale Multi-Agent Systems"; in: A. Garcia, C. Lucena, A. Omicini, F. Zambonelli, J. Castro (Eds). "Software Engineering for Large-Scale Multi-Agent Systems". Springer-Verlag, LNCS 2603, April, 2003.

[17] Antón, A. I.; Earp, J. B. and , Reese, A. ; "Analyzing Website Privacy Requirements Using a Privacy Goal Taxonomy", IEEE Joint International Requirements Engineering Conference 2002.

[18] OME3 Tool, "http://www.cs.toronto.edu/km/GRL/."

[19] http://www.math.yorku.ca/~cysneiro/nfrs/PrivacyCanada/canadiancatalogue.png
[20] VanLamsweerde, A. *Goal-Oriented Requirements Engineering: A Guided Tour*. in *Proc. of 5th IEEE Int. Symp. on Requirements Engineering*. 2001: IEEE.
[21] Potts, C., K. Takahashi and A. I. Antón., *Inquiry-Based Requirements Analysis.* IEEE Software, 1994. march: p. 21-32.