Instituto de Computação
Universidade Estadual de Campinas

Políticas de Privacidade Semânticas para Descrição e Descoberta de Serviços na Arquitetura Orientada a Serviços

Este exemplar corresponde à redação final da Tese devidamente corrigida e defendida por Diego Zuquim Guimarães Garcia e aprovada pela Banca Examinadora.

Campinas, 10 de agosto de 2011.

Maria Beatriz Filgar de Toledo
Maria Beatriz Felgar de Toledo (Orientadora)

Miriam A. M. Capretz
Miriam Akemi Manabe Capretz (Co-orientadora)

Tese apresentada ao Instituto de Computação, UNICAMP, como requisito parcial para a obtenção do título de Doutor em Ciência da Computação.

FICHA CATALOGRÁFICA ELABORADA POR
MARIA FABIANA BEZERRA MÜLLER - CRB8/6162
BIBLIOTECA DO INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E
COMPUTAÇÃO CIENTÍFICA - UNICAMP

Garcia, Diego Zuquim Guimarães, 1982-
G165p Políticas de privacidade semânticas para descrição e descoberta de serviços na arquitetura orientada a serviços / Diego Zuquim Guimarães Garcia. – Campinas, SP: [s.n.], 2011.

Orientador: Maria Beatriz Felgar de Toledo.
Coorientador: Miriam Akemi Manabe Capretz.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Computação.

1. Direito à privacidade. 2. Arquitetura orientada a serviços. 3. Serviços na Web – Semântica. I. Toledo, Maria Beatriz Felgar de, 1959-. II. Capretz, Miriam Akemi Manabe. III. Universidade Estadual de Campinas. Instituto de Computação. IV. Título.

Informações para Biblioteca Digital

Título em Inglês: Semantic privacy policies for service description and discovery in service-oriented architecture
Palavras-chave em Inglês:
Right of privacy
Service-oriented architecture (Computer science)
Web services - Semantics
Área de concentração: Ciência da Computação
Titulação: Doutor em Ciência da Computação
Banca examinadora:
Maria Beatriz Felgar de Toledo [Orientador]
Abdelkader Ouda
Arlindo Flavio da Conceição
Edmundo Roberto Mauro Madeira
Jagath Samarabandu
Data da defesa: 10-08-2011
Programa de Pós Graduação: Ciência da Computação

TERMO DE APROVAÇÃO

Tese Defendida e Aprovada em 10 de agosto de 2011, pela Banca examinadora composta pelos Professores Doutores:

Prof. Dr. Abdelkader Ouda
Department of Electrical and Computer Engineering / UWO

Prof. Dr. Arlindo Flávio da Conceição
Departamento de Ciência da Computação / UNIFESP

Prof. Dr. Jagath Samarabandu
Department of Electrical and Computer Engineering / UWO

Edmundo Madeira

Prof. Dr. Edmundo Roberto Mauro Madeira
IC / UNICAMP

Maria Beatriz Felgar de Toledo

Profª. Drª. Maria Beatriz Felgar de Toledo
IC / UNICAMP

Instituto de Computação
Universidade Estadual de Campinas

Políticas de Privacidade Semânticas para Descrição e Descoberta de Serviços na Arquitetura Orientada a Serviços

Diego Zuquim Guimarães Garcia

Agosto de 2011

Banca Examinadora:

- Dra. Maria Beatriz Felgar de Toledo – Orientadora (Presidente) (IC-Unicamp)
- Dr. Abdelkader Ouda (ECE-UWO)
- Dr. Arlindo Flavio da Conceição (DCT-UNIFESP)
- Dr. Edmundo Roberto Mauro Madeira (IC-Unicamp)
- Dr. Jagath Samarabandu (ECE-UWO)
- Dr. André Santanchè (Suplente) (IC-Unicamp)
- Dra. Islene Calciolari Garcia (Suplente) (IC-Unicamp)
- Dra. Olga Nabuco (Suplente) (CTI)

Resumo

A privacidade pode ser definida como o direito de um indivíduo de ter informações sobre ele acessadas e usadas em conformidade com aquilo que ele considera aceitável. A preservação da privacidade é um problema em aberto na Arquitetura Orientada a Serviços (AOS). Uma solução para esse problema deve incluir características que apoiem a preservação da privacidade em cada área da AOS. Esta tese foca as áreas da descrição e descoberta de serviços. Os problemas nessas áreas são que não é possível descrever como um provedor de serviços usa as informações recebidas de um consumidor de serviços e descobrir serviços que satisfaçam as preferências de um consumidor. Diversos trabalhos de pesquisa têm sido realizados nessas áreas, mas ainda não existe um framework que ofereça uma solução que apoie uma rica descrição das políticas de privacidade e a sua integração no processo de descoberta de serviços. Consequentemente, o principal objetivo desta tese é propor um framework de preservação de privacidade para as áreas de descrição e descoberta de serviços na AOS. O framework aprimora a descrição e descoberta de serviços com a especificação e interseção das políticas de privacidade usando uma ontologia base de privacidade e ontologias de privacidade de domínios específicos. Além disso, o framework melhora essas áreas com uma extensão da AOS tradicional, a qual inclui dois novos papéis responsáveis por implementar um repositório de políticas de privacidade e intermediar as interações entre os consumidores e provedores e o componente de preservação de privacidade. A infra-estrutura proposta pelo framework foi implementada e avaliada através de um cenário no domínio da saúde, uma vez que a preservação da privacidade é uma questão importante nesse domínio.

Abstract

Privacy can be defined as the right of an individual to have information about them accessed and used in conformity with what they consider acceptable. Privacy preservation in Service-Oriented Architecture (SOA) is an open problem. A solution for this problem must include features that support privacy preservation in each area of SOA. This thesis focuses on the areas of service description and discovery. The problems in these areas are that currently it is not possible to describe how a service provider deals with information received from a service consumer as well as discover a service that satisfies the privacy preferences of a consumer. Research has been carried out in these areas, but there is currently no framework which offers a solution that supports a rich description of privacy policies and their integration in the process of service discovery. Thus, the main goal of this thesis is to propose a privacy preservation framework for the areas of service description and discovery in SOA. The framework enhances service description and discovery with the specification and intersection of privacy policies using a base and domain-specific privacy ontologies. Moreover, the framework enhances these areas with an extension to basic SOA that includes roles responsible for implementing a privacy registry as well as mediating the interactions between service consumers and providers and the privacy preservation component. The framework is evaluated through a health care scenario as privacy preservation is an important issue in this domain.

Agradecimentos

Agradeço às minhas orientadoras, Profa. Beatriz Toledo e Profa. Miriam Capretz, pelo apoio.

Este trabalho foi apoiado por FAPESP e CAPES.

Sumário

Lista de Figuras

Lista de Tabelas

Chapter 1

Introduction

This chapter introduces the work by presenting its motivation, giving an overview of the proposal and discussing its goals. Finally, the chapter presents the organization of the rest of this thesis.

1.1 Motivation

Service-Oriented Architecture (SOA) [18] is a software architecture based on the concept of service, a loosely coupled, abstract and discoverable software component. SOA has been an intense area of research because of its potential to facilitate the development and management of software solutions. However, SOA still has open problems [31] that must be addressed in order to enable its wider application. Privacy preservation is one of these problems. Privacy [46] can be defined as the right of an individual to have information about them accessed and used in conformity with what is considered acceptable by that particular individual.

SOA includes two mandatory roles: service consumer and provider. A consumer uses a service provided by a provider. The service provider usually requires information from the service consumer so that the consumer can use the service supplied by the provider. This can include private information. Thus, the consumer needs to know how the provider will use its information so that the consumer can decide whether to disclose the information to that provider or try another alternative. This is the general problem of privacy preservation in SOA [20] and it is related to the concern of the consumer that disclosed information can be misused by providers receiving it.

The problem of privacy preservation in SOA demands solutions that include privacy enhancing mechanisms in the different areas of SOA. This thesis focuses on the areas of service description and discovery. In basic SOA, service description is restricted to functional characteristics of services. As a consequence, service discovery is based on functionality of services. Extensions to SOA were proposed in order to include non-functional or Quality of Service (QoS) characteristics of services in service description. These extensions allow for service discovery that considers not only the functionality of the service but also the non-functional characteristics of the service. However, there still is a lack of an extension for privacy preservation [44]. Thus, the privacy preservation problems in the areas of service description and discovery are that it is not possible to describe how a service provider deals with private information received from a service consumer and discover a service that satisfies the privacy preferences of the consumer.

Work that has been done on privacy in SOA does not offer a proper solution for the problems in the areas of service description and discovery. Privacy frameworks proposed in the literature have limitations including limited privacy policy model, privacy vocabulary as well as support for privacy policy specification and intersection as they do not use, for example, concepts defined in ontologies for creating policies. Furthermore, existing privacy preservation frameworks have no service discovery integration. Finally, such frameworks do not have proper support for the inclusion of other QoS attributes and for the consideration of domain-specific privacy preservation issues.

1.2 Overview

This thesis proposes a solution for the problems of privacy preservation in the areas of service description and discovery in SOA. The proposed solution is a privacy preservation framework that addresses the limitations identified in privacy frameworks for SOA proposed in the literature.

The privacy framework proposed in this thesis includes a policy model, which enables the description of privacy practices and preferences of service providers and consumers. In the policy model, policy assertions refer to ontological concepts. Thus, policies are created from concepts defined in privacy ontologies. This semantic information supports the matching between the policies of a consumer and provider. Moreover, the framework includes privacy-aware service discovery, which enables the discovery of services that meet privacy preferences of consumers.

In the approach proposed in this thesis, service providers and consumers describe their privacy preservation practices and preferences in policies. Thus, policy intersection enhances service discovery so that discovered services are from providers whose privacy practices match the privacy preferences of the consumer. The use of policies for service discovery is accomplished by extending SOA with two new roles: privacy and mediator. The privacy role is responsible for the publication and discovery of privacy policies. The mediator role mediates the interactions of service publication and discovery between the provider or consumer and the publication and discovery space, which includes the service registry and the privacy.

Privacy preservation is a problem in several domains. Some privacy preservation issues are common to different domains, but it is important to consider that each domain includes specific privacy issues. Typically, a general privacy preservation regulation [9] deals with common issues and a separate privacy regulation [28] can complement it with domain issues. In order to address this aspect of privacy preservation, the solution proposed in this thesis follows an approach in which general privacy issues are represented by a base privacy ontology and domain-specific issues are captured by ontologies that extend the base ontology.

Among the different domains, health care is an example in which privacy preservation is particularly important, as health information is usually regarded as sensitive. Thus, the health care domain was chosen to evaluate the framework. The evaluation involves the demonstration of cases in which service consumers, which look for services in a health care scenario, have their privacy preservation preferences checked against the privacy preservation practices of service providers so that the consumers can decide whether to select or not the services offered by those providers.

The main contribution of this thesis is a framework that supports privacy preservation in service description and discovery in SOA. The framework allows service consumers to select services that not only meet the functionality required by the consumers but also satisfy their privacy preservation preferences. Specifically, the contributions of this thesis are a model for semantic privacy policy, which enables the specification of policies using concepts defined in a base privacy ontology and domain-specific privacy ontologies, as well as privacy-aware service discovery, which enables the use of privacy policies of consumers and providers as well as their intersection in service publication and discovery. Differently from existing privacy frameworks, the policy model of the proposed framework enables a flexible specification of privacy practices and preferences, defines a comprehensive privacy vocabulary, allows for the use of privacy ontologies and takes domain-specific issues into consideration. In terms of the SOA extension of the proposed framework, the differences from existing privacy frameworks are that it keeps compatibility with basic SOA, integrates privacy policies in service discovery and supports its extension in order to deal with other non-functional characteristics.

This work follows an approach that is used in Web service technology in order to deal with security. In Web service technology, security (Web Services Security – WS-Security [27]) and policy (Web Services Policy – WS-Policy [42]) standards are used together in order to create security policies for Web services. The privacy policies created in this work can be used in combination with policies for other aspects in order to improve the non-functional support in SOA. Thus, the privacy preservation framework proposed in this thesis should be considered as one component of a set of components that would create a comprehensive security framework for SOA.

1.3 Goals

The main goal of this thesis is to propose a privacy preservation framework for the areas of service description and discovery in SOA. Specifically, the goals are:

- The creation of a privacy policy model using ontologies to enhance service description with privacy preservation practices and service request with privacy preservation preferences. This goal can be accomplished by defining elements and their organization in a format that enables intersection and the use of an ontological approach to support a rich description of privacy policies.

- The integration of privacy preservation-awareness in service publication and discovery in order to enable the publication of privacy practices of service providers and a process of service discovery that considers privacy preferences of service consumers. This goal can be accomplished by extending SOA with new roles and interactions, which enable the use of the proposed policy model in order to support the consideration of privacy preservation practices of providers and consumer preferences in the process of service discovery.
- The application of the privacy preservation framework to a scenario in the domain of health care in order to evaluate the effectiveness of the proposed SOA privacy framework. This goal can be accomplished by developing a health care privacy ontology that extends the base ontology as well as creating a health care scenario that enables the definition and execution of evaluation cases to demonstrate the privacy preservation capabilities of the framework, which includes the solutions for the first two goals.

1.4 Organization

The rest of this thesis is organized as follows:

- Chapter 2 presents background information. It contextualizes the thesis by introducing the concepts of SOA, privacy and ontology. It also presents the main technologies used for implementing the proposed framework.
- Chapter 3 presents related work. This chapter reviews the literature in SOA privacy preservation by surveying existing SOA privacy frameworks. It also elaborates on the necessity of a privacy preservation solution by discussing the limitations of existing frameworks.
- Chapter 4 gives an overview of the framework proposed in this thesis that offers solutions for the identified limitations.
- Chapter 5 presents the first part of the framework. It describes the semantic privacy policy model that enhances service description, including the policy format and base privacy ontology.
- Chapter 6 presents the second part of the proposed framework. It describes the extensions to basic SOA that support the use of the privacy policy model for enhancing service discovery.
- Chapter 7 presents the implementation and evaluation of the proposed privacy framework. It introduces the health care ontology, scenario and cases that were developed in order to evaluate the effectiveness of the framework.
- Chapter 8 presents conclusions. It describes the contributions of this thesis and discusses possible future work.

Chapter 2
Background

This chapter presents basic concepts involved in this thesis. In Section 2.1, Service-Oriented Architecture (SOA) is described as it establishes the context for this work. The concept of privacy is discussed in Section 2.2 as this work tackles the problem of privacy preservation in the areas of service description and discovery in SOA. Finally, Section 2.3 presents the concept of computational ontology as the use of ontologies is proposed in order to improve the solution for privacy preservation in SOA proposed in this work.

2.1 Service-Oriented Architecture (SOA)

SOA [31] is a software architecture based on the concept of service. A service is a software component with three main characteristics: abstraction, discoverability and loose coupling. As shown in Figure 2.1, SOA [18] has three main roles: service provider, service consumer and service registry. A service provider hosts a service and publishes a description of the service to a service registry. A service consumer that needs a service to accomplish a task discovers a service from a service registry and uses the description of the discovered service in order to bind and interact with the service provider.

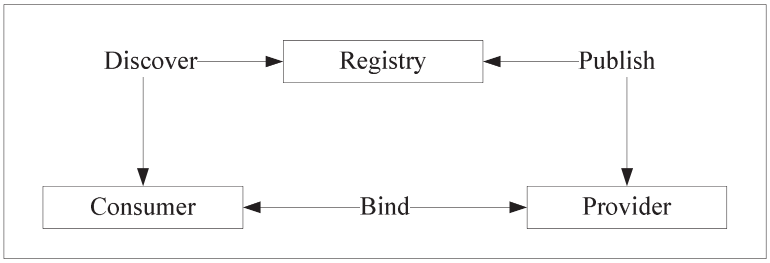

Figure 2.1. SOA roles.

2.1.1 Layers and Infrastructure

SOA [6] facilitates the development and management of services that cross the boundaries of applications. SOA [23] features a set of layers with a clear separation between presentation, business processes, services and applications (Figure 2.2).

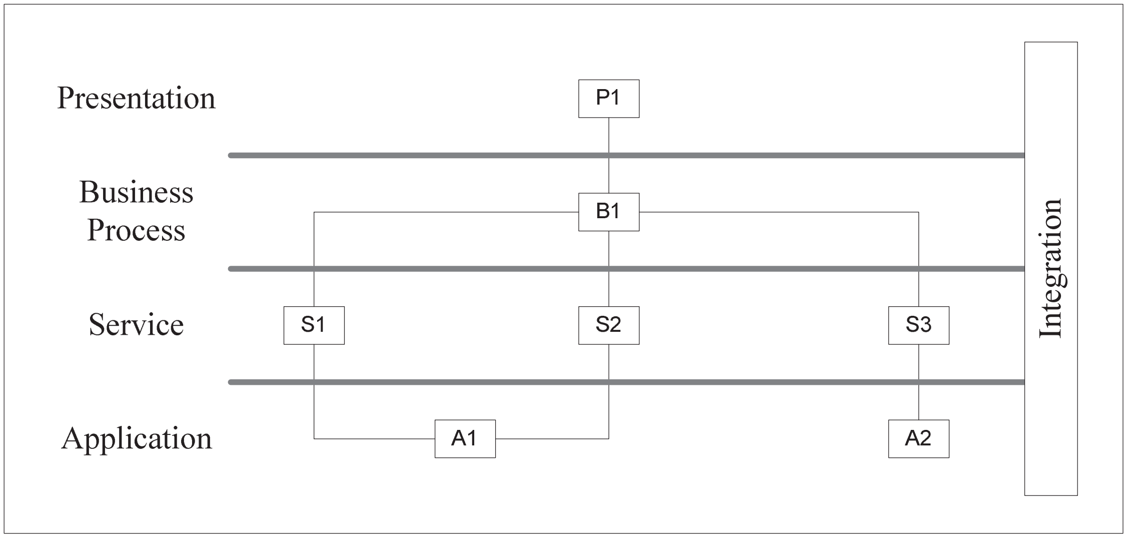

Figure 2.2. SOA layers.

The layers of SOA are described as follows:

- Presentation: is the entry point for end users and business partners, comprising user interfaces and externally accessible services.
- Business Process: comprises business processes that model solutions exposed in the Presentation layer and are created from services contained in the Service layer. In Figure 2.2, a business process (B1) is exposed by an interface (P1) in the Presentation layer.
- Service: provides standardized interfaces that enable services implemented by different applications to be composed and interoperate in a business process. In Figure 2.2, the three services (S1, S2 and S3) in the Service layer create the business process (B1) in the Business Process layer.

- Application: includes software applications that constitute implementations of services. In Figure 2.2, an application (A1) implements two services (S1 and S2) and another application (A2) implements the third service (S3) contained in the Service layer. Thus, each service interfaces a different operation or operation set realized by an application.
- Integration: deals with concerns that cut across the other SOA layers, such as Quality of Service (QoS), monitoring and management. QoS refers to the non-functional characteristics of services, for example, security and availability. Monitoring and management involve the use of techniques to detect problems and to improve solutions.

The infrastructure of SOA is supported by an Enterprise Service Bus (ESB) [33], which is responsible for connecting services that represent applications. The ESB provides features, such as message delivery, service publication and discovery (service registry) as well as the features included in the Integration layer of SOA. The features provided by an ESB are usually needed for different services and they are also modeled as services. The ESB features can be implemented using the most suitable solution available and they can be added to the ESB as needed. Thus, the ESB abstracts common concerns of services in SOA, further facilitating the development and management of services.

SOA includes several areas of research, for example, service description, discovery and composition. This thesis focuses on the areas of description and discovery. Service description is a document that includes information on a service. This information can include the functionality of the service, its non-functional characteristics as well as information on where and how to access and use the service. This document can be directly passed to a service consumer by a service provider so that they can interact. In this case, the parties should know each other in advance. When this is not the case, then a service registry can be used, which facilitates service publication and discovery. The registry offers providers a mechanism for making service descriptions available to consumers. Thus, a provider can use this mechanism to publish its service so that it can be discovered by consumers. In order to discover a service, a consumer uses another mechanism provided by the registry. This mechanism allows the consumer to inform its requirements for the service, which can include functional and non-functional requirements. The registry is responsible for performing the discovery process, searching for a service that matches the requirements of the consumer.

2.1.2 Web Services

One of the strengths of SOA is Web service technology. Web service [5] is a technology that can be used to implement SOA. Web service technology has been supported by major software companies, including Hewlett-Packard (HP), International Business Machines (IBM), Microsoft, Oracle and Sun Microsystems. These companies, together with several other companies, have delivered standards for Web services [7] in order to accomplish the vision of seamless application integration. The vision of seamless application integration is supported by the standardization of several aspects of the service life cycle, such as security (Web Services Security – WS-Security [27]) and policy (Web Services Policy – WS-Policy [42]).

Web service technology comprises three basic standards:

- Web Services Description Language (WSDL) [8]: WSDL is a language for describing the functionality of a service.
- SOAP [26] (formerly Simple Object Access Protocol): SOAP is a protocol for message exchange among services.
- Universal Description Discovery & Integration (UDDI) [11]: UDDI is a registry that supports service publication and discovery.

2.2 Privacy

A paper [45] published in 1890 is often cited in the literature in order to provide a definition of the concept of privacy. According to the authors of the paper, the right to be left alone is considered to define privacy. The paper by Warren and Brandeis is often cited in the literature because the authors first discussed the issue that privacy includes injury of feelings, as a result from disclosing private information to the public, in addition to the concept of physical privacy.

In another influential work [46], the claim of individuals and groups for determining for themselves how information is communicated defines privacy. The definition by Westin of the concept of privacy suggests that an individual should have a means to control the access to information about the individual. The definition of the concept of privacy is valid offline and online. However, the range of privacy risks is broader in electronic environments than offline. The actions of the individuals are typically recorded over a long period of time online. Furthermore, a large amount of information pieces of the individuals is collected by a number of organizations. Moreover, the capabilities of information processing are getting higher and higher. All of these possibilities increase the risks to privacy.

Thus, giving the individuals a means to control the access to their information is a part of privacy. Another important part of privacy is to control the use of information that is no longer under the control of the individuals in order to avoid that private information is used in an unacceptable way. In this thesis, privacy is defined as the right of an individual to have information about them accessed and used in conformity with what that particular individual considers acceptable.

2.2.1 Individuals – Surveys

In 2009, a survey [12] was conducted in Canada in order to understand the views of individuals on privacy issues. The survey examined the levels of awareness, understanding and concerns of the individuals. The results of the survey showed a general concern among the respondents about the protection of their private information. Two thirds of the respondents were not confident organizations can adequately safeguard information. Furthermore, the majority of the respondents agreed on the statement that privacy preservation would be one of the most important issues in the next decade. Regarding new technologies, the results of the survey showed that almost half of the respondents were concerned about the impact of the new technologies on privacy preservation.

In the United States of America, another survey [39] was conducted in 2009 in order to determine the opinions of individuals about the use of behavioral targeting by marketer. The use of behavioral targeting has been a controversial issue before government policy makers. Behavioral targeting involves tracking the actions of the individuals and then tailoring advertisements for the individuals based on their actions. The survey discovered that most adult respondents did not accept tailoring advertisements to their interests, in opposition to the claim of many marketers. This finding was valid even among young adults (between 18 and 24 years of age), who have often been portrayed by advertisers as caring little about privacy. A high percentage of adult respondents rejected the gathering of information about individuals for tailoring advertisements by marketers. Moreover, another finding of the survey was that a large proportion of respondents rejected even anonymous behavioral targeting.

The two surveys [12], [39] and other surveys [17], [13], [32] on privacy provide information that allows us understanding the impact of privacy concerns on the behaviors of the individuals and the acceptability of the new technologies. For example, the surveys report that a high percentage of the respondents have decided not to use a service due to concerns about the use of private information.

Although it could be thought that privacy was not regarded as essential by many individuals due to the widespread adoption of information-intensive services and the lack of sufficient protection of the personal information of the individuals, a study [37] has shown that privacy is an important issue for the majority of the individuals. In the study, some participants were provided with simple information on the privacy policies of websites while other participants were not provided with the information. The first group of participants was more likely to use websites with better policies than the second group of participants. Moreover, a survey [22] on mobility pricing systems has investigated the willingness-to-pay for privacy of individuals. It has shown that the majority of the respondents have accepted paying a higher cost in order to maintain a higher level of privacy.

2.2.2 Individuals – Concerns

A study [34] was conducted in order to develop a measurement instrument for information privacy research. The instrument helps measure the concerns of the individuals about the privacy practices of the organizations. The concerns are listed and described as follows:

- Collection: a large amount of information is collected and stored.
- Internal Unauthorized Secondary Use: the information is collected for a purpose, but the information is used for another purpose internally within the organization that has collected the information.
- External Unauthorized Secondary Use: the information is collected for a purpose, but the information is used for another purpose by an external party after disclosure by the collecting organization.
- Improper Access: the information about the individual is readily available to people not properly authorized to access the information.
- Errors: the protection against deliberate and accidental errors in information is inadequate.
- Reduced Judgment: the excessive automation of the decision-making process leads to inadequate decisions.
- Combining Data: the information from different databases is combined in larger databases.

A more recent study [24] drew on the theory of social contract in order to characterize the notion of information privacy concerns of the Internet users. The social contract theory defines that contracts must be grounded in informed consent, must be reinforced by exit and must voice rights. Thus, the notion of information privacy concerns of the Internet users was characterized in terms of three factors as follows:

- Collection: represents the central theme of fair information exchange based on an agreed social contract.
- Control: represents the freedom to give an opinion or exit.
- Awareness: indicates understanding about the accepted conditions and actual practices.

2.2.3 Organizations

New regulations and concerns of individuals have motivated organizations to take into account privacy-preserving systems. Furthermore, there is a cost to the lack of privacy preservation. Organizations may have to pay fines for privacy preservation breaches, for instance. In addition to this cost, an analysis [1] on information security economics investigated the impact of privacy incidents on the market values of organizations and showed that privacy breaches can have a negative impact on the stock market. This study gathered several examples of private information breaches and executed various empirical analyses, whose results allow seeing that there was a relation between some privacy incidents of organizations and their market values.

Thus, it is important that organizations implement measures in order to preserve the privacy of individuals. However, on the other hand, the collection and use of private information is frequently a requirement in order for organizations to provide their services and can be an important component for achieving competitiveness. This creates a challenge for organizations, as organizations have to balance the attitude of privacy preservation and the necessity of taking business advantage from collecting and using private information of individuals.

2.2.4 Preservation

Privacy preservation is maintaining the privacy of an individual at the level required by the individual, that is, keeping the right of the individual to have information about them accessed and used in conformity with what the individual considers acceptable. Two different research lines can be identified in the area of privacy preservation [35]:

- Access prevention: the research line of access prevention focuses on developing protection mechanisms that prevent access to private information of individuals, for example, by making individuals anonymous. This is usually effective, as high levels of privacy can be maintained by restricting the identification of collected information. However, access prevention cannot always be used, since it may limit the functionality of services and hinder their marketing.
- Awareness and control: the research line of awareness and control focuses on increasing awareness of individuals and their control over information activities. This can lead to inadequate protection against privacy preservation attackers, as identifiable information continues to be collected, disclosed, retained and used. However, the application of awareness and control is typically wider than access prevention, because the identification of collected information is usually important for organizations in order to provide value-added services.

2.2.5 Regulations

A number of privacy regulations [40], [29], [14], [9] have been created around the world. The privacy regulations define several principles in order to support the preservation of the privacy of the individuals:

- Accountability: an organization is responsible for the information under its control.

- Identifying purposes: the purposes for which the information is collected are identified by the organization.
- Consent: the consent of the individual is necessary for the collection and use of the information.
- Limiting collection: the collection of the information is limited to the information which is needed for the purposes identified by the organization. Fair and lawful means is employed for information collection.
- Limiting use, disclosure and retention: the information is not used for purposes other than the purposes for which the information was collected. The information is retained only for the time period that is necessary for the fulfillment of the purposes.
- Accuracy: the information is correct, comprehensive and current as it is necessary for the purposes for which the information is to be used.
- Safeguards: the information is protected by the security safeguards appropriate to the sensitivity of the information.
- Openness: an organization makes readily available to individuals its information management practices.
- Individual access: upon request, an individual is informed of the existence and use of their information and information access is given to that individual. An individual can challenge the accuracy of the information and have the information corrected as appropriate.

- Challenging compliance: an individual is able to address a challenge concerning the compliance with privacy principles to a party accountable for the compliance of the organization.

2.3 Ontology

The definition of the concept of computational ontology by Gruber [16] is often cited in the literature. The author defines a computational ontology as a formal, explicit specification of a shared conceptualization. Each part of this definition indicates a characteristic of ontologies as follows:

- Conceptualization: an ontology is an abstract model of a domain in the world, which identifies the concepts and relationships among concepts of the target application domain.
- Explicit: an ontology defines the concepts and their relationships explicitly.
- Formal: an ontology is computer-processable.
- Shared: an ontology represents consensual knowledge.

There are different types of formal languages [36] that are used for specifying ontologies, including description logics and frame logics. Computational ontologies were created in the area of artificial intelligence mainly aiming at supporting knowledge sharing. Ontologies have been an intense subject of research in different fields of artificial intelligence, such as knowledge engineering and natural-language processing. More recently, the notion of ontology has become popular in other areas, such as information retrieval and integration as well as cooperative information systems. The reason for the widespread use of the concept of ontology [15] is due to the support it provides for the establishment of common understandings of domains that can be communicated among people and software applications.

2.3.1 Types

An ontology is created mainly to construct a model of a target domain. It provides a vocabulary that can be used to model the application domain. However, there are different ontology [41] types:

- Domain ontology: represents knowledge specific to a domain, for example, an ontology for the domain of health care.
- Metadata ontology: offers a vocabulary for describing the content of information sources, for example, an ontology for digital material such as video.

- Common sense ontology: captures general knowledge about the world, providing basic concepts that are valid across domains, for example, an ontology for the concept of time.
- Representational ontology: provides representational constructs in a domain-independent way, for example, an ontology for concepts of object orientation.

2.3.2 Web Ontology Language

As a result of the work of the World Wide Web Consortium (W3C) in the context of the Web Ontology Working Group as part of the W3C Semantic Web Activity, the Web Ontology Language (OWL) [43] was developed as an ontology standard for the Web. The OWL specification is endorsed as a W3C Recommendation. OWL extends the Resource Description Framework (RDF) and RDF Schema (RDFS) standards. OWL is a language that supports the creation of ontologies on the Web. The formal foundation of OWL is based on the description logics.

2.4 Summary

This chapter presented basic concepts involved in this thesis, including SOA, privacy and ontology. The chapter started with SOA by describing its layers and infrastructure as well as Web service technology. Then, the chapter discussed the concept of privacy as well as presented privacy preservation and regulations. Finally, ontologies and the OWL standard were presented.

Chapter 3
Related Work

This chapter reviews privacy frameworks for Service-Oriented Architecture (SOA) proposed in the literature. Two aspects were considered in the review of the frameworks:

- **Policy model**: how are privacy policies of service consumers and providers expressed in the framework?
- **SOA extension**: how is the basic architecture of SOA extended by the framework?

3.1 Policy Model

The following questions were considered in order to review the privacy policy model of the frameworks:

- **Format**: does the policy format defined by the framework allow for flexible specification of privacy policies?

A policy format is a standard structure that has to be followed by privacy policies defined by service consumers and providers. Thus, this first question asks if the framework defines a language that is used to structure policies in a way that they can be processed by computers. Several frameworks [21], [38], [4], [2], [30] assume the use of privacy policies by service consumers and providers, but these frameworks do not define a format for the privacy policies. Thus, these frameworks do not have a format or the format is not available and consequently the frameworks do not allow for the specification of computer-processable privacy policies. The existing frameworks [47], [3], [25] that define a format for privacy policies do not include support for flexibility in the policy format. Thus, these frameworks do not define rules that convert privacy policies to the standard structure and consequently the format is rigid. When these rules are present, consumers and providers can create flexible privacy policies that are converted to the standard structure before being processed. A flexible format includes constructs, for example, alternatives and optional assertions, which allow for richer privacy policy specifications.

- **Vocabulary**: does the privacy vocabulary defined by the framework cover the principles of privacy regulations?

A privacy vocabulary is a set of terms related to privacy and relationships among the terms that are used in the specification of privacy policies by service consumers and providers. Some frameworks [21], [2], [30] assume the use of a privacy vocabulary together with a format for privacy policies, but these frameworks do not define a privacy vocabulary. Thus, these frameworks do not include a vocabulary or the vocabulary is not available and consequently the frameworks do not allow for the specification of interoperable privacy policies. Several frameworks define a privacy vocabulary, but the vocabulary is limited. The privacy vocabulary of some frameworks [38], [4] includes the concepts of information and collector only. Other existing frameworks [47], [3], [25] define a privacy vocabulary that misses the concepts related to collection means, owner access and use record as well as the categorization of some concepts. Thus, these frameworks do not include terms and relationships that capture the principles defined in privacy preservation regulations and consequently the vocabulary is limited. When the principles of regulations are present, consumers and providers can create comprehensive privacy policies that cover a wide range of requirements and guarantees related to privacy preservation. A comprehensive privacy vocabulary, which includes concepts such as owner access and use record, allows for the specification of policies that can provide a higher level of privacy preservation.

- **Semantics**: does the support for semantics of the framework allow for the specification and intersection of semantic policies?

Meaning can be added to the information in a privacy vocabulary by including support for semantics in the framework. Several frameworks [21], [47], [4], [3], [2] do not include support for semantics. Thus, these frameworks do not have a privacy vocabulary enriched with semantic information or the semantics is not available and consequently the frameworks allow for the matching between the privacy policies of a service consumer and provider based on syntax only. The frameworks [38], [25], [30] that include support for semantics do not allow for the specification and intersection of semantic policies as these frameworks extend service ontologies. Thus, in these frameworks the privacy policy is a part of the service description and consequently the policy is not a separate document. When a privacy ontology is present, consumers and providers can create privacy policies that are easier to maintain as they are likely to change more often than the service descriptions. An ontology-based policy, such as an annotated policy, allows for the reuse of policies and the use of policy intersection for verifying the compatibility of privacy policies.

- **Domain**: does the framework define an approach to deal with domain-specific privacy issues?

Different domains, such as health and learning, have specific privacy issues in addition to the privacy issues that cross multiple domains. Several frameworks [38], [47], [4], [3], [25], [30] do not consider domain-specific privacy preservation issues. Thus, these frameworks do not have support for extension and consequently the frameworks do not allow for the specification of privacy policies that include concepts from a given domain. Some existing frameworks [21], [2] include placeholders for dealing with domain-specific privacy issues, but these frameworks do not define an approach to the application of the framework to different domains. Thus, these frameworks consider the importance of dealing with domain-specific privacy issues and consequently the frameworks are open for extensions. However, they do not define any approach as a part of the framework that drives the extension of the framework with concepts derived from domain-specific issues. The lack of a mechanism to implement the extension of the framework requires the definition of one by the user, which can affect the interoperability of the framework negatively.

3.2 SOA Extension

The following questions were considered in order to review the extension to the basic architecture of SOA of the frameworks:

- **Modification**: how does the framework modify the roles and interactions of basic SOA?

Some frameworks [21], [38], [47] modify basic roles of SOA, whereas other frameworks [4], [3], [2], [25], [30] add new roles to SOA. Between these two design choices, the second choice is the better one as it facilitates the deployment of the extension to an SOA environment. The new roles are added as services that are used by consumers and providers the same way as they use other services in the environment. The modification of basic roles, including consumer, provider and registry, is hard to deploy as the entities that are active in the environment need to be modified. Interactions related to privacy preservation are needed between the service consumer and provider in some frameworks [21], [3], [30]. This setting is not a good design choice as in basic SOA the decision on which service to use is done at discovery time and the consumer and provider start interacting after the decision. Thus, privacy-related interactions should involve a third party at publication and discovery times. All existing frameworks require direct interaction with the components responsible for privacy preservation. This setting is not a good design decision as it affects the scalability of the framework negatively when other non-functional characteristics are dealt with. Thus, direct interaction with the privacy components should be avoided.

- **Discovery**: does the framework integrate privacy policies in the process of service discovery?

No framework that integrates privacy policies in the process of service discovery has been identified in the literature. In the surveyed frameworks [21], [38], [47], [4], [3], [2], [25], [30], the service consumer has to perform actions after service discovery in order to receive services that meet the privacy preservation preferences of the consumer, for example, the consumer has to request the policy from the provider as well as forward it to the privacy component for verification or do it itself. Due to the lack of integration, consumers and providers may have to perform additional tasks or the number of interactions needed for a consumer to use a service may increase. The integration of privacy policies in the process of service discovery may lead to modifications to the registry, but they can be avoided. Thus, if the integration can be implemented without modifications to the registry, then it is a better design decision as it keeps compatibility with basic SOA as well as alleviates the burden on service consumers and providers.

- **Quality of Service (QoS)**: does the framework enable the inclusion of other QoS attributes with the separation of the different attributes?

QoS is a set of non-functional characteristics of services such as privacy, security and reliability. Although the framework proposed in this thesis has been developed specifically to deal with privacy preservation, it has to be prepared for working with other QoS attributes. The QoS attributes required in different environments and interactions vary. They should be dealt with separately as they are processed differently, for example, they need different matching rules. No framework that supports the inclusion of other QoS attributes with the separation of the different attributes has been identified in the literature. In order to deal with other QoS attributes in the surveyed frameworks [21], [38], [47], [4], [3], [2], [25], [30], the service consumer and/or the service provider have to interact with a set of components responsible for the QoS attributes or a single component is responsible for all QoS attributes in the framework. These two settings are not good design decisions. The first one affects the scalability of the framework negatively regarding consumers and providers, which have to interact with an increasing number of components that have to be discovered and bound to. The second design choice affects the performance of the framework negatively as a heavy component, which is responsible for processing all the requested QoS attributes, is included in the framework. In addition, new matching rules have to be added to the component when a new attribute is included in the framework.

3.3 Summary

The following limitations were identified in the privacy frameworks for SOA proposed in the literature:

- **Inflexible format for privacy policies**. A flexible format for privacy policies is required in order to support the specification of alternative privacy preservation practices and compact privacy policies.
- **Limited vocabulary of privacy preservation**. A privacy vocabulary that covers the principles of privacy preservation regulations is required in order to support the expression of complete privacy preservation practices.
- **Poor support for semantics**. A privacy framework that includes semantics, such as, by ontological annotation of privacy policies is required in order to support rich specification of privacy policies of service consumers and providers and intersection between the privacy policies of a consumer and provider.
- **Incomplete support for domain-specific issues of privacy preservation**. A privacy framework that enables the consideration of domain-specific privacy preservation

issues is required in order to support the application of the framework to different domains.

- **Inadequate modifications of basic roles of SOA and inclusions of interactions in basic SOA**. A privacy framework that does not modify basic roles of SOA as well as does not require direct interaction with the privacy preservation component and privacy-related interactions between the consumer and provider is required in order to support the deployment of the framework.
- **No integration in the process of service discovery**. A privacy framework that integrates privacy policies in the process of service discovery is required in order to support a privacy-aware process of service discovery.
- **Improper support to QoS extension**. A privacy framework that enables the inclusion of other QoS attributes with the separation of the different attributes is required in order to support QoS management without extra impacts on scalability and performance.

A privacy framework that addresses the limitations identified in the existing frameworks is proposed in this thesis. The proposed framework includes a flexible format for privacy policies and a privacy vocabulary that covers the principles of privacy preservation regulations. Additionally, the framework includes semantics by ontological annotation of privacy policies and enables the consideration of domain-specific privacy preservation issues.

The framework does not modify basic roles of SOA as well as does not require direct interaction with the privacy preservation component and consumer-provider privacy-related interactions. It integrates privacy policies in the process of service discovery. Finally, the framework enables the inclusion of other QoS attributes with the separation of the different attributes.

Chapter 4
Privacy Preservation Framework

This chapter describes the privacy preservation framework for the areas of service description and discovery in SOA proposed in this thesis.

4.1 Overview

Privacy preservation is an open problem in SOA. The framework proposed in this thesis employs policies in order to support the right of an individual to have information about them accessed and used in conformity with what is considered acceptable by them. Consumers and providers use policies to express their preferences and practices regarding privacy. The framework defines a model for enhancing service description with privacy through the use of policies, so that service description includes information on privacy practices of providers as well as request includes information on preferences of consumers. In addition, the framework uses these policies in order to enhance service discovery with privacy-awareness, so that service publication includes privacy policies as well as discovery considers the preferences and practices of consumers and providers when selecting services. This way, the framework supports privacy preservation in SOA.

In the framework, in addition to the mandatory and optional basic roles, new roles are added to SOA in order to deal with privacy policies. A consumer uses a service provided by a provider. The provider usually requires information from the consumer so that the consumer can use the service supplied by the provider. This can include private information. If the consumer needs to know how the provider will use its information so that the consumer can decide whether to disclose the information to that provider or not, the consumer can create a policy that specifies its privacy preferences. Providers can have policies describing their privacy practices in the context of services provided by them. Provider policies are published so that they can be considered when discovering services for consumers concerned with their privacy. Thus, the framework offers a solution for the problem of privacy preservation in SOA, which is related to the concern of the consumer that disclosed information can be misused by providers receiving it. Specifically, the framework is aimed at providing a solution to the areas of service description and discovery as the problem of privacy preservation in SOA encompasses several areas and thus demands solutions that include privacy enhancing mechanisms in the different areas of SOA.

The framework deals with the privacy preservation problems in the areas of service description and discovery, that is, it is not possible to describe how a provider deals with private information received from a consumer and discover a service that satisfies the privacy preferences of the consumer. The solution to these problems offered by the framework was designed in order to address the limitations identified in surveyed privacy frameworks [21], [38], [47], [4], [3], [2], [25], [30] for SOA proposed in the literature that deal with these issues. These limitations include limited policy model, privacy vocabulary as well as support for privacy policy specification and intersection. Furthermore, the frameworks have no service discovery integration as well as proper support for the inclusion of other QoS attributes and for the consideration of domain-specific issues.

With these limitations in mind, the proposed framework includes a model (Chapter 5) for policies with elements (Section 5.1), which enables the specification of policies that define different aspects of privacy preservation (components) of different information items (assertions) in different settings (alternatives). A format (Section 5.2) that considers the proposed policy elements is defined, which is used by consumers and providers for the specification of their preferences and practices as policies. These policies offer the base for the proposed framework as, in addition to improving service description, they are used in order to improve the process of service discovery. In order to use policies in the process of service discovery, a mechanism of policy intersection (Section 5.3) is defined, which indicates how privacy policies are matched. One of the main characteristics of the policy model is that it includes semantics by enabling the use of ontologies (Section 5.4) in the definition of privacy policies. The ontologies define the vocabulary used to create policies and are developed according to an approach that separates general (base ontology) and domain-specific (domain ontologies) privacy issues. This approach is employed to address the aspect that privacy preservation is a problem in several domains and some privacy preservation issues are common to different domains, but it is important to consider that each domain includes specific issues. In the policy model, policy assertions refer to ontological concepts. Thus, policies are created from concepts defined in privacy ontologies. This semantic information supports the matching between the policies of a consumer and provider.

The framework includes privacy-aware service discovery (Chapter 6), which enables the discovery of services that meet privacy preferences of consumers. In basic SOA, service description is restricted to functional characteristics of services. As a consequence, service discovery is based on functionality of services. Thus, the framework extends SOA in order to include privacy preservation characteristics of services in service description. This extension allows for service discovery that considers not only the functionality of the service but also the privacy characteristics of the service. Thus, in the proposed approach, providers and consumers describe their privacy preservation practices and preferences in policies, policy intersection enhances service discovery so that discovered services are from providers whose privacy practices match the privacy preferences of the consumer. The use of privacy policies for service discovery is accomplished by extending SOA with two new roles: mediator (Section 6.1) and privacy (Section 6.2). The privacy role is responsible for the publication and discovery of privacy policies. The mediator role mediates the interactions of service publication and discovery between the provider or consumer and the publication and discovery space, which includes the service registry and the privacy. The privacy role is responsible for privacy preservation and complements the registry. These two roles define a publication and discovery space in which they are responsible for the services of publication and discovery of services, where the registry is responsible for functional characteristics of services and the privacy, for privacy characteristics. The mediator is added to the architecture so that the publication and discovery space is transparent to the consumer and provider and support to additional QoS characteristics can be added by following the same approach used to deal with privacy.

An implementation and evaluation (Chapter 7) of the framework is presented. The implementation (Section 7.1) includes the mediator (Section 7.1.1) and privacy (Section 7.1.2) extensions. Among the different domains, health care is an example in which privacy preservation is particularly important, as health information is usually regarded as sensitive. Thus, the health care domain was chosen for the framework evaluation (Section 7.2). The evaluation involves the extension of the base ontology for the domain of health care (Section 7.2.1) and the creation of a health care scenario (Section 7.2.2), which were used to demonstrate cases (Section 7.2.3) in which consumers have their policies checked against the policies of providers to verify if the practices of a provider satisfy the preferences of a consumer.

An overview of the framework is shown in Figure 4.1.

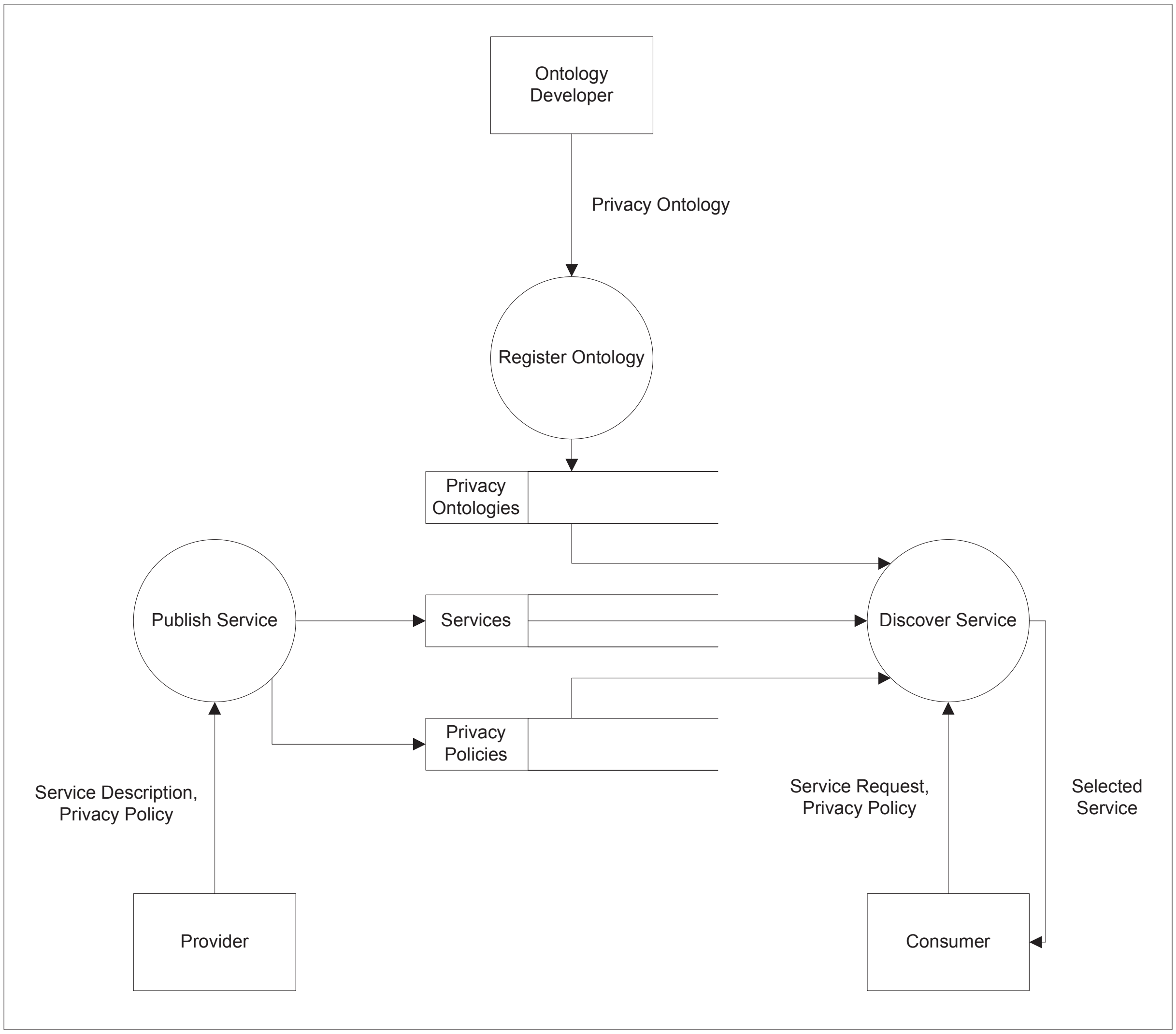

Figure 4.1. Privacy preservation framework.

As shown in Figure 4.1, the privacy preservation framework includes a model for semantic privacy policies and a process of privacy-aware service discovery through an extension to the basic architecture of SOA. The model for semantic privacy policies enables the description of privacy preservation practices of service providers and privacy preservation preferences of service consumers in policies. The policy model follows an approach in which general privacy preservation issues are represented by a base privacy ontology and domain-specific privacy issues are captured by privacy ontologies that extend the base ontology. Privacy-aware service discovery enables the discovery of services that meet privacy preferences of consumers. Privacy-aware service discovery uses the model for semantic privacy policies. At service discovery, privacy policies are intersected to select services from providers whose policies match the consumer's policy. Thus, the framework proposed in this thesis provides privacy preservation support for the areas of service description and discovery in SOA. The model for semantic policies enhances service description with privacy preservation practices and service request with privacy preservation preferences. The privacy policies complement basic service description and request that include information on service functionality and use. Privacy-aware service discovery integrates privacy-awareness in the processes of service publication and discovery in order to enable the publication of privacy practices and service discovery that considers privacy preferences. The process of privacy-aware service discovery is accomplished by extending basic SOA with new roles and activities that support the idea of different registry types, including registries for service descriptions and privacy policies.

4.2 Summary

This chapter gave an overview of the privacy preservation framework proposed in this thesis. The framework is further described in Chapters 5 and 6. Chapter 5 describes the model for semantic privacy policies for service description and Chapter 6 presents the extension to basic SOA for privacy-aware service discovery. Finally, the evaluation of the framework is presented in Chapter 7.

Chapter 5
Semantic Privacy Policies Model for Service Description

This chapter presents the framework's policy model. The privacy framework includes a policy model to enhance service descriptions and requests with privacy preservation properties of providers and requirements of consumers. In the policy model, policies specify privacy preferences and practices of consumers and providers. The subject of a policy can be a service request or a service as a consumer policy is associated with a service request and a provider policy with a service.

By investigating existing privacy frameworks (Chapter 3), some problems were identified regarding the policy model, mainly the limited privacy vocabulary. Thus, the model for policies proposed in this thesis organizes elements in a format that enables the use of ontological concepts. The policies improve service descriptions, which include functionality information, in the framework, and are used in order to improve the process of service discovery, through the use of a mechanism of policy intersection. In the policy model, the ontologies define the vocabulary used to create policies and are developed according to an approach that separates general and domain-specific privacy issues. This approach is employed as it is important to consider that each domain includes specific privacy issues.

The policy model is based on WS-Policy. WS-Policy is the standard for Web service policies and, thus, its format was used in order to make the privacy policy model interoperable. The main difference between the proposed privacy policy model and WS-Policy is that WS-Policy does not support the use of ontologies, whereas in the proposed framework, ontologies are used to define a privacy vocabulary whose concepts are used to specify policies.

5.1 Policy Elements

The policy model includes four elements: component, assertion, alternative and policy. Figure 5.1 shows an example of a privacy policy, which is going to be used to illustrate the policy elements.

```
01 Policy
02   ExactlyOne
03     All
04       Name
05         LegalRetention
06     All
07       Name
08         NoRetention
```

Figure 5.1. Example of privacy policy.

In Figure 5.1, Line 1 indicates a policy. Line 2 shows that the policy includes alternatives. The first alternative is defined from Line 3 and the second one from Line 6. Each alternative includes an assertion on the name information piece (Lines 4 and 7). Each assertion includes a component, which defines the retention period of the information piece (Lines 5 and 8). The elements of the policy model are described as follows:

- **Component and Assertion**

```
    All
      Name
        NoRetention
```

Figure 5.2. Example of component and assertion.

An assertion deals with a set of information pieces, which is its subject. An assertion includes components and each component restricts an aspect of the handling of the assertion's subject. Figure 5.2 includes an assertion and a component. The assertion's subject is the name information piece and the component restricts its retention.

Each assertion restricts the handling of a set of information pieces. This way consumers and providers can define assertions for a single information piece or a set with more than one information piece. Thus, by including components to an assertion according to their needs, consumers and providers can express different restrictions to information pieces in different settings and establish different privacy preservation levels based on what each consumer and provider consider as an acceptable practice.

Assertions are expressed using concepts defined in ontologies. These concepts define component types. They create a terminology for expressing policies and indicate general as well as domain-specific privacy semantics. Thus, assertions associated with different services and referring to the same concepts are interpreted similarly. A concept is referred to by an assertion and a component through its qualified name, including the Uniform Resource Identifier (URI) of the ontology that represents the namespace and its local identification. For readability, assertions are expressed using local identifications. In the examples used in this chapter, the policy components are from the base ontology (Section 5.4) and some components are used to enrich the examples and would have to be defined in domain ontologies. In Figure 5.2, the *Name* assertion subject and the *NoRetention* component are defined in a domain and the base ontologies, respectively.

- **Alternative**

```
ExactlyOne
  All
    Name
      LegalRetention
  All
    Name
      NoRetention
```

Figure 5.3. Example of alternative.

Assertions are grouped in collections called alternatives. An alternative is an ordered assertion collection. It indicates the preferences or practices represented by its assertions and its privacy preservation level depends on the assertions' level. Assertions are processed in the order in which they appear in the alternative. Figure 5.3 has two alternatives with an assertion each.

This element is included in the policy model to offer providers and consumers the possibility to specify alternative settings of privacy preservation practices and preferences. This way the likelihood to successfully intersect policies when discovering services is higher.

- **Policy**

```
Policy
  ExactlyOne
    All
      Name
        LegalRetention
    All
      Name
        NoRetention
```

Figure 5.4. Example of policy.

A policy is created by grouping alternatives. It is an ordered collection of alternatives. A policy with more than one alternative indicates that there are choices of preferences or practices. Alternatives are processed in the order in which they appear in the policy. While processing a policy, the first alternative is checked, then, if needed, the second one and so on. Figure 5.4 shows a policy with two alternatives.

Policies restrict interactions between consumers and providers. Provider policies specify practices of providers and consumer policies preferred practices or preferences of consumers. Policies apply to information pieces disclosed by consumers to providers in order to use their services. Figure 5.4 can represent a consumer or provider policy. Thus, it can define a consumer's preferences or provider's practices regarding the retention of the name information piece.

A provider exposes a policy describing conditions under which it performs its activities in the context of a service. A behavior that reflects those conditions is presented by the provider in order to satisfy the policy. A consumer can use the policy exposed by the provider in order to decide whether or not to use the service. It can choose any alternative in the policy, as each one represents valid conditions under which the service can be used. As each alternative represents an alternative set of conditions, the consumer can choose only one for each interaction with the service. A provider supports an assertion if it performs the practice represented by it. An alternative is supported by a provider if all of its assertions are supported by it. A provider supports a policy if it supports all the alternatives of the policy. Thus, it must be able to operate under the different conditions represented by the alternatives in a policy so that it can support the policy. According to Figure 5.4, the provider has to be able to provide the service with legal retention or no retention of the name information piece in order to support the policy. In the case of the consumer, the policy indicates that the consumer accepts services from providers with no retention or legal retention practices.

5.2 Policy Format

This section describes the policy format, which defines a standard structure for the specification of policies. The items of the format as well as rules to map additional items to the format are described in this section. Policies follow the format shown in Figure 5.5.

```
Policy Name=“” Id=“”
  ExactlyOne
    All
      Assertion
```

Figure 5.5. Policy format.

The items of the policy format are described as follows:

- **Policy**: a policy.
- **Name**: the identity of the policy in the form of an absolute Internationalized Resource Identifier (IRI). The name of a policy is referred to by a service description or request in order to associate them.
- **Id**: the policy’s identity in the form of an identifier within its enclosing document. An IRI-reference is composed using the identifier of a policy and the IRI of the enclosing document in order to refer to the policy externally.
- **ExactlyOne**: the collection of all the alternatives of the policy. This item indicates that only one alternative can be selected at a time.
- **All**: an alternative. This item groups the assertions of an alternative and indicates that all assertions are valid when the alternative is selected.
- **Assertion**: a preference in the case of a consumer policy or a practice in the case of a provider policy.

An example of a policy named http://www.privpol.com/Policy1 in the policy format is shown in Figure 5.6. The assertions are illustrative and their definitions are not necessary at this point as the focus is on the description of the format.

```
01  Policy Name="http://www.privpol.com/Policy1"
02    ExactlyOne
03      All
04        Name, Contact
05      All
06        Name
```

Figure 5.6. Formatted policy.

This example includes two alternatives. The first one states that name and contact information is collected by the provider (Lines 03-04), whereas name information only is collected for the second alternative (Lines 05-06).

A formatted policy lists all the alternatives, whereas each alternative lists all the assertions. As this format can lead to extensive policies, two constructs were added so that it is possible to express policies in a more compact way, which can then be converted to the basic format in order to keep interoperability. These constructs, optional assertion and policy referencing, are described as follows.

- **Optional Assertion**

A compact construct for the policy format is the optional assertion. The result of including an optional assertion in a provider policy is that the formatted policy is going to have two alternatives, one with the assertion and another without it. The example in Figure 5.7 includes an optional assertion.

```
01 Policy
02   ExactlyOne
03     All
04       Name
05         Optional: BusinessRecipient
```

Figure 5.7. Policy with optional assertion.

The example in Figure 5.7 is equivalent to the example of formatted policy in Figure 5.8.

```
01 Policy
02   ExactlyOne
03     All
04       Name
05         BusinessRecipient
06     All
07       Name
08         NoRecipient
```

Figure 5.8. Formatted policy with optional assertion.

In Figure 5.7, the optional assertion in Line 5 indicates that there are two possible alternatives, one with the assertion and another one without it or with an assertion that nullifies the original one. The first alternative is shown in Lines 3-5 in Figure 5.8 and states that name information is going to be disclosed to third-party businesses. The second alternative, in Lines 6-8, states that the information is not going to be disclosed to third parties.

In consumer policies, an optional assertion indicates that the assertion is checked only if all mandatory ones have been matched. If an optional assertion is not matched, this does not prevent a policy from being determined as compatible. An example of a consumer policy with an optional assertion is shown in Figure 5.9. It shows that name information cannot be retained (Line 5) and should not be disclosed to third parties (Line 6).

```
01  Policy
02    ExactlyOne
03      All
04        Name
05          NoRetention
06          Optional: NoRecipient
```

Figure 5.9. Consumer policy with optional assertion.

- **Policy Referencing**

A policy can include another one through the mechanism of policy referencing. Thus, this mechanism supports inter-policy assertion sharing. The policy reference can be an IRI or IRI-reference and is placed at the place where an assertion is included in a policy. When a policy references another one, the portion of the referenced policy wrapped by the item that represents a policy replaces the item that represents a reference in the referencing policy.

Then, the new content is wrapped by an item that represents an alternative in the referencing policy.

An example of policy referencing is shown in Figure 5.10. Two policies are defined in the same document. The first one is identified as Policy1 (Line 1). The second one, Policy2 (Line 4), references Policy1 through an IRI-reference so that it includes its content (Line 5).

```
01  Policy Id=“Policy1”
02    All
03
04  Policy Id=“Policy2”
05    PolicyReference IRI=“#Policy1”
06    All
```

Figure 5.10. Policy with IRI-reference.

Another example is shown in Figure 5.11. In this example, the two policies are defined in separate documents. The first one is named http://www.privpol.com/Policy1 (Line 1). The second policy, named http://www.privpol.com/Policy2 (Line 4), references the first one through an IRI (Line 5).

```
01  Policy Name=“http://www.privpol.com/Policy1”
02    All
03
04  Policy Name=“http://www.privpol.com/Policy2”
05    PolicyReference IRI=“http://www.privpol.com/Policy1”
06    All
```

Figure 5.11. Policy with IRI.

5.3 Policy Intersection

Intersection is matching between policies, which identifies compatibility between two policies in order to verify if their owners can interact with each other. The input of the process of policy intersection is a consumer and provider policy. The output is a policy including a compatible alternative from the provider policy or empty if the policies are incompatible.

Two policies are compatible if at least one consumer alternative is compatible with at least one provider alternative. Two alternatives are compatible if each consumer mandatory assertion is compatible with a provider assertion as well as each provider assertion is compatible with a consumer mandatory assertion. If there is more than one compatible provider policy, the compatibility of the optional assertions is checked in order to rank the compatible providers. The provider policy with more compatible optional assertions is the highest one on the ranking. The result of the intersection process between two compatible alternatives is the provider alternative or empty otherwise. Two assertions are compatible according to matching rules defined for ontologies.

In the case of compatible policies, the alternative in the resulting policy comes from the provider policy. The selected provider is the one at the highest position on the ranking, if more than one compatible provider is identified. The selected provider has to support all practices indicated by the result of the process of policy intersection. If the provider policy has different assertions in other alternatives, the provider cannot support the practices represented by them.

An example of a policy intersection is shown as follows. Figure 5.12 and Figure 5.13 present a consumer and provider policy, respectively. These policies are the intersection input.

```
01 Policy
02   ExactlyOne
03     All
04       Name
05         NoRecipient
06         LegalRetention
07     All
08       Name
09         AnyRecipient
10         NoRetention
```

Figure 5.12. Consumer policy.

Figure 5.12 includes two alternatives. The first one (Lines 3-6) indicates that *Name* information can be retained as required by law (*LegalRetention*) but the information cannot be disclosed to third parties (*NoRecipient*). The second alternative (Lines 7-10) indicates that *Name* information can be disclosed to any third parties (*AnyRecipient*) but it cannot be retained (*NoRetention*).

```
01 Policy
02   ExactlyOne
03     All
04       Name
05         BusinessRecipient
06         LegalRetention
07     All
08       Name
09         BusinessRecipient
10         NoRetention
```

Figure 5.13. Compatible provider policy.

Figure 5.13 includes two alternatives. The first alternative (Lines 3-6) indicates that *Name* information is retained as required by law (*LegalRetention*) and disclosed to third-party businesses (*BusinessRecipient*). The second one (Lines 7-10) indicates that *Name* information is disclosed to third-party businesses (*BusinessRecipient*) and not retained (*NoRetention*).

The first consumer alternative (Figure 5.12) is not supported by any provider alternative (Figure 5.13) as it requires no disclosure (*NoRecipient*) and both provider alternatives disclose *Name* information (*BusinessRecipient*). The second consumer alternative is not supported by the first provider alternative as it requires no retention (*NoRetention*) and the first provider alternative retains *Name* information (*LegalRetention*). The intersection result includes the second provider alternative as it supports the second consumer alternative (*NoRetention*). It is shown in Figure 5.14.

```
Policy
  ExactlyOne
    All
      Name
        BusinessRecipient
        NoRetention
```

Figure 5.14. Policy intersection result.

For another example of intersection result, Figure 5.15 presents a provider policy, which is the input for the intersection process along with the policy in Figure 5.12.

```
01 Policy
02   ExactlyOne
03     All
04       Name
05         BusinessRecipient
06         LegalRetention
07     All
08       Name
09         GovernmentRecipient
10         LegalRetention
```

Figure 5.15. Incompatible provider policy.

Figure 5.15 includes two alternatives. The first one (Lines 3-6) indicates that *Name* information is retained as required by law (*LegalRetention*) and disclosed to third-party businesses (*BusinessRecipient*). The second alternative (Lines 7-10) indicates that *Name* information is retained as required by law (*LegalRetention*) and disclosed to governmental third parties (*GovernmentRecipient*). None of the provider alternatives supports any of the consumer alternatives in Figure 5.12 as one of the consumer alternatives requires no retention of *Name* information (*NoRetention*) and the other one requires no disclosure (*NoRecipient*), but both provider alternatives disclose (*BusinessRecipient* and *GovernmentRecipient*) and retain (*LegalRetention*) *Name* information. The intersection result between the policies is empty.

5.4 Base Ontology

The semantic approach that supports the policy model includes a base and domain-specific ontologies. The base ontology includes general privacy concepts. Domain-specific ontologies extend the base one and include domain-specific privacy concepts. This section presents the base ontology. A domain-specific ontology is presented in Chapter 7. An overview of the base ontology is shown in Figure 5.16.

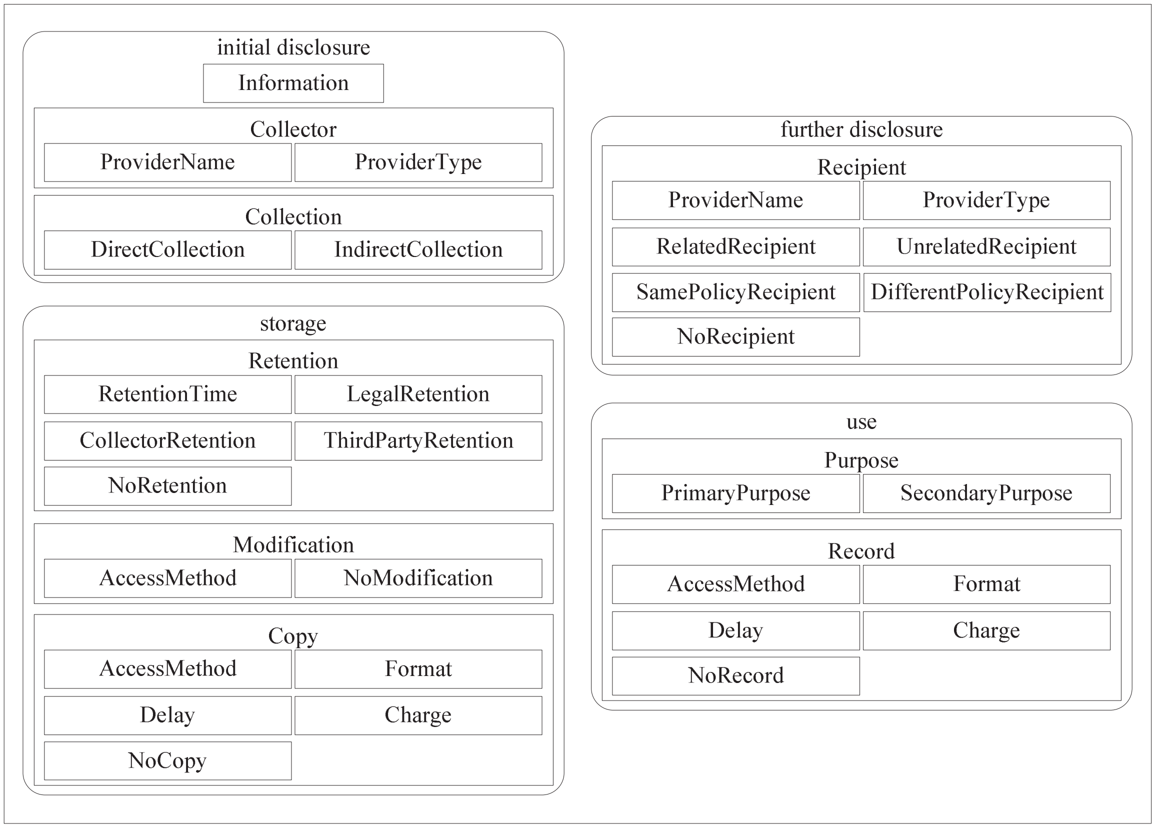

Figure 5.16. Base ontology.

The base concepts are described under types of information activities to which they relate. Four activity types can be identified in privacy regulations [40], [29], [14], [9]: initial disclosure, further disclosure, storage and use.

5.4.1 Initial Disclosure

In this information activity, a consumer discloses information to a provider. It is important to give the consumer the ability to control the initial disclosure of information. Firstly, it is necessary to ensure that the consumer is aware of the initial disclosure. It is also important to ensure that it is aware of the implications of the disclosure so that it can balance these privacy implications and the benefits it is going to get from the disclosure. Three concepts were identified in this activity type: *Information*, *Collector* and *Collection*.

- **Information**

This concept represents the type of the information piece to be disclosed by the consumer (in a consumer policy) or collected by the provider (in a provider policy).

In a consumer and provider assertion, the *Information* component includes one or more information types. Figure 5.17 shows the basic structure of *Information* with examples of information types.

```
Information
  Identifier,
  Name,
  Contact,
  ServiceUse,
  Finance
```

Figure 5.17. Information.

There is a match between a consumer and provider assertion if the condition for each component is true. A provider *Information* matches the one of a consumer if each of its information types is the same class or a subclass of a consumer information type.

- **Collector**

This concept represents the provider that is allowed by the consumer to collect its information (in a consumer policy) and the provider that is going to collect the consumer's information (in a provider policy). *Collector* includes the following concepts:

  - **ProviderName**: identifies the providers allowed by the consumer (in a consumer policy) and the one that is going to collect the information (in a provider policy).
  - **ProviderType**: indicates the types of the providers allowed by the consumer (in a consumer policy) and the type of the one that is going to collect the information (in a provider policy).

In a consumer assertion, the *Collector* component includes one or more provider names and/or types. An assertion without *Collector*, *ProviderName* or *ProviderType* indicates that any provider, any provider of a given type (*ProviderType*) or only the given providers (*ProviderName*) are allowed. In a provider assertion, *Collector* includes one provider name and can include one type. An assertion without the *ProviderType* part indicates that the provider does not identify its type. Figure 5.18 shows the basic structure of *Collector* with examples of provider types.

```
Collector
  ProviderName=“”
  ProviderType
    Business
    Government
    NGO
```

Figure 5.18. Collector.

A provider and consumer *Collector* match if the *ProviderName* and *ProviderType* parts match. *ProviderName* or *ProviderType* matches if the provider name or type in the provider assertion is in the name or type set in the consumer assertion; or the consumer does not specify any *ProviderName* and *ProviderType*.

- **Collection**

This concept represents the information collection means, that is, the means the provider employs to collect information from the consumer, allowed by the consumer (in a consumer policy) and used by the provider (in a provider policy). Types of collection means include:

  - **DirectCollection**: indicates that the information can be collected directly (in a consumer policy) and is going to be collected directly (in a provider policy).
  - **IndirectCollection**: indicates that the information can be collected indirectly; for example, using information provided by the consumer to obtain publicly-

available information (in a consumer policy), and is going to be collected indirectly (in a provider policy).

In a consumer and provider assertion, *Collection* includes zero or more direct and/or indirect types of information to be collected. A consumer assertion without *Collection* indicates that any collection means is allowed, whereas an empty *DirectCollection* or *IndirectCollection* indicates that any direct or indirect means is allowed. A provider assertion without *Collection* indicates that any means can be used, whereas an empty *DirectCollection* or *IndirectCollection* indicates that any direct or indirect means can be used. Figure 5.19 shows the basic structure of *Collection* with examples of indirect collection types.

```
Collection
  DirectCollection
  IndirectCollection
    DataCapture
    DataDerivation
```

Figure 5.19. Collection.

A provider and consumer *Collection* match if the *DirectCollection* and *IndirectCollection* parts match. *DirectCollection* or *IndirectCollection* matches if each direct or indirect type in the provider assertion is the same class or a subclass of a direct or indirect type in the consumer assertion; or the direct or indirect type set in the consumer assertion is empty; or there are no *DirectCollection* and *IndirectCollection* in the consumer assertion.

5.4.2 Further Disclosure

A further information disclosure occurs between two providers. In this type of information activity, the provider that collected the information from the consumer shares it with another one. Different indirectness levels can occur, as the third-party provider can share the information received from its collector with another provider. Thus, a provider receives the information of the consumer from the provider with which the consumer directly interacted or, in additional indirectness levels, it receives the information from a provider that is not the collector. The *Recipient* concept was identified in this activity type.

- **Recipient**

This concept represents the recipient of a further disclosure of information allowed by the consumer (in a consumer policy) and the third parties that are going to receive from the collector the information disclosed by the consumer (in a provider policy). *Recipient* includes the following concepts:

  - **ProviderName**: identifies the recipients of further information disclosures allowed by the consumer (in a consumer policy) and the third parties that are going to be recipients of further disclosures by the provider (in a provider policy).
  - **ProviderType**: indicates the types of the recipients of further disclosures allowed by the consumer (in a consumer policy) and the types of the third

parties that are going to be recipients of further information disclosures by the provider (in a provider policy).

- **RelatedRecipient**: indicates that the recipients must behave on behalf of the collector (in a consumer policy) and are going to do so (in a provider policy).
- **UnrelatedRecipient**: indicates that the recipients can behave on their own behalf (in a consumer policy) and are going to do so (in a provider policy).
- **SamePolicyRecipient**: indicates that the recipients must perform the same practices as the collector regarding the disclosed information (in a consumer policy) and are going to do so (in a provider policy).
- **DifferentPolicyRecipient**: indicates that the recipients can perform different practices from the collector regarding the disclosed information (in a consumer policy) and are going to do so (in a provider policy).
- **NoRecipient**: indicates that no recipient is allowed by the consumer (in a consumer policy) and the collector does not disclose the information to any third party (in a provider policy).

In a consumer and provider assertion, *Recipient* includes one or more provider names and/or types. In addition, it can include a *RelatedRecipient* or *UnrelatedRecipient* as well as a *SamePolicyRecipient* or *DifferentPolicyRecipient*. Instead, it can include a *NoRecipient*. A consumer assertion without *Recipient*, *ProviderName* or *ProviderType* indicates that any recipient, any recipient of a given type or only the given recipients are allowed respectively. No *RelatedRecipient* and *UnrelatedRecipient* indicate that the consumer does not impose any restriction on the relationship between collector and recipient, whereas no *SamePolicyRecipient* and *DifferentPolicyRecipient* indicate that it does not impose any restriction on the recipient policy. Figure 5.20 shows the basic structure of *Recipient*.

```
Recipient
  ProviderName=“”
  ProviderType
  RelatedRecipient
  UnrelatedRecipient
  SamePolicyRecipient
  DifferentPolicyRecipient
  NoRecipient
```

Figure 5.20. Recipient.

A provider and consumer *Recipient* match if *ProviderName*, *ProviderType*, *RelatedRecipient*, *UnrelatedRecipient*, *SamePolicyRecipient*, *DifferentPolicyRecipient* and *NoRecipient* match. *ProviderName* or *ProviderType* matches if each name or type in the provider assertion is in the name or type set in the consumer assertion; or there are no *ProviderName* or *ProviderType* and *NoRecipient* in the consumer assertion. *RelatedRecipient* or *UnrelatedRecipient* matches if there is *RelatedRecipient* or *UnrelatedRecipient* in the consumer and provider assertions; or there are no *RelatedRecipient*, *UnrelatedRecipient* and *NoRecipient* in the consumer assertion. The logic of *SamePolicyRecipient* and *DifferentPolicyRecipient* are similar. *NoRecipient* matches if there is *NoRecipient* in the consumer and provider assertions; or there is no *NoRecipient* in the consumer assertion.

5.4.3 Storage

Two types of information storage can occur. In the first one, information is stored beyond service completion. The second type refers to information that is stored only for the time period of the transaction. Another dimension that can classify information storage is who is going to store it. Information can be stored by the provider with which the consumer interacted or by a third party provider. Three concepts were identified in this type of information activity: *Retention*, *Modification* and *Copy*.

- **Retention**

This concept represents the time period of the information retention and the provider responsible for it. *Retention* includes the following concepts:

  - **RetentionTime**: indicates the maximum time period the information can (in a consumer policy) and is going to be retained (in a provider policy).
  - **LegalRetention**: indicates that the information can (in a consumer policy) and is going to be retained as required by law (in a provider policy).
  - **CollectorRetention**: indicates that the information must (in a consumer policy) and is going to be retained by the collector (in a provider policy).
  - **ThirdPartyRetention**: indicates that the information must (in a consumer policy) and is going to be retained by a third party (in a provider policy).

- **NoRetention**: indicates that the information cannot (in a consumer policy) and is not going to be retained beyond service completion (in a provider policy).

In a consumer and provider assertion, *Retention* includes a retention time or *LegalRetention*. In addition, it can include *CollectorRetention* or *ThirdPartyRetention*. Instead, it can include *NoRetention*. A consumer assertion without *Retention*, *RetentionTime* or *CollectorRetention* and *ThirdPartyRetention* indicates that the information can be retained indefinitely by any provider, the given provider or for the given time period by any provider respectively. Figure 5.21 shows the basic structure of *Retention*.

```
Retention
  RetentionTime="”
  LegalRetention
  CollectorRetention
  ThirdPartyRetention
  NoRetention
```

Figure 5.21. Retention.

Provider and consumer *Retention* match if *RetentionTime*, *LegalRetention*, *CollectorRetention*, *ThirdPartyRetention* and *NoRetention* match. *RetentionTime* matches if the retention time in the provider assertion is less than or equal to the consumer one; or there are no *RetentionTime*, *LegalRetention* and *NoRetention* in the consumer assertion. *LegalRetention* matches if there is one in the consumer and provider assertions; or there are no consumer *LegalRetention*, *RetentionTime* and *NoRetention*. *CollectorRetention* or *ThirdPartyRetention* matches if there is one in the consumer and provider assertions; or there are no consumer *CollectorRetention*, *ThirdPartyRetention* and *NoRetention*. *NoRetention* matches if there is consumer and provider *NoRetention*; or there is no consumer one.

- **Modification**

This concept represents the capability of the consumer to request to the provider the modification of the retained information. *Modification* includes the following concepts:

  - **AccessMethod**: identifies the means required by the consumer (in a consumer policy) and supported by the provider to request the retained information modification (in a provider policy).
  - **NoModification**: indicates that the consumer does not require (in a consumer policy) and the provider does not allow for modification request (in a provider policy).

In a consumer and provider assertion, *Modification* includes zero or more access method types or *NoModification*. A consumer assertion without *Modification* or with an empty *AccessMethod* indicates that it does not require modification or any access method type is allowed. Figure 5.22 shows the basic structure of *Modification* with examples of access method types.

```
Modification
  AccessMethod
    EService
    EMail
    Telephone
    Fax
    Mail
  NoModification
```

Figure 5.22. Modification.

Provider and consumer *Modification* match if *AccessMethod* and *NoModification* match. *AccessMethod* matches if each access method type in the provider assertion is the same class or a subclass of an access method type in the consumer assertion; or the set of access method types in the consumer assertion is empty; or there are no consumer *AccessMethod* and *NoModification*. *NoModification* matches if there is one in the consumer and provider assertions; or there is no consumer *NoModification* so that the consumer does not care about this aspect of privacy preservation.

- **Copy**

This concept represents the consumer's capability to request a copy of the retained information to the provider. *Copy* includes the following concepts:

  - **AccessMethod**: identifies the means required by the consumer (in a consumer policy) and supported by the provider to request retained information copy (in a provider policy).

- **Format**: identifies the format of the copy required by the consumer (in a consumer policy) and supported by the provider (in a provider policy).
- **Delay**: identifies the maximum time period the consumer is willing to wait for the receipt of the requested copy (in a consumer policy) and the delay the provider demands to make it available (in a provider policy).
- **Charge**: identifies the maximum charge the consumer is willing to pay for the receipt of the requested copy (in a consumer policy) and the charge the provider demands to make it available (in a provider policy).
- **NoCopy**: indicates that the consumer does not require (in a consumer policy) or the provider does not allow for copy request (in a provider policy).

In a consumer and provider assertion, *Copy* includes zero or more access method and format types. In addition, it can include a delay and charge. Instead, it can include *NoCopy*. A consumer assertion without *Copy*, delay, charge, with an empty *AccessMethod* or *Format* indicates that it does not require copy, any specific delay, charge, any access method or format type is allowed respectively. Figure 5.23 shows the basic structure of *Copy*.

```
Copy
  AccessMethod
  Format
  Delay
  Charge
  NoCopy
```

Figure 5.23. Copy.

Provider and consumer *Copy* match if *AccessMethod*, *Format*, *Delay*, *Charge* and *NoCopy* match. *AccessMethod* or *Format* matches if each access method or format type in the provider assertion is the same class or a subclass of an access method or format type in the consumer assertion; or the access method or format type set in the consumer assertion is empty; or there are no *AccessMethod* or *Format* and *NoCopy* in the consumer assertion. *Delay* or *Charge* matches if the provider delay or charge is less than or equal to the consumer delay or charge; or there are no consumer *Delay* or *Charge* and *NoCopy*. *NoCopy* matches if there is *NoCopy* in the consumer and provider assertions; or there is no consumer *NoCopy* so that the consumer does not care about this aspect of privacy preservation.

5.4.4 Use

Two types of information use can occur. The first one includes the uses that are necessary for accomplishing the service, while the second one includes secondary uses. Another classification dimension for use is the provider that performs it. Information can be used by the provider with which the consumer directly interacted or third parties to which the collector disclosed it. Two concepts were identified in this activity type: *Purpose* and *Record*.

- **Purpose**

This concept represents the purposes for information collection allowed by the consumer (in a consumer policy) and the purposes for which the provider is going to collect the information (in a provider policy). *Purpose* includes the following concepts:

    - **PrimaryPurpose**: indicates that the collected information can (in a consumer policy) and is going to be used for service completion only (in a provider policy).
    - **SecondaryPurpose**: indicates that the collected information can (in a consumer policy) and is going to be used for secondary purposes (in a provider policy).

In a consumer and provider assertion, *Purpose* includes *PrimaryPurpose* and/or zero or more secondary purpose types. A consumer assertion without *Purpose* or with an empty *SecondaryPurpose* indicates that the collected information can be used for any purpose type or secondary purpose type. Figure 5.24 shows the basic structure of *Purpose* with examples of secondary purpose types.

```
01 Purpose
02   PrimaryPurpose
03   SecondaryPurpose
04     AdministrationPurpose
05     MaintenancePurpose
06     CustomizationPurpose
07     ProfilePurpose
08     MarketingPurpose
09     LegalPurpose
```

Figure 5.24. Purpose.

Provider and consumer *Purpose* match if *PrimaryPurpose* and *SecondaryPurpose* match. *PrimaryPurpose* matches if there are consumer and provider *PrimaryPurpose*; or there are no *PrimaryPurpose* and *SecondaryPurpose* in the consumer assertion. *SecondaryPurpose* matches if each secondary purpose type in the provider assertion is the same class or a subclass of a secondary purpose type in the consumer assertion; or the secondary purpose type set in the consumer assertion is empty; or there are no consumer *SecondaryPurpose* and *PrimaryPurpose*.

- **Record**

This concept represents the capability of the service consumer to request to the service provider a record of the use of the collected information. *Record* includes the following concepts:

- **AccessMethod**: identifies the means required by the consumer (in a consumer policy) and supported by the provider to record request (in a provider policy).
- **Format**: identifies the record format required by the consumer (in a consumer policy) and supported by the provider (in a provider policy).
- **Delay**: identifies the maximum time period the consumer is willing to wait for the receipt of the requested record (in a consumer policy) and the delay the provider demands to make it available (in a provider policy).
- **Charge**: identifies the maximum charge the consumer is willing to pay for the receipt of the requested record (in a consumer policy) and the charge the provider demands to make it available to the consumer (in a provider policy).
- **NoRecord**: indicates that the consumer does not require (in a consumer policy) and the provider does not allow for record request (in a provider policy).

In a consumer and provider assertion, *Record* includes zero or more access method and format types. In addition, it can include a delay and charge. Instead, it can include *NoRecord*. A consumer assertion without *Record*, delay, charge, with an empty *AccessMethod* or *Format* indicates that it does not require record request, any specific delay, charge, any access method or format type is allowed respectively. Figure 5.25 shows the basic structure of *Record*.

```
Record
  AccessMethod
  Format
  Delay
  Charge
  NoRecord
```

Figure 5.25. Record.

Service provider and consumer *Record* match if *AccessMethod*, *Format*, *Delay*, *Charge* and *NoRecord* match. *AccessMethod* or *Format* matches if each access method or format type in the service provider assertion is the same class or a subclass of an access method or format type in the consumer assertion; or the access method or format type set in the service consumer assertion is empty; or there are no service consumer *AccessMethod* or *Format* and *NoRecord*. *Delay* or *Charge* matches if the service provider delay or charge is less than or equal to the service consumer delay or charge; or there are no consumer *Delay* or *Charge* and *NoRecord*. *NoRecord* matches if there is *NoRecord* in the consumer and provider assertions; or there is no consumer *NoRecord* so that the service consumer does not care about this aspect of privacy preservation.

5.5 Summary

This chapter presented the policy model of the proposed framework. It described the policy format and base ontology. The policy model described in this chapter supports the process of privacy-aware service discovery presented in Chapter 6.

Chapter 6

Privacy-aware Service Discovery

Another problem identified after investigating existing frameworks in Chapter 3 was the use of privacy policies for enhancing service discovery, mainly the limited integration in the service discovery mechanism. Thus, the proposed framework includes a privacy-aware mechanism of service discovery. In the framework, providers and consumers use policies to describe their privacy practices and preferences. Then, policy intersection is used to enhance service discovery by matching provider and consumer policies in order to discover providers with practices that are compatible to the preferences of the consumers. Service-Oriented Architecture (SOA) is extended, in the framework, with the privacy and mediator roles. A publication and discovery space is defined, which includes a new role, named privacy, in addition to the basic role of registry. The services in the publication and discovery space are responsible for the publication and discovery of services. Whereas the registry service is responsible for functional characteristics of services, the privacy service is responsible for privacy characteristics. The second new role, the mediator, is added to make the publication and discovery space transparent to the consumers and providers as well as support additional

Quality of Service (QoS) characteristics. The mediator service facilitates the interactions of the providers and consumers with the services in the publication and discovery space.

The provider uses the privacy extension to SOA by sending its policy together with the service description to the mediator. In the case of the consumer, the extension is used by sending to the mediator its policy together with the service request. The mediator can then be added to SOA and interacted with the same way the registry is used in traditional SOA, by selecting a service in an Enterprise Service Bus (ESB) and using an Application Programming Interface (API), for example. If consumers and providers do not want to use the privacy feature, then they can still interact similarly to how they do so in traditional SOA.

This chapter discusses privacy-aware service discovery by presenting the extensions to SOA roles and interactions. The process of privacy-aware service discovery uses the policy model described in Chapter 5. It allows for consumers to have their privacy preferences considered when looking for services. In order to enable the process, two new roles were included in SOA: mediator and privacy. As with the service registry, these roles should be played by trusted third parties to ensure that their activities are unbiased. SOA extended with these roles is shown in Figure 6.1.

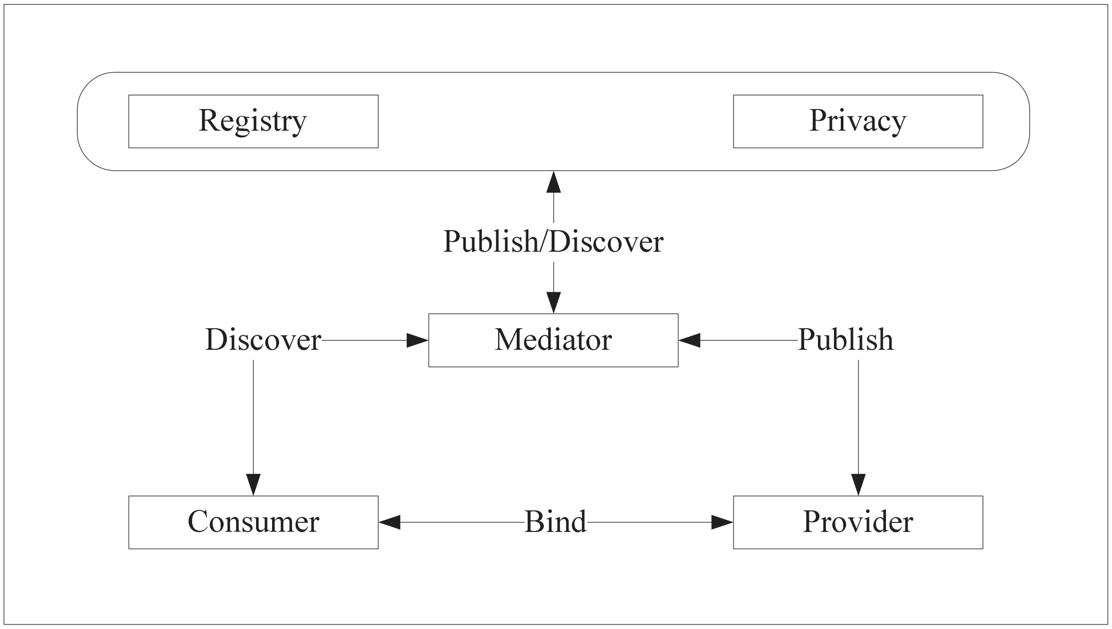

Figure 6.1. SOA new roles.

The new roles (mediator and privacy) and their interactions with the basic ones (consumer, provider and registry), shown in Figure 6.1, are presented as follows.

6.1 Mediator

The mediator service is included in SOA in order to facilitate the interactions between the provider or consumer and the publication and discovery services, including registry and privacy services, by making these services transparent to consumers and providers. Together with the service registry and the privacy, the mediator is responsible for service publication and discovery. It uses them to execute these activities. The mediator has a registry of publication and discovery services, which is used to register addresses of service registries and privacies. Service registry and privacy providers are responsible for registering their services in the registry of the mediator. Based on the message received from the provider or consumer, the mediator decides which publication or discovery services are needed in order to execute the requested activity. It retrieves the addresses of the service registry and privacy from its registry so that it can use them.

The activities of registration and deregistration of publication and discovery services performed by the mediator are shown in Figure 6.2.

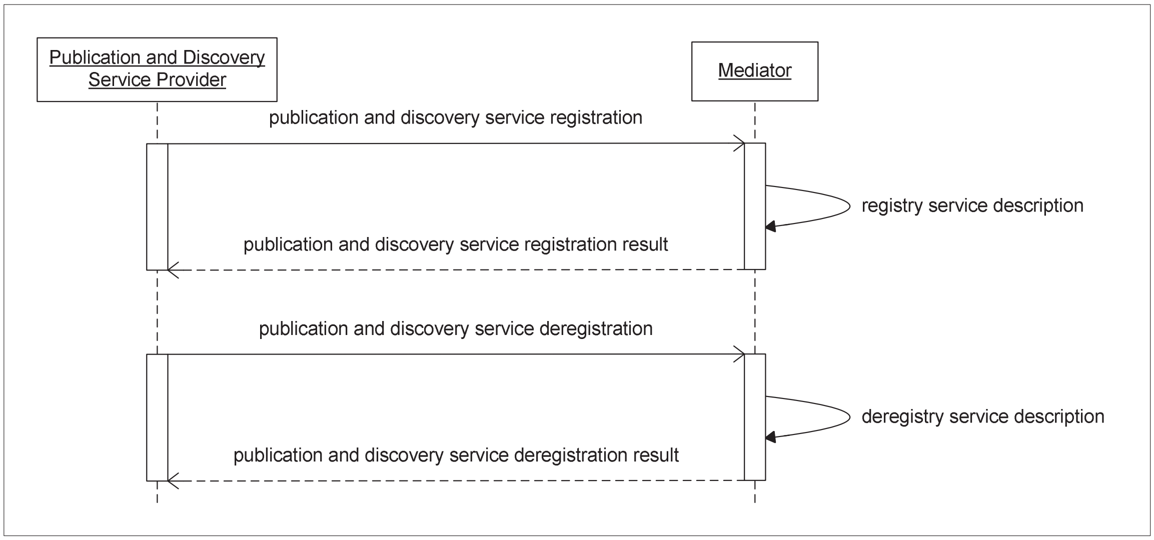
Publication and Discovery Service Provider
Mediator
publication and discovery service registration
registry service description
publication and discovery service registration result
publication and discovery service deregistration
deregistry service description
publication and discovery service deregistration result

Figure 6.2. Registration and deregistration of publication and discovery services.

At publication and discovery service registration/deregistration, the mediator receives a registration/deregistration message from the provider including a description of the service. Then, the service description is registered/deregistered. Finally, it sends a result message to the provider.

The tasks under the responsibility of the mediator at service publication and unpublication are shown in Figure 6.3.

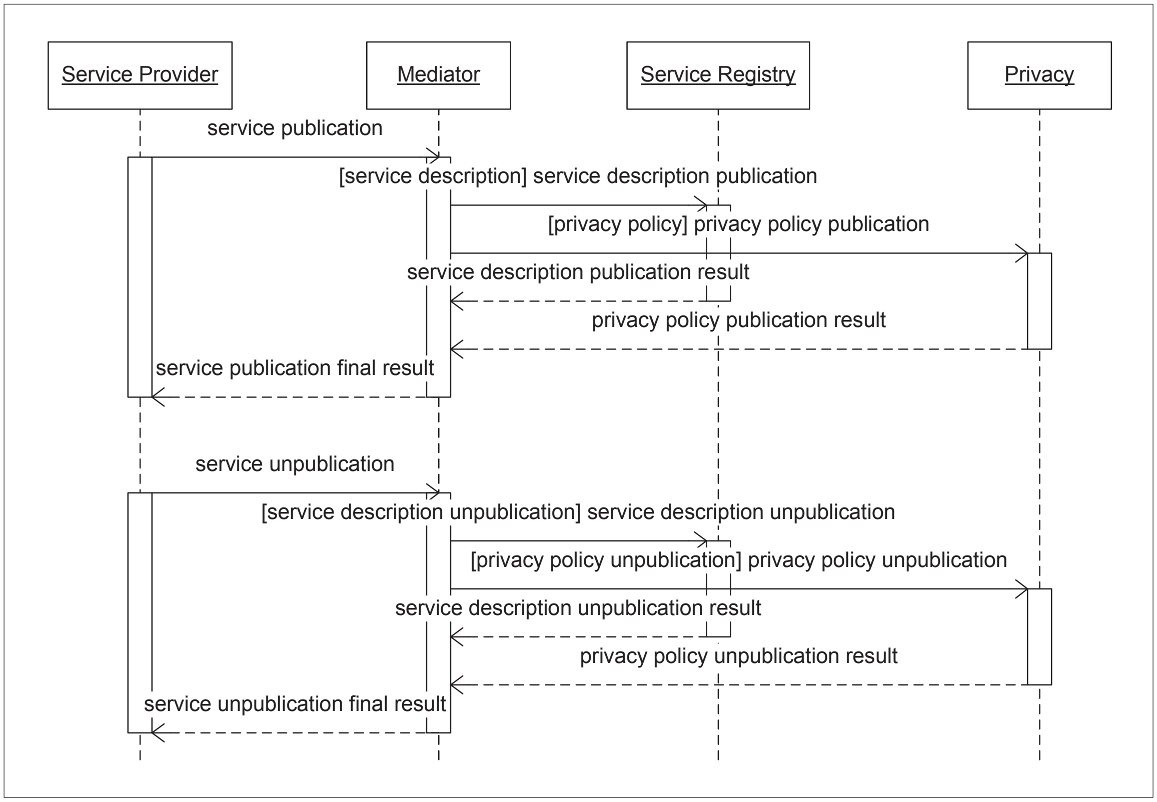
Service Provider
Mediator
Service Registry
Privacy
service publication
[service description] service description publication
[privacy policy] privacy policy publication
service description publication result
privacy policy publication result
service publication final result
service unpublication
[service description unpublication] service description unpublication
[privacy policy unpublication] privacy policy unpublication
service description unpublication result
privacy policy unpublication result
service unpublication final result

Figure 6.3. Mediator tasks at service publication and unpublication.

At service publication/unpublication, the mediator receives a publication/unpublication message from the provider. It sends a service description message to the service registry and a privacy policy message to the privacy if the publication/unpublication message includes a service description and privacy policy. Then, the mediator receives a service description and privacy policy result message from the service registry and privacy. Finally, it sends a final result message to the provider.

The tasks under the responsibility of the mediator at service discovery are shown in Figure 6.4.

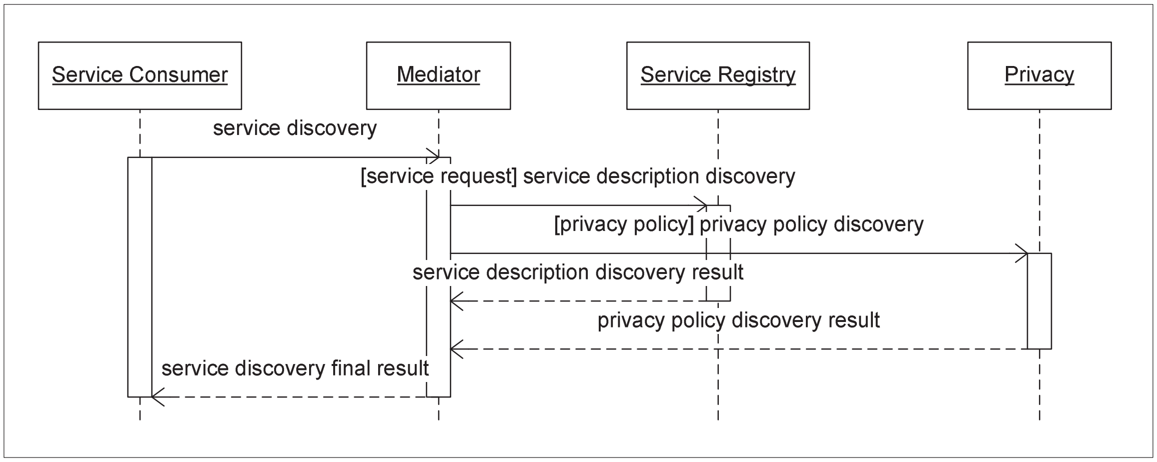
Service Consumer
Mediator
Service Registry
Privacy
service discovery
[service request] service description discovery
[privacy policy] privacy policy discovery
service description discovery result
privacy policy discovery result
service discovery final result

Figure 6.4. Mediator tasks at service discovery.

At service discovery, the mediator receives a discovery message from the consumer. It sends a service description and privacy policy message to the service registry and privacy if the discovery message includes a service request and privacy policy. Then, the mediator receives a service description and privacy policy result message from the service registry and privacy. Finally, it sends a final result message to the consumer.

6.2 Privacy

The privacy service is the service which guarantees that the selected service has a policy compatible with the privacy preferences. This service is used only if the provider wants to publish a service with a privacy policy and the consumer wants to discover a service that satisfies its privacy preferences. This is done by sending a policy to the mediator to publish and discover services. Thus, as long as the mediator receives a policy, it forwards the policy to the privacy, which can then perform its tasks.

The privacy service is responsible for the publication, unpublication and discovery of privacy policies. It provides these activities to the provider and consumer through the mediator. The privacy includes a policy registry, which is used to register provider policies at policy publication. These policies are retrieved by the privacy at policy discovery so that it can intersect them with the consumer policy. The mediator is responsible for sending the policies to the privacy. The privacy also includes an ontology registry, which is used to register the base and domain ontologies and query them in order to determine compatibility between consumer and provider policies at policy intersection. In order to verify policy compatibility, the privacy service retrieves the ontological concepts associated to each assertion in the policies. Then, it checks the relationship between the concepts in the ontologies. Domain representative organizations are responsible for developing domain-specific privacy ontologies and registering them in the privacy's ontology registry.

The activities of registration and deregistration of privacy ontologies, which are defined in order to apply the framework to specific domains, performed by the privacy are shown in Figure 6.5.

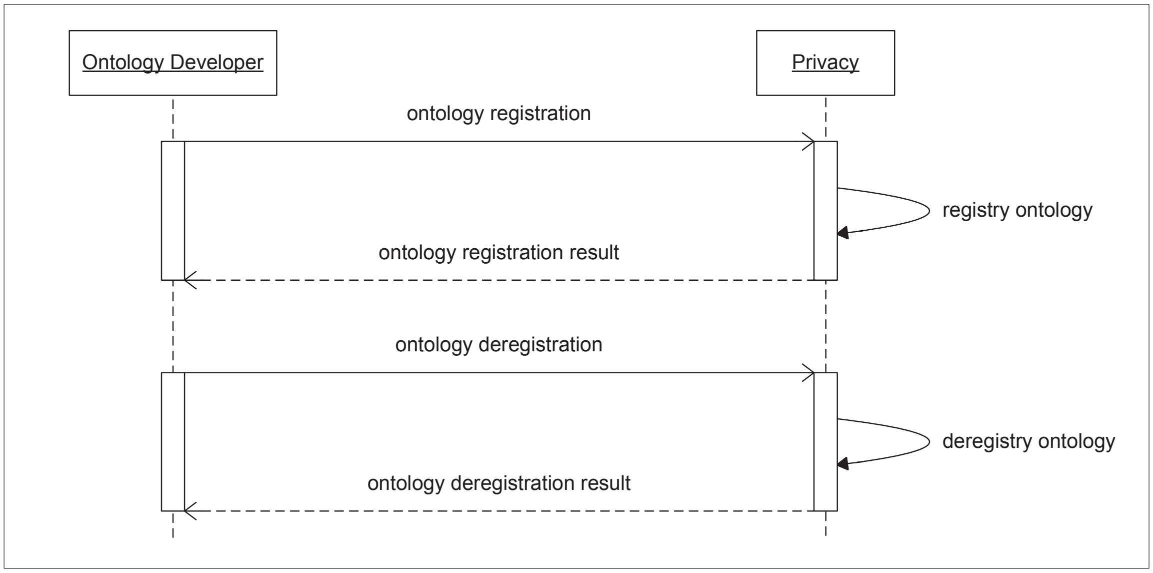

Figure 6.5. Registration and deregistration of ontologies.

At privacy ontology registration/deregistration, the privacy receives an ontology message from the ontology developer. Then, it registries/deregistries the ontology. Finally, the privacy sends an ontology result message, indicating the outcome of the activity, to the ontology developer.

The activities of privacy policy publication, unpublication and discovery performed by the privacy are shown in Figure 6.6.

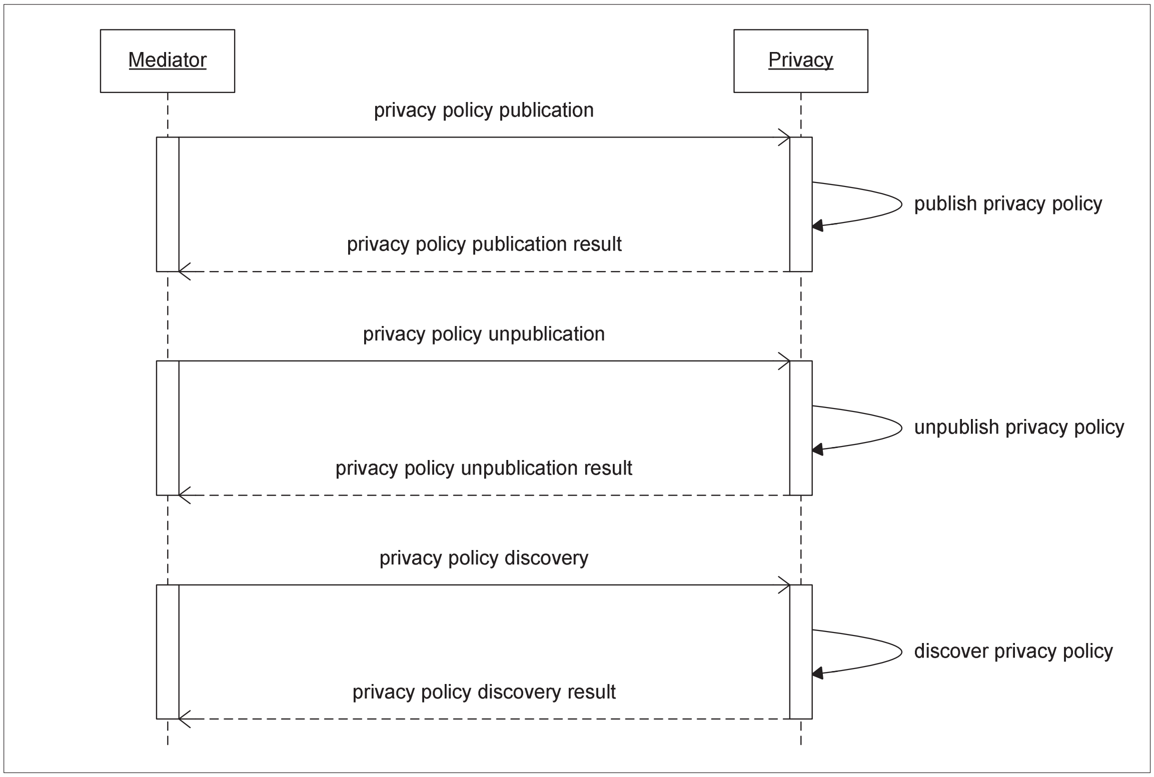
Mediator
Privacy
privacy policy publication
publish privacy policy
privacy policy publication result
privacy policy unpublication
unpublish privacy policy
privacy policy unpublication result
privacy policy discovery
discover privacy policy
privacy policy discovery result

Figure 6.6. Publication, unpublication and discovery of policies.

At service publication/unpublication/discovery, the privacy receives a privacy policy message from the mediator. Then, it publishes/unpublishes/discovers the privacy policy. Finally, the privacy sends a privacy policy result message, indicating the outcome of the activity, to the mediator.

6.3 Summary

This chapter presented the extensions to SOA roles and interactions for the framework proposed in this thesis. It described the two new roles included in SOA: mediator and privacy. The process of privacy-aware service discovery enabled by these extensions uses the policy model described in Chapter 5.

Chapter 7

Implementation and Evaluation

This chapter presents in Section 7.1 an implementation of the framework proposed in this thesis. In Section 7.2, it presents an evaluation of the framework in the domain of health care. The implementation and evaluation were performed in order to evaluate the effectiveness of the proposed framework.

In order to evaluate the proposed framework, a prototype was implemented so that tests could be executed. The goal of the evaluation was to check the effectiveness of the privacy preservation extension to Service-Oriented Architecture (SOA) and the advantage of using ontologies for comparing privacy policies. Thus, the parts of the framework regarding the definition of privacy policies as well as the publication and discovery of services based on policies were implemented. However, no Graphical User Interface (GUI) for consumers and providers was developed. Web services were employed in order to represent different consumers and providers with policies defined directly in Extensible Markup Language (XML) files.

A graphical interface for the proposed framework is important for its application. However, the emphasis of this thesis is on the development of an infrastructure for using semantic policies, which can be applied to enhancing privacy preservation in the areas of service description and discovery in SOA. Thus, the infrastructure was implemented and used to test the definition of policies and the discovery of services considering privacy policies. The user interface for the proposed framework is included in future work. Moreover, the particular way the infrastructure is going to be employed was not the focus of this implementation and evaluation is going to vary among domains and scenarios. The emphasis of this implementation and evaluation was, thus, on the integration of privacy preservation in service description and discovery through the use of semantic policies.

In the proposed framework, consumers and providers use policies to express their privacy preferences and practices. In the implementation and evaluation of the framework, policies were defined in XML files and these files were linked to service descriptions and requests through the use of Web services, which implemented the activities that are performed by providers and consumers in the proposed framework. Thus, the activities of requesting service publication and discovery were implemented by these Web services.

The framework defines a model in which service description includes information on provider practices as well as service request includes information on consumer preferences. This model was fully tested in the implementation and evaluation of the framework even though the policies were created manually, without tool support. In addition, the framework uses policies in order to enhance service discovery with privacy-awareness and this aspect of the framework was also fully tested. Thus, in the implementation and evaluation of the framework, service publication included privacy policies as well as service discovery considering the preferences and practices of consumers and providers expressed in their policies.

7.1 Implementation

A prototype of the framework was developed using Web service technology in order to evaluate it. Web services were implemented in the Java programming language, including a mediator and privacy Web services, which added the proposed privacy preservation features to the areas of service description and discovery. Other implemented Web services defined a SOA environment and represented a service provider, consumer and registry. The databases of the service registry for storing service descriptions were created using the Structured Query Language (SQL). Policies were created in order to demonstrate different cases in the domain scenario that was proposed for the evaluation of the framework. They were written using an extended version of the Web Services Policy Framework (WS-Policy), which was created in order to support the proposed policy model. The base and domain-specific ontologies created for the evaluation of the framework were written in the Web Ontology Language (OWL). The mediator, privacy and registry Web services were deployed on an application server. The following software products were used in order to build the prototype:

- Sun Java Development Kit Version 1.5: Java support.
- Apache Tomcat Version 4.0: an application server.
- MySQL AB MySQL Version 5.0: a database management system.
- Apache Axis Version 1.3: Web Services Description Language (WSDL) support and a SOAP engine.
- Apache jUDDI Version 0.9: a Universal Description Discovery & Integration (UDDI) registry.
- HP/IBM/SAP UDDI4J Version 2.0: a UDDI Java Application Programming Interface (API).
- Apache WS-Commons/Policy Version 0.9: WS-Policy support.
- Stanford Protégé 4.0: OWL support.

The prototype created an environment formed by a set of Web services (Figure 7.1). A Web service was used to provide the registry operations through the UDDI API and another Web service implemented the privacy service by using the OWL API. These Web services were encapsulated by a third Web service that implemented the mediator service, which provided an interface to the consumers and providers. In this setting, the consumers and providers were represented by Web services that used the operations provided by the mediator Web service in order to publish and discover services. The privacy policies of the consumers and providers were defined in files that were linked to ontologies through Protégé and processed in Java code through the Eclipse Integrated Development Environment (IDE).

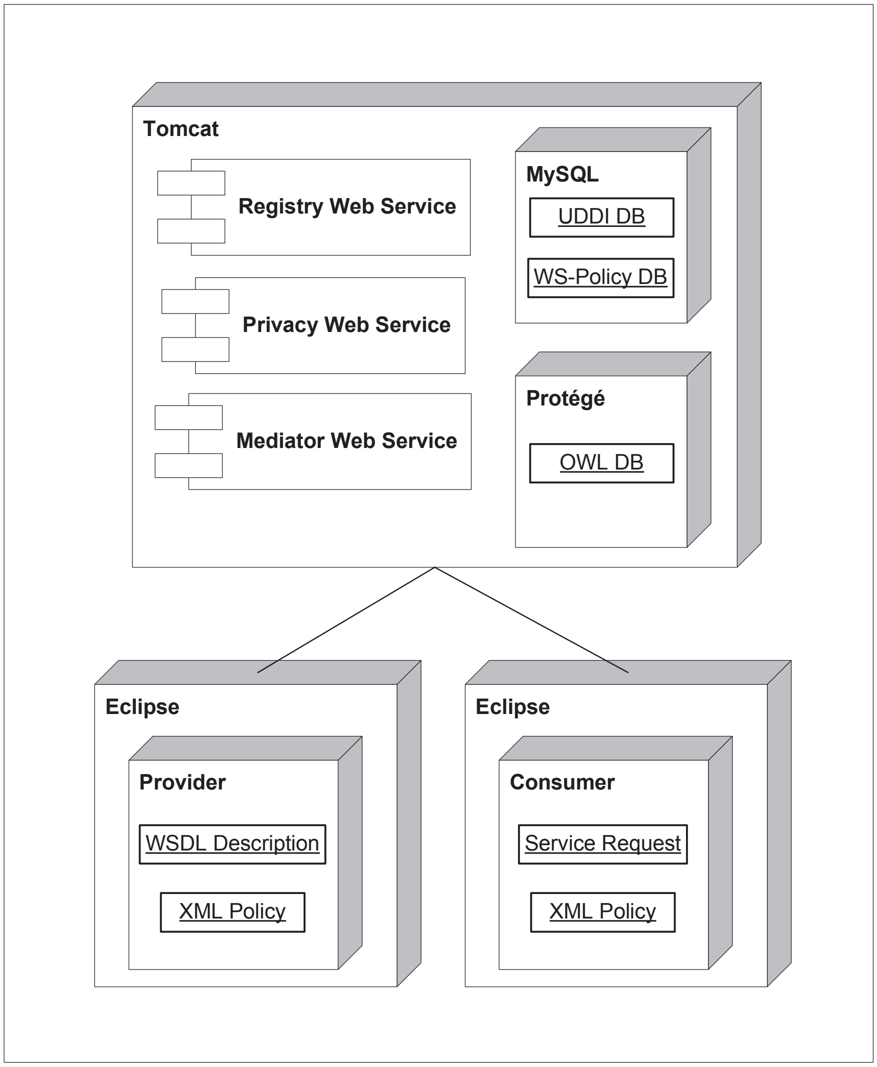

Figure 7.1. Prototype overview.

In order to test the framework, a SOA environment was developed with Web services to provide registry functionality as well as consumer and provider behaviors. Then, the privacy preservation extension was implemented and integrated in the environment. The privacy Web service was developed so that it had access to the ontologies, which were used to perform policy intersection at service discovery. The mediator Web service was the point from which the consumers and providers had access to the privacy preservation features that were added to SOA. The privacy Web service was one of the main parts under evaluation as it was the service responsible for performing the comparison between the policies, which were defined in order to represent different privacy requirements and guarantees that could be present in the scenario used for the evaluation. The consumers and providers were represented by Web services that interacted with the mediator Web service and sent it their service descriptions, requests and privacy policies. The mediator service was, thus, another of the main parts under evaluation, specifically its capability of integrating the new features in the traditional SOA environment set for the development of the prototype.

In the implementation and evaluation, the consumers used the services provided by the providers. The providers required private information from the consumers so that they could use the services supplied by the providers. In the evaluation of the framework, the consumers needed to know how the providers would use their information so that the consumers could decide whether to disclose the information to the providers or not. Policies were created to the consumers that specified their privacy preferences. Additionally, the providers had policies that were created and linked to them describing their privacy practices in the context of services provided by them. Provider policies were published so that they could be considered when discovering services for the consumers, which were concerned with their privacy.

The implementation of the framework included the proposed model for semantic privacy policies and the process of privacy-aware service discovery through the extension of the traditional SOA environment. The evaluation of the model for semantic privacy policies was carried out by specifying the description of the privacy preservation practices of the service providers and the privacy preservation preferences of the service consumers in their policies. These specifications used the base privacy ontology, which represented general privacy preservation issues, and a privacy ontology that extended the base ontology, which captured domain-specific privacy issues.

The evaluation of the process of privacy-aware service discovery was carried out by discovering services that met the privacy preferences of the consumers through the use of the model for semantic privacy policies. The privacy policies were intersected in order to select the services from the providers whose policies matched the consumer policies as a refinement to service discovery. Thus, the implementation enabled the evaluation of the proposed framework regarding its privacy preservation support for the areas of service description and discovery in SOA.

The process of privacy-aware service discovery was accomplished by extending basic SOA with the new roles and the activities that supported the idea of different registry types, including the registries for service descriptions and privacy policies, which constituted the publication and discovery space of the prototype. The new roles proposed to extend SOA, that is, the mediator and privacy roles, were implemented by services and added one in the publication and discovery space and the other interfacing this space. The operations and messages that implement the behaviors of the mediator and privacy roles presented in Chapter 6 are described as follows.

7.1.1 Mediator

In the implementation and evaluation of the proposed framework, the mediator role was played by a Web service. This Web service was included in a traditional SOA environment in order to facilitate the interactions between the providers and the services in the publication and discovery space as well as the consumers and these services. The publication and discovery services included a registry and privacy Web service. In the implementation of the prototype, the mediator Web service was responsible for service publication and discovery through the coordination of the services in the publication and discovery space. It used the registry and privacy Web services in order to extend the activities of publication and discovery with privacy preservation features.

The mediator Web service included a registry of publication and discovery services. This registry was used to store the addresses of the registry and privacy Web services employed in the evaluation of the framework. In the proposed framework, registry and privacy providers are responsible for registering their services in the registry of the mediator. However, in the implementation and evaluation of the framework, this task was performed manually. In order to evaluate the extension to SOA, the mediator Web service was employed to check the messages received from the providers or consumers and to decide which publication or discovery services were needed in order to execute the requested operations. The mediator Web service then retrieved the addresses of the registry and privacy Web services from its registry so that it could use them in order to coordinate the execution of the required operations.

The WSDL interface of the mediator is shown in Appendix A.1. It includes operations and messages to perform the activities of registration and deregistration of publication and discovery services as well as the tasks the mediator is responsible for at service publication, unpublication and discovery. They are described as follows.

- **Publication and Discovery Service Registration/Deregistration**

The registration process of publication and discovery service is executed by the Register Publication and Discovery Service operation. This operation receives a registration message of publication and discovery service and triggers the registration process. This message includes a description of publication and discovery service. This description includes the type of information managed by the service and its address. The operation inserts a description in the registry of publication and discovery service. The service address is set as its identifier. The operation sends a registration result message of publication and discovery service and ends the process. This message indicates if a registration was successful or failed. The deregistration process of publication and discovery service is executed by the Deregister Publication and Discovery Service operation, which performs the reverse operation.

- **Service Publication/Unpublication**

The tasks of service publication performed by the mediator are executed by the Publish Service operation. The operation receives a message of service publication and triggers the process. It can start the process of publication of a service description, privacy policy or a service description and privacy policy. The type of publication process this operation starts depends on the content of the received message. A message can include a service description and privacy policy. A service description includes information on the functionality of the service and how to use it, including the service address. A privacy policy includes information on the practices of the provider in the context of the service. A SOAP example of this message is shown in Figure 7.2.

```
<Envelope>
  <Body>
    <servicePublication>
      <serviceDescription>
        http://lh:8080/axis/ServiceDescription.wsdl
      </serviceDescription>
      <privacyPolicy>
        http://lh:8080/axis/ProviderPrivacyPolicy.xml
      </privacyPolicy>
    </servicePublication>
  </Body>
</Envelope>
```

Figure 7.2. Service publication message.

The Publish Service operation verifies if a service publication message includes a service description and privacy policy. It sends publication messages of service description and privacy policy. These messages include a service description with a service identifier and a privacy policy with a service identifier. The operation receives publication result messages of service description and privacy policy. These indicate if publications of service description and privacy policy were successful or failed. The operation joins the messages together in a final result message of service publication. This message indicates the result of each publication of service description and privacy policy. The operation sends it and ends the service publication process. The service unpublication tasks performed by the mediator are executed by the Unpublish Service operation, which performs the reverse operation.

- **Service Discovery**

The service discovery tasks performed by the mediator are executed by the Discover Service operation. It receives a service discovery message and triggers the process. This operation can start the discovery process of a service description, privacy policy or a service description and privacy policy. The type of discovery process depends on the content of the received message. A message can include a service request and privacy policy. A service request includes information on the required functionality of the service. A privacy policy includes information on the preferences of the consumer. A SOAP example of this message is shown in Figure 7.3.

```
<Envelope>
 <Body>
  <serviceDiscovery>
   <serviceRequest>
    http://lh:8080/axis/ServiceRequest.xml
   </serviceRequest>
   <privacyPolicy>
    http://lh:8080/axis/ConsumerPrivacyPolicy.xml
   </privacyPolicy>
  </serviceDiscovery>
 </Body>
</Envelope>
```

Figure 7.3. Service discovery message.

The Discover Service operation verifies if a service discovery message includes a service request and privacy policy. It sends discovery messages of service description and privacy policy. These messages include a service request and privacy policy. The operation receives result messages of service description and privacy policy. These messages include the identifier of the services that own the discovered service descriptions and privacy policies along with the associated intersection policies. They are empty if no service description and privacy policy were discovered. The operation intersects the service identifier sets from a non-empty result message of service description and privacy policy and orders the intersection set by following the order of the result message of privacy policy, if there is one, starting from the first received message of each message type, if there is more than one type of discovery result message. It creates a final result message of service discovery. This message includes the first service identifier from the intersection set and its intersection policy, if there is a non-empty intersection set; the first service identifier from the first received result message, if there is only one type of result message, and its intersection policy, if the message type is privacy policy; or is empty, if the intersection set is empty. An empty message indicates the reason, which can be no service description discovered, if all result messages of service description are empty; no privacy policy discovered, if all result messages of privacy policy are empty; or, otherwise, no service description-privacy policy pair discovered, if the intersection set is empty. The operation sends a final result message and ends the service discovery process.

7.1.2 Privacy

The privacy role was played by a Web service that was employed in the implementation and evaluation of the framework in order to guarantee that the selected services had policies compatible with the preferences of the consumers. The goal of the framework evaluation was to test several cases. In these evaluation cases, the providers wanted to publish services with privacy policies and the consumers wanted to discover services that satisfied their privacy preferences. Thus, the privacy Web service was used. In order to perform these tests, the Web services that represented the consumers and providers sent their policies to the mediator Web service in order to publish and discover the services. Then, the mediator Web service forwarded the received policies to the privacy Web service. Upon the receipt of the policies from the mediator Web service, the privacy Web service was responsible for the publication, unpublication and discovery of the policies.

In the implementation of the framework, the privacy Web service included a policy registry, which was used to store the provider policies. Then, these policies were retrieved by the privacy Web service in order to intersect them with the consumer policies. The mediator Web service was responsible for sending the provider and consumer policies to the privacy service. In addition to the policy registry, the privacy Web service included an ontology registry in order to store the base and domain ontologies as well as query them to determine policy compatibility.

Policy compatibility verification was carried out in the framework implementation and evaluation by retrieving the ontological concepts associated to each assertion in the policies through the privacy Web service. Then, the privacy service checked the relationship between the concepts in the ontologies. In the proposed framework, domain representative organizations are responsible for developing domain-specific privacy ontologies. However, in the implementation and evaluation of the proposed framework, the domain ontology was developed based on a domain-specific privacy regulation and manually registered in the privacy Web service.

The WSDL interface of the privacy is shown in Appendix A.2. It includes operations and messages to perform the activities of registration and deregistration of privacy ontologies as well as the activities of publication, unpublication and discovery of privacy policies. They are described as follows.

- **Privacy Ontology Registration/Deregistration**

The registration process of privacy ontology is executed by the Register Privacy Ontology operation. This operation receives a registration message of privacy ontology and triggers the process. This message includes the ontology address. The operation registries an ontology in the ontology registry. The ontology address is set as its identifier. The operation sends a registration result message of privacy ontology and ends the process. This message indicates if a registration was successful or failed. The deregistration process of privacy ontology is executed by the Deregister Privacy Ontology operation, which performs the reverse operation.

- **Privacy Policy Publication/Unpublication**

The publication process of privacy policy that supports service publication is executed by the Publish Privacy Policy operation. This operation receives a publication message of privacy policy and triggers the process. This message includes a privacy policy with a service identifier. A SOAP example of this message is shown in Figure 7.4.

```
<Envelope>
 <Body>
  <privacyPolicyPublication>
   <privacyPolicy>
    http://lh:8080/axis/ProviderPrivacyPolicy.xml
   </privacyPolicy>
   <serviceIdentifier>
    http://lh:8080/axis/Service.jws
   </serviceIdentifier>
  </privacyPolicyPublication>
 </Body>
</Envelope>
```

Figure 7.4. Policy publication message.

The operation publishes a privacy policy in the policy registry along with a service identifier. The address of the privacy policy is set as its identifier. The operation sends a publication result message of privacy policy and ends the process. This message indicates if a privacy policy publication was successful or failed. The unpublication process of privacy policy that supports service unpublication is executed by the Unpublish Privacy Policy operation, which performs the reverse operation.

- **Privacy Policy Discovery**

The discovery process of privacy policy that supports service discovery is executed by the Discover Privacy Policy operation. This operation receives a message and triggers the discovery process. This message includes a privacy policy. A SOAP example of this message is shown in Figure 7.5.

```
<Envelope>
  <Body>
    <serviceDiscovery>
      <privacyPolicy>
        http://lh:8080/axis/ConsumerPrivacyPolicy.xml
      </privacyPolicy>
    </serviceDiscovery>
  </Body>
</Envelope>
```

Figure 7.5. Policy discovery message.

The operation intersects the received policy with the policies in the policy registry. It sends a discovery result message of privacy policy and ends the process. This message includes a set of intersection policies along with their associated service identifiers or is empty if the policy intersection result is empty.

7.2 Evaluation

The implementation of the proposed framework included the mediator and privacy extensions. This implementation provided the features of service publication and discovery considering privacy policies and, thus, offered the necessary infrastructure to evaluate the effectiveness of the proposed framework through different evaluation cases with several comparisons among policies of consumers and providers in a fictitious scenario.

Among the different domains, health care is an example in which privacy preservation is particularly important, as health information is usually regarded as sensitive. Thus, the domain of health care [19] was chosen in order to evaluate the effectiveness of the framework.

The evaluation of the proposed framework involved the extension of the base privacy ontology for the domain of health care and the creation of a health care scenario, which were used to demonstrate cases in which the consumers had their policies checked against the policies of the providers in order to verify if the practices of the providers satisfied the preferences of the consumers. Thus, the framework evaluation included the following main activities:

- Development of a domain-specific privacy ontology, with the use of a health care privacy regulation in order to extend the base ontology.
- Creation of a health care scenario, with the inclusion of interactions that could demonstrate the capabilities of the proposed SOA extension.
- Definition of evaluation cases, with the specification of policies by following the created scenario and using the developed health care ontology.

This section presents the evaluation by describing the privacy ontology for the domain of health care as well as the scenario and cases of privacy preservation in health care defined in order to perform the evaluation.

7.2.1 Health Care Ontology

In order to include semantics in the policy model, the proposed framework enables the use of ontologies for the specification of privacy policies. Thus, the ontologies define the vocabulary for the creation of the privacy policies. The approach followed by the framework considers that privacy preservation issues can be separated in general and domain-specific issues.

At the first step in order to evaluate the proposed framework, in addition to the base privacy ontology, a domain-specific privacy ontology was developed in order to deal with the privacy preservation issues that are specific to a particular domain. The domain that was chosen in order to carry out the evaluation of the framework was the health care domain. The concepts from the health care privacy ontology were referred to together with the concepts from the base privacy ontology in the policy assertions in order to restrict different aspects of the use of private information.

The health care ontology for privacy preservation is based on the Personal Health Information Protection Act (PHIPA) [28]. This regulation provides useful definitions for extending the base ontology in order to create a domain-specific one for the health care domain. The definitions extend some aspects captured in the base ontology (Figure 5.16), including *Information*, *Collector*, *Collection*, *Recipient* and *Purpose*. These definitions along with their associated concepts in the health care ontology are presented as follows.

The concepts related to *Information* are shown in Table 7.1. The information types are divided in two categories: *Personal Health Information* (Concept 01) and *Non Personal Health Information* (Concept 11). *Personal Health Information* is defined by a set of information types (Concepts 02-10).

	Information	Definition
01	Personal Health Information	Health-related information.
02	Patient Identification	Information that can be used to identify the individual on its own or linked to another piece of information, including the individual's health insurance number.
03	Health	Information that relates to the individual's primary or mental health.
04	Family Health History	Information about the individual's family history that relates to health.
05	Health Care	Information on the health care received by the individual.
06	Health Care Provider Identification	Information that can be used to identify the health care provider responsible for providing health care to the individual.
07	Health Care Payment	Information that relates to the individual's payment for health care as well as the individual's eligibility for health care or for coverage for health care under a health insurance plan.

08	Body Part Donation	Information on the individual's donation of body parts or bodily substances.
09	Substitute Decision-Maker Identification	Information that can be used to identify the individual's substitute decision-maker.
10	Personal Health Information Accompanying Information	Information that belongs to none of the previous categories but is part of a record that contains personal health information.
11	Non Personal Health Information	Non health-related information.

Table 7.1: Health Care Ontology – Information

The concepts related to *Collector* are shown in Table 7.2. The collector types are divided in two categories: *Health Care Provider* (Concept 01) and *Non Health Care Provider* (Concept 09). These categories are defined by sets of collector types: *Health Care Provider* (Concepts 02-08) and *Non Health Care Provider* (Concepts 10-12).

	Collector	Definition
01	Health Care Provider	A service provider that supplies a service related to health care.
02	Health Information Custodian	A health information custodian is a person whose primary purpose is the provision of health care.
03	Health Information Custodian Agent	An agent of a health information custodian is a person who performs activities (collection, disclosure, retention and use) over personal health information on behalf of the health information custodian.
04	Substitute Decision-Maker	A substitute decision-maker is a person legally entitled to make decisions for the individual that are necessary for, or auxiliary to, the individual's health care.

05	Privacy Commissioner	The Privacy Commissioner is the authority responsible for managing compliance with the PHIPA.
06	Primary Health Care Provider	A service provider that supplies a service related to primary health care.
07	Mental Health Care Provider	A service provider that supplies a service related to mental health care.
08	Pharmacy	A service provider that supplies a service related to medication.
09	Non Health Care Provider	A service provider that supplies a service unrelated to health care.
10	Insurer	A service provider that supplies a service related to health insurance.
11	Housing Provider	A service provider that supplies a service related to housing to a patient.
12	Employment Provider	A service provider that supplies a service related to employment to a patient.

Table 7.2: Health Care Ontology – Collector

The concepts related to *Collection* are shown in Table 7.3. Two aspects are used in order to categorize collection types: consent and source. Consent indicates whether or not consent is required and the type of consent (Concepts 01-03). The source aspect indicates from whom information is collected (Concepts 04-05).

	Collection	Definition
01	Express Consent	Express consent is required when giving personal health information to a person who is not a health information custodian or when giving personal health information to a health information custodian for a purpose unrelated to health care.
02	Implied Consent	Implied consent can be relied on when consent is required, but not express consent.

03	Without Consent	Personal health information can be collected without consent as required in order to comply with laws.
04	Direct	Personal health information can be collected directly from its subject.
05	Indirect	Personal health information can be collected from a person other than the subject individual or its substitute decision-maker.

Table 7.3: Health Care Ontology – Collection

The concepts related to *Recipient* are shown in Table 7.4. There are three recipient types (Concepts 01-03).

	Recipient	Definition
01	Recipient Health Information Custodian	A health information custodian can disclose personal health information under its control to another health information custodian.
02	Recipient Health Information Custodian Agent	A health information custodian can share personal health information under its control with its agent.
03	Personal Health Information Recipient	A health information recipient is a person who receives personal health information from a health information custodian but does not act on the health information custodian's behalf and does not use the personal health information for a purpose related to health care.

Table 7.4: Health Care Ontology – Recipient

The concepts related to *Purpose* are shown in Table 7.5. These concepts are divided in two categories: *Health Care-Related* (Concept 01) and *Non Health Care-Related* (Concept 08).

These categories are defined by sets of purpose types: *Health Care-Related* (Concepts 02-07) and *Non Health Care-Related* (Concepts 09-11).

	Purpose	Definition
01	Health Care-Related	Any purpose that is related to health care.
02	Primary Health Care	Any purpose that is related to primary care.
03	Mental Health Care	Any purpose that is related to mental health care.
04	Health Treatment	Any observation, examination, assessment, care, service or procedure that is provided to diagnose, treat or maintain the individual's primary or mental health condition.
05	Health Prevention and Promotion	Any observation, examination, assessment, care, service or procedure that is provided to prevent disease or injury or to promote health.
06	Palliative Health Care	Any observation, examination, assessment, care, service or procedure that is provided as part of palliative care.
07	Medication	Any purpose that is related to medication.
08	Non Health Care-Related	Any purpose that is not related to health care.
09	Health Insurance	Any purpose that is related to health insurance.
10	Patient Housing	Any purpose that is related to housing for a patient.
11	Patient Employment	Any purpose that is related to employment for a patient.

Table 7.5: Health Care Ontology – Purpose

7.2.2 Evaluation Scenario

The second step of the three main steps for the evaluation of the proposed framework using the implemented prototype was the creation of a scenario, which could be used to execute the tests. A fictitious scenario was created considering the domain of health care so that the domain-specific privacy ontology developed at the first step could by applied to the framework evaluation.

One of the main constraints for the definition of the evaluation scenario was to include interactions among the different parts involved in the scenario, which could be explored in the evaluation cases in order to demonstrate different capabilities of the proposed SOA extension. This constraint shaped the evaluation scenario, which also considered the health care privacy ontology. The evaluation scenario provided the basis for the definition, publication and discovery of services on the prototype during the execution of the evaluation of the proposed framework.

Figure 7.6 shows the scenario in the domain of health care created in order to evaluate the framework. It is based on examples of privacy preservation from a PHIPA toolkit [10].

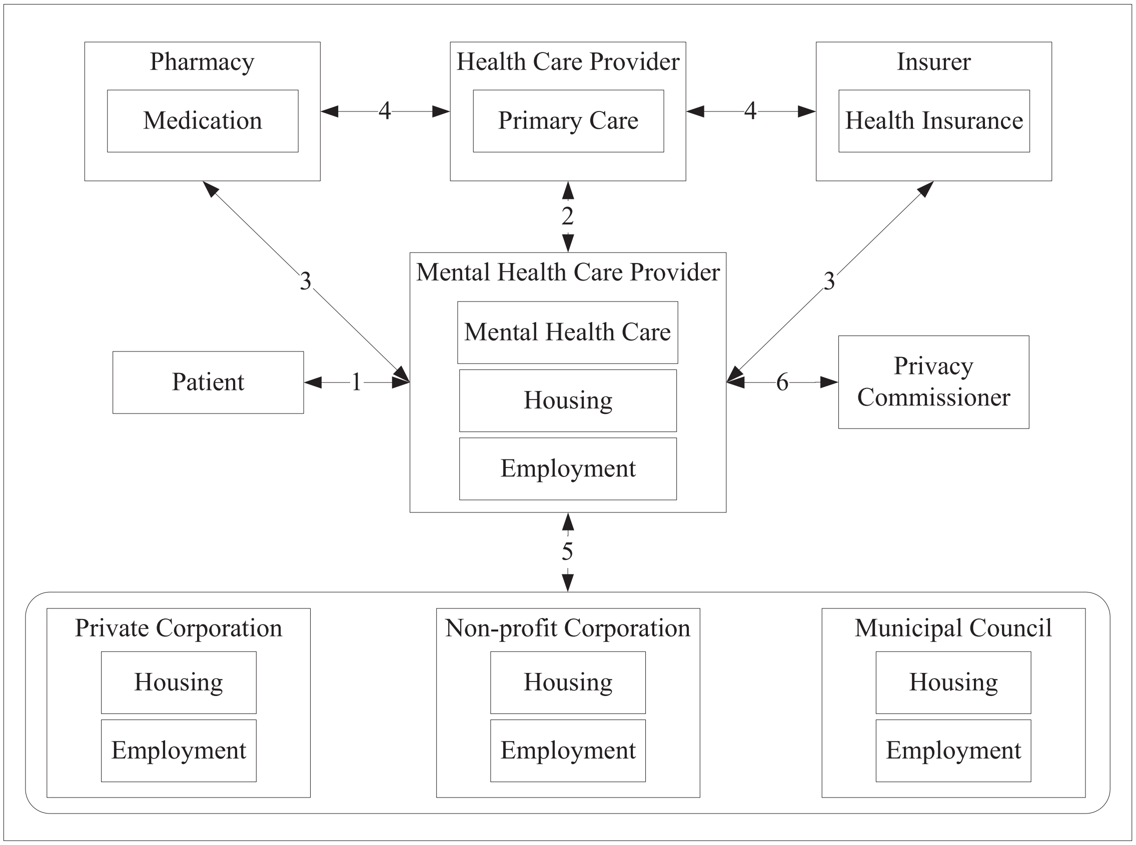

Figure 7.6. Evaluation scenario.

In the scenario, a patient uses services provided by a mental health care service provider. In order to use the services, the patient discloses some of its health information (*Collection*). This interaction is labeled as 1 in Figure 7.6. In addition to mental health care services, it uses other health care-related services offered by the service provider, including primary health care, as well as services unrelated to health care, such as housing and employment services.

The mental health care service provider employs a holistic approach for health care, that is, it provides primary health care along with mental health care. The primary care services are not provided directly by the mental health care service provider, but by a third-party health care service provider (Interaction 2). In this case, the mental health care service provider, which is a custodian, discloses the health information of the patient to another custodian, the health care service provider (*Recipient*).

In order to supply its services, the mental health care service provider uses health insurance services from insurers and medication services from pharmacies (Interaction 3). Insurers and pharmacies can be third-party organizations or part of the mental health care service provider. In the first case, the insurer is a recipient and the pharmacy is a custodian as both service providers receive the health information of the patient from the mental health care service provider (*Collector*). In the second one, both the insurer and pharmacy service providers are agents of the custodian. Third-party health care service providers can also use services provided by insurer and pharmacy service providers so that multiple levels of disclosure of health information can take place (Interaction 4).

The mental health care service provider offers housing and employment services to its patients. These services are offered so that the patient can get help in looking for house and employment (*Purpose*). The mental health care service provider offers several options for the services. They can be supplied by the mental health care service provider itself or by a third party, which is a recipient in this case. The third-party service provider can be a private corporation, non-profit corporation or municipal council (Interaction 5).

As the Privacy Commissioner is the authority responsible for managing compliance with the PHIPA, there can be an interaction between any service provider and the Privacy Commissioner. In the evaluation scenario, there is an interaction between the mental health care service provider and the Privacy Commissioner (Interaction 6). The mental health care service provider may have to send information (*Information*) about its patients to the Privacy Commissioner in response to requests from it as required by public regulations.

7.2.3 Evaluation Cases

The last step of the evaluation that was carried out using the prototype of the proposed framework was the definition of evaluation cases. The evaluation cases were defined in order to demonstrate the capabilities of the proposed framework regarding the discovery of services considering privacy policies. Thus, the cases were defined based on the evaluation scenario created at the previous step. The main part of the definition of the evaluation cases was the creation of the provider and consumer policies, which used the base and health care privacy ontologies. These policies were created according to the interactions included in the evaluation scenario. The evaluation cases were then executed in order to demonstrate which of the interactions were possible to happen based on the policies that were defined for each of the involved parts.

Several evaluation cases were created in order to cover the interactions in the evaluation scenario. The cases show consumers and providers with different preferences and practices specified by following the policy format. The policies were created using the vocabulary defined by the base (Section 5.4) and health care (Section 7.2.1) ontologies. For readability, the policy format is not shown in the policies.

- **Evaluation Case - Health Care Provider**

This case considers Interactions 1 and 2 in the scenario (Figure 7.6). It aims at exemplifying the use of domain-specific knowledge for the verification of compatibility between policies. A mental health care service provider can disclose health information about their patients to a health care service provider for the purpose of primary health care if it is authorized to do so by the original owner of the information. A third party can have the same status as the information owner for that purpose as a substitute decision maker. Thus, that third party would be able to grant the required disclosure authorization to the mental health care service provider.

In this case, a patient named *Patient* publishes a policy. In its policy, it states that a third party named *ThirdParty* is its substitute decision maker for the purpose of health care. Figure 7.7 shows this statement.

Policy Owner: Patient

Information = PersonalHealthInformation
Collector.ProviderName = ThirdParty
Collector.ProviderType = SubstituteDecisionMaker
Recipient
Purpose = HealthCareRelated

Figure 7.7. Patient policy for substitute decision maker.

Additionally, a mental health care service provider named *MentalProvider* publishes a policy, which states that it discloses health information collected from its patients to a primary health care service provider for the provision of a primary health care service if the patient allows doing so. Figure 7.8 shows this statement.

Policy Owner: MentalProvider

Information = PersonalHealthInformation
Collector.ProviderName = MentalProvider
Collector.ProviderType = MentalHealthCareProvider
Recipient.ProviderType = PrimaryHealthCareProvider
Purpose = PrimaryHealthCare

Figure 7.8. Provider policy for primary health care.

Continuing the case, *ThirdParty*, looking for a mental health care service provider that follows a holistic approach for *Patient*, publishes its policy. It states that health information about the patient can be disclosed by the mental health care service provider to a health care service provider for purposes related to health care if the patient allows doing so. This statement is shown in Figure 7.9.

```
Policy Owner: Patient

Information = PersonalHealthInformation
Collector.ProviderType = MentalHealthCareProvider
Recipient.ProviderType = HealthCareProvider
Purpose = HealthCareRelated
```

Figure 7.9. Patient policy for mental health care.

In this case, the mediator selected the mental health care service supplied by *MentalProvider* for *Patient* because the privacy known that *ThirdParty* was a substitute decision maker for *Patient* and it could make decisions on behalf of a patient if authorized to do so.

- **Evaluation Case - Pharmacy and Insurer**

This case considers Interactions 1, 3 and 4 from Figure 7.6. It aims at exemplifying the execution of compositional reasoning for the verification of compatibility between policies assuming that the patient allowed sharing information among the several services. A mental health care service provider uses medication and health insurance services supplied by third-party service providers in order to provide its mental health care service. For these purposes, it discloses health information of its patients to a pharmacy service provider and an insurer service provider. In addition to these external services, it uses a service of primary health care offered by a health care service provider. The third-party health care service provider requires health information about the patient from the mental health care service provider and discloses it to an external pharmacy service provider and an insurer service provider. A patient allows health information to be used and disclosed by the health information collector for any health care-related purpose, but the patient does not accept it to be disclosed by third-party service providers, even for health care-related purposes. Thus, that patient would not be able to disclose health information to that mental health care service provider.

In this case, a health care service provider named *HealthProvider* publishes a policy. It states that health information collected from patients of its primary health care service is disclosed to a third-party pharmacy and insurer service provider for the purpose of medication and health care insurance. Figure 7.10 shows these statements.

```
Policy Owner: HealthProvider

Information = PersonalHealthInformation
Collector.ProviderName = HealthProvider
Collector.ProviderType = HealthCareProvider
Recipient.ProviderType = Pharmacy
Purpose = Medication

Information = PersonalHealthInformation
Collector.ProviderName = HealthProvider
Collector.ProviderType = HealthCareProvider
Recipient.ProviderType = Insurer
Purpose = Health Insurance
```

Figure 7.10. Provider policy for primary health care.

In addition, a mental health care service provider named *MentalProvider* publishes a policy, which states that it discloses personal health information of its patients to a pharmacy service provider and third-party insurer for the purpose of medication and health care insurance. The mental health care service provider follows a holistic approach to health care and uses the primary health care service offered by *HealthProvider*. Figure 7.11 shows these statements.

```
Policy Owner: MentalProvider

Information = PersonalHealthInformation
Collector.ProviderName = MentalProvider
Collector.ProviderType = MentalHealthCareProvider
Recipient.ProviderType = Pharmacy
Purpose = Medication

Information = PersonalHealthInformation
Collector.ProviderName = MentalProvider
Collector.ProviderType = MentalHealthCareProvider
Recipient.ProviderType = Insurer
Purpose = Health Insurance

Information = PersonalHealthInformation
Collector.ProviderName = MentalProvider
Collector.ProviderType = MentalHealthCareProvider
Recipient.ProviderType = HealthProvider
Purpose = PrimaryHealthCare
```

Figure 7.11. Provider policy for mental health care.

Continuing the case, a patient named *Patient*, looking for a mental health care service, publishes its policy, which states that it allows health information to be disclosed by the mental health care service provider for any health care-related purpose. However, it does not allow information to be disclosed by third-party service providers employed by the mental health care service provider. These statements are shown in Figure 7.12.

```
Policy Owner: Patient

Information = PersonalHealthInformation
Collector.ProviderType = MentalHealthCareProvider
Recipient
Purpose = HealthCareRelated

Information = PersonalHealthInformation
Collector.ProviderType = Recipient
NoRecipient
```

Figure 7.12. Patient policy for mental health care.

In this case, the mediator did not select the mental health care service provided by *MentalProvider* to *Patient*. This happened because it used a primary health care service offered by *HealthProvider* and this provider disclosed information to others. As the health care service provider was a third-party provider in the interaction between the patient and the mental health care service provider, the preferences of the patient were not satisfied by its practices.

- **Evaluation Case - Housing and Employment**

This case considers Interactions 1 and 5 from Figure 7.6. It aims at exemplifying the check of relationship between different terms for the verification of compatibility between policies. A mental health care service provider offers housing and employment services to its patients. The patients can choose among several provider types, including the mental health care service provider itself and different types of third-party providers. A patient needs housing and employment services, but it requires that they be supplied by a mental health care service provider directly. Thus, that patient would be able to use the housing and employment services provided directly by that mental health care service provider.

In this case, a mental health care service provider named *MentalProvider* publishes a policy, which states that it uses health information of its patients for purposes unrelated to health care. Alternatively, the information is disclosed by the mental health care service provider to a private corporation, non-profit corporation or municipal council for the provision of the services of housing and employment. Figure 7.13 shows the statements for the housing service. A similar set of statements would be necessary for the employment service.

```
Policy Owner: MentalProvider

Alternative 1

Information = PersonalHealthInformation
Collector.ProviderName = MentalProvider
Collector.ProviderType = MentalHealthCareProvider
Purpose = PatientHousing

Alternative 2

Information = PersonalHealthInformation
Collector.ProviderName = MentalProvider
Collector.ProviderType = MentalHealthCareProvider
Recipient.ProviderType = PrivateCorporation
Purpose = PatientHousing

Alternative 3

Information = PersonalHealthInformation
Collector.ProviderName = MentalProvider
Collector.ProviderType = MentalHealthCareProvider
Recipient.ProviderType = PrivateCorporation
Purpose = NonProfitCorporation

Alternative 4

Information = PersonalHealthInformation
Collector.ProviderName = MentalProvider
Collector.ProviderType = MentalHealthCareProvider
Recipient.ProviderType = PrivateCorporation
Purpose = MunicipalCouncil
```

Figure 7.13. Provider policy for housing.

In addition, a patient named *Patient*, looking for housing and employment services supplied directly by a mental health care service provider, publishes its policy. It states that the patient allows health information to be used by the mental health care service provider but not disclosed for purposes unrelated to health care. This statement is shown in Figure 7.14.

```
Policy Owner: Patient

Information = PersonalHealthInformation
Collector.ProviderType = MentalHealthCareProvider
NoRecipient
Purpose = NonHealthCareRelated
```

Figure 7.14. Patient policy for housing and employment.

In this case, the mediator selected the housing and employment services supplied by *MentalProvider* for *Patient*. This happened because the privacy was able to check that the purposes of housing and employment were non health care-related purposes and, thus, the patient policy included a requirement that was more general than the guarantees offered by the mental health care service provider. The alternative chosen by the patient was included in the intersection policy so that the mental health care service provider could decide which source to use for the provision of the services.

- **Evaluation Case - Privacy Commissioner**

This case considers Interactions 1 and 6 from Figure 7.6. It aims at exemplifying the performance of conclusion for the verification of compatibility between policies. A group of health care service providers can disclose health information about their patients under certain regulations. A provider is a member of that group. Thus, a patient would not send their information to that provider if it does not want the information to be disclosed to any third-party provider.

In this case, the Privacy Commissioner publishes a policy. It states that mental health care service providers can disclose personal health information of their patients to the Privacy Commissioner in order to comply with public regulations. Figure 7.15 shows this statement.

```
Policy Owner: PrivacyCommissioner

Information = PersonalHealthInformation
Collector.ProviderType = MentalHealthCareProvider
Recipient.ProviderType = PrivacyCommissioner
Purpose = Legal
```

Figure 7.15. Privacy Commissioner policy for mental health care providers.

In addition, a provider named *MentalProvider* publishes a policy, which states that it is a mental health care service provider and collects health information from patients of its mental health care service. Figure 7.16 shows this statement.

Policy Owner: MentalProvider

Information = PersonalHealthInformation
Collector.ProviderName = MentalProvider
Collector.ProviderType = MentalHealthCareProvider

Figure 7.16. Provider policy for mental health care.

Continuing the case, a patient named *Patient*, looking for a mental health care service, publishes its policy. It states that the patient does not allow the disclosure of health information. This statement is shown in Figure 7.17.

Policy Owner: Patient

Information = PersonalHealthInformation
Collector.ProviderType = MentalHealthCareProvider
NoRecipient

Figure 7.17. Patient policy for mental health care.

In this case, the privacy was able to conclude that *MentalProvider* could disclose health information of its patients and, thus, *Patient* could not use the mental health care service supplied by the mental health care service provider. The mediator returned no service to the patient in this case. This shown that, in such a case, it was important to inform the reason for returning no service to the patient so that it could change its policy accordingly.

The implementation and evaluation of the proposed framework showed that it is able to deal with the privacy preservation problems in the areas of service description and discovery in SOA. These privacy problems are that, in traditional SOA, it is not possible to describe how providers deal with consumer private information and discover services that satisfy consumer privacy preferences. The proposed framework provided solutions to these problems that addressed the limitations identified in other privacy frameworks for SOA proposed in the literature as discussed in Chapter 3. The evaluation showed that the framework provides a policy model with a comprehensive vocabulary as well as support for the specification and intersection of policies. Furthermore, the evaluation of the framework showed that it offers service discovery integration as well as support for the consideration of domain-specific privacy preservation issues.

The elements of the policy model of the proposed framework enabled the specification of policies in the implementation and evaluation of the framework. These policies defined different aspects of privacy preservation of different information items in different settings, through the definition of components, assertions and alternatives, respectively. The format for semantic privacy policies that considers the proposed policy elements was used for the specification of preferences and practices of consumers and providers as policies. These privacy policies offered the base for the framework by improving service description and discovery, through the use of privacy policies in the processes of service description and discovery as well as the integration of a mechanism of policy intersection that indicated how privacy policies were to be matched.

Semantics support was included in the proposed policy model by enabling the use of privacy ontologies. In the evaluation of the proposed framework, the ontologies defined the vocabulary used to create the policies. They were developed according to the proposed approach that separated general and domain-specific privacy preservation issues, through the definition of the base and health care privacy ontologies. This approach addressed the aspect that each application domain includes specific privacy preservation issues. In the semantic privacy policy model, the policy assertions referred to ontological concepts and the policies were created from concepts defined in the base and health care privacy preservation ontologies.

In the implementation and evaluation of the proposed framework, the privacy preservation practices and preferences of the service providers and consumers were described in the privacy policies sent to the mediator Web service. Privacy policy intersection was employed through the privacy Web service in order to enhance service discovery so that the discovered services were from the providers whose privacy preservation practices matched the privacy preferences of the service consumers.

The use of the privacy policies for service discovery was accomplished by using the privacy and mediator Web services. The privacy Web service complemented the registry Web service with the feature of privacy preservation. These two Web services defined the publication and discovery space in which they were responsible for the publication and discovery of the services, through the coordination of the mediator Web service. The mediator Web service was added to the architecture so that the publication and discovery space was made transparent to the consumers and providers as well as support to additional Quality of Service (QoS) characteristics could be added by following the same approach used to deal with the privacy preservation issues.

Service description is typically restricted to functional characteristics of services and, consequently, service discovery is based on functionality of services. Thus, in order to include the privacy preservation characteristics of the services in the service descriptions, the framework extended traditional SOA. In the evaluation of the proposed framework, the extension allowed for service discovery that considered not only the functionality of the services but also their privacy characteristics.

7.3 Summary

This chapter presented an implementation and evaluation of the framework described in Chapters 5 and 6. The framework provided the following benefits:

- **Flexible format for privacy policies**. The framework provides a flexible format for privacy policies. Thus, it supports the specification of alternative practices and compact policies.
- **Extensive vocabulary of privacy preservation**. The framework provides a privacy vocabulary that covers the principles of privacy regulations. Thus, it supports the expression of complete practices.
- **Rich support for semantics**. The framework includes semantics by ontological annotation of privacy policies. Thus, it supports rich specification of privacy policies of consumers and providers and intersection between privacy preservation policies.
- **Complete support for domain-specific issues of privacy preservation**. The proposed framework enables the consideration of domain-specific issues of privacy preservation. Thus, it supports its application to different application domains.
- **Adequate modifications of SOA**. The framework does not modify the basic roles of SOA, does not require direct interaction with the component responsible for privacy preservation as well as does not require interactions related to privacy preservation between consumers and providers. Thus, it supports its deployment in SOA environments.
- **Integration in the process of service discovery**. The framework integrates privacy policies in the process of service discovery. Thus, it supports a privacy-aware process of service discovery.

- **Proper support to QoS extension**. The framework enables the inclusion of other QoS attributes with the separation of them. Thus, it supports QoS management without extra impacts.

Chapter 8
Conclusions

This chapter concludes the thesis. In Section 8.1, the chapter summarizes the solution proposed in this thesis for the problems of privacy preservation in the areas of service description and discovery in Service-Oriented Architecture (SOA). Then, the chapter presents the contributions of the thesis in Section 8.2. Finally, Section 8.3 discusses future work.

8.1 Summary

SOA can facilitate the development and management of software solutions. SOA has been an intense area of research, but the preservation of privacy in SOA still includes open problems. For example, two of these problems are that it is not possible to describe how a service provider deals with private information received from a service consumer as well as discover a service that satisfies the privacy preservation preferences of a service consumer in addition to the required service functionality.

Several privacy preservation frameworks for SOA were proposed in the literature. The surveyed frameworks offer a limited solution for the problems of privacy preservation in the areas of service description and discovery. Some existing frameworks do not define a format for privacy policies, whereas other frameworks define a policy format that has no flexibility. Furthermore, the privacy vocabularies of the surveyed frameworks miss important concepts, such as means of information collection, access to retained information by the owner of the information as well as availability of use record of collected information. Additionally, some existing frameworks do not include semantics, whereas other frameworks follow an approach that does not allow for a comprehensive privacy vocabulary. Moreover, none of the surveyed frameworks integrates privacy policies in the process of service discovery completely. In addition, some existing frameworks modify basic roles of SOA, require direct interaction with the component responsible for privacy preservation as well as require privacy preservation-related interactions between the service consumer and provider. Moreover, the surveyed frameworks do not enable the inclusion of other Quality of Service (QoS) attributes in a way that separates the different attributes. Finally, the surveyed frameworks do not fully support the consideration of domain-specific privacy preservation issues.

The framework proposed in this thesis provides a novel solution for the problems of privacy preservation in the areas of service description and discovery in SOA. The framework addresses the limitations identified in frameworks presented earlier.

The proposed framework includes a model for semantic privacy policies and support for privacy-aware service discovery. The policy model enables the description of privacy preservation practices of service providers and privacy preservation preferences of service consumers. In the policy model, policy assertions refer to ontological concepts. Thus, semantic policies are created from concepts defined in privacy ontologies. This semantic information enriches the matching between the privacy policies of a service consumer and provider. The policy matching supports the process of privacy-aware service discovery, which enables the discovery of services that meet privacy preferences of service consumers.

The proposed framework considers that service providers and consumers describe their privacy preservation practices and preferences in semantic privacy policies. Thus, the operation of policy intersection enhances the process of service discovery so that the services selected for the service consumer are from providers whose privacy practices match the privacy preferences of the consumer. In the proposed framework, the use of policies for service discovery is accomplished by extending SOA with two new roles, the privacy role and the mediator role. The privacy role is responsible for the publication and discovery of policies. The mediator role mediates the interactions related to service publication and discovery between the provider or the consumer and the registry and privacy services.

An aspect of privacy preservation is that despite the fact that some privacy preservation issues are common to different application domains, there are privacy preservation issues that are specific to each particular domain. In order to address this aspect of privacy preservation, the proposed framework follows an approach in which a base privacy ontology represents general privacy preservation issues, whereas domain-specific privacy issues are captured by domain-specific privacy ontologies that extend the base ontology.

The framework was implemented in order to evaluate the effectiveness of the proposal. Health-related information is usually regarded as sensitive information, thus the effectiveness of the framework was evaluated using a health care privacy ontology and an evaluation scenario of privacy preservation defined in the domain of health care. This evaluation was carried out in order to verify if the privacy preservation solution for the areas of service description and discovery satisfies the specific goals of the thesis. The evaluation involved the test of several privacy preservation cases in which service consumers had their privacy preservation preferences checked against the privacy preservation practices of service providers so that the service consumers could decide whether to select or not the services offered by the service providers in the evaluation scenario. The results of the evaluation demonstrated that the proposed framework offers an effective solution for the privacy preservation problems in the areas of service description and discovery. The framework provides a means for describing the privacy preservation practices of service providers and the preferences of consumers in a semantics-enriched way that enhances service description with privacy preservation practices and service request with privacy preservation preferences. Moreover, such enhancements support the integration of privacy preservation-awareness in service publication and discovery in order to enable the publication of privacy practices of service providers and a process of service discovery that considers privacy preferences of service consumers.

8.2 Contributions

The main contribution of this thesis is a privacy preservation framework for the areas of service description and discovery in SOA. The proposed framework enables service providers to describe their privacy preservation practices and service consumers to describe their privacy preservation preferences. In addition, the framework allows consumers to discover services that do not only meet their functional requirements but also satisfy their privacy preferences. Thus, specifically, the contributions of this thesis are a model for semantic privacy policies and a process of service discovery that considers privacy policies. These contributions offer solutions for the problems identified in the areas of service description and discovery in SOA. The policy model enables the specification of practices regarding the handling of private information. It is in conformity with the privacy definition as it allows for consumers and providers to define the information activities that are considered acceptable by them. The discovery integration enables the use of the policy model in service discovery. It is important as, at service discovery, relationships between consumers and providers can be formed and, thus, the suitability of the interactions has to be checked.

Regarding the area of service description, the proposed framework enhances service descriptions of providers with the inclusion of the description of privacy practices of service providers. The framework also enhances service requests of consumers with the inclusion of the description of privacy preferences of service consumers. This improvement to the area of service description is supported by the model for semantic privacy policies included in the proposed framework. This model enables the specification of semantic privacy policies of service providers and consumers using concepts defined in a base privacy ontology and domain-specific privacy ontologies. Thus, the policy model supports the use of semantic information on privacy preservation practices and preferences when matching policies of service providers and consumers.

Regarding the area of service discovery, the framework enhances service publication with the inclusion of the publication of semantic privacy policies of service providers. The framework also enhances service discovery with the inclusion of the intersection of privacy policies of service consumers and providers. This improvement to the area of service discovery is supported by the process of privacy-aware service discovery included in the proposed framework. This process enables the use of the semantic privacy policies of a service consumer and provider and the intersection between the policies in service publication and discovery for the selection of services that meet the privacy preservation requirements of consumers. Thus, the process of privacy-aware service discovery supports a dynamic approach that completely integrates semantic privacy policies of service consumers and providers in service discovery.

Compared to other privacy preservation frameworks for SOA proposed in the literature, the framework developed in this thesis offers the following new features and benefits for the preservation of privacy in the areas of service description and service discovery:

- A flexible format for privacy policies that supports the specification of alternative privacy practices and compact privacy policies of service consumers and providers.
- A privacy vocabulary that covers the principles in privacy preservation regulations and supports the expression of complete actual and acceptable privacy preservation practices of service providers and consumers, respectively.
- A semantic privacy policy model that includes ontological concepts in policy assertions and supports semantics-enriched privacy policy specification and intersection.
- An architecture that integrates semantic privacy policies in service discovery and supports a privacy-aware process of service discovery.
- An architecture that does not modify the basic roles of SOA, does not require direct interaction between the service consumer or the service provider and the component responsible for privacy preservation as well as does not require interactions related to privacy preservation between the service consumer and the service provider.

- An architecture that enables the inclusion of other QoS attributes keeping the separation of the different attributes as well as supports QoS management without extra impacts on the scalability and performance of the framework.
- An approach that includes a base privacy ontology and domain-specific privacy ontologies in order to take domain-specific privacy preservation issues into consideration as well as supports the application of the framework to different domains.
- The application of the framework to the domain of health care, including the definition of a health care scenario and an ontology based on a privacy preservation regulation for the domain of health care.

This work has produced the following papers:

- Diego Garcia, M. Beatriz F. Toledo. Semantic Policies for Web Services. IC-Unicamp PhD Thesis Workshop. Campinas, Brazil, 2007.
- Diego Garcia, M. Beatriz F. Toledo. A Web Service Privacy Framework Based on a Policy Approach Enhanced with Ontologies. Workshop Web2Touch - Living Experience Through Web, IEEE International Conference on Computational Science and Engineering. São Paulo, Brazil, 2008.
- Diego Garcia, M. Beatriz F. Toledo. Ontology-based Security Policies for Supporting the Management of Web Service Business Processes. IEEE International Conference on Semantic Computing. Santa Clara, USA, 2008.

- Diego Garcia, M. Beatriz F. Toledo. Quality of Service Management for Web Service Compositions. IEEE International Conference on Computational Science and Engineering. Sao Paulo, Brazil, 2008.
- Diego Garcia, M. Beatriz F. Toledo. An Approach for Establishing Trust Relationships in the Web Service Technology. IFIP Working Conference on Virtual Enterprises. Poznan, Poland, 2008.
- Diego Garcia, M. Beatriz F. Toledo. Web Service Security Management Using Semantic Web Techniques. ACM Symposium on Applied Computing. Fortaleza, Brazil, 2008.
- Diego Garcia, M. Beatriz F. Toledo, Paul Grace, Gordon S. Blair, Miriam A. M. Capretz, David S. Allison. Towards Protecting Consumer's Privacy in Service-Oriented Architecture. IEEE Toronto International Conference - Science and Technology for Humanity. Toronto, Canada, 2009.
- Diego Garcia, M. Beatriz F. Toledo, Gordon S. Blair, Paul Grace, Carlos Flores, Miriam A. M. Capretz, David S. Allison. Towards a Base Ontology for Privacy Protection in Service-Oriented Architecture. IEEE International Conference on Service-Oriented Computing and Applications. Taipei, Taiwan, 2009.
- Diego Garcia, Gordon S. Blair, Paul Grace, Carlos Flores, M. Beatriz F. Toledo, Miriam A. M. Capretz. A Configurable Approach to Privacy Ontology and its Application to Mobile e-Health Services. IFIP/PrimeLife International Summer School - Privacy and Identity Management for Life. Nice, France, 2009.

- Hany F. EL Yamany, Miriam A. M. Capretz, David S. Allison, Diego Garcia, M. Beatriz F. Toledo. QoSS Policies within SOA. IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies. Milan, Italy, 2009.
- M. Beatriz F. Toledo, Olga Nabuco, Marcos Rodrigues, Diego Garcia, Miriam A. M. Capretz, Marcelo Fantinato, Itana Gimenes, Rodrigo Bonacin, Ana Guerra, Tarcísio da Rocha, Laura Viana. An SOA-based Collaborative Environment for Clinical Trials on Neglected Diseases. IEEE International Workshop on Service Oriented Architectures in Converging Networked Environments. Bradford, UK, 2009.
- Miriam A. M. Capretz, M. Beatriz F. Toledo, Marcelo Fantinato, Diego Garcia, Shuying Wang, David S. Allison, Olga Nabuco, Marcos Rodrigues, Rodrigo Bonacin, Emma Chen Sasse, Itana Gimenes, Americo Brigido Cunha. Web technologies in a collaborative platform for clinical trials. Electronic journal of communication information and innovation in health. Vol. 3, 2009.
- Diego Garcia, David S. Allison, Miriam A. M. Capretz, M. Beatriz F. Toledo. Privacy Protection Mechanisms for Web Service Technology. ACIS International Conference on Software Engineering Research, Management and Applications. Montreal, Canada, 2010.
- Diego Garcia, Miriam A. M. Capretz, M. Beatriz F. Toledo. Using Contract and Ontology for Privacy Protection in Service-Oriented Architecture. International Conference on Digital Information Management. Thunder Bay, Canada, 2010.

8.3 Future Work

Future work includes developing tools for the specification and publication of semantic privacy policies. The privacy preservation approach proposed in this thesis does not require service providers to create a privacy policy for each service consumer individually. This could be done through negotiation if necessary. However, service providers have to define a privacy policy for each service they offer, which can still be difficult to some providers. As privacy policies usually follow a similar specification, a tool could be provided to facilitate the specification of these policies. For instance, feature modeling could be employed by such tool in order to manage policy commonalities and help in the specialization of a privacy policy to different services. In the case of service consumers, it can be difficult to specify and publish their privacy preservation preferences as it is necessary to understand the privacy ontologies to do so. Again, a tool to guide service consumers through the specification and publication of their policies could be used. Privacy policy templates could be created and the tool would support a service consumer to configure a policy template and generate its privacy policy. Such tool could help the service consumer to understand the different information activities and their privacy preservation implications. Moreover, it would be important to have domain representative organizations for service consumers and service providers defining these templates for each service type in a particular application domain, which would work as default privacy preservation preferences and practices that then could be specialized according to the needs of consumers and providers.

In addition, the privacy preservation approach proposed in this thesis requires service providers to adhere to the practice of specifying semantic privacy policies. Furthermore, the mediator and privacy roles included in SOA must have the capability of using semantic privacy policies for service publication and discovery. Thus, regulatory mechanisms are necessary in order to enforce these behaviors and guarantee that they are unbiased.

Another future work is the inclusion of a protocol for privacy policy negotiation in the proposed framework in order to help a service provider and a service consumer reaching an agreement in the case of incompatible policies.

The inclusion of a mechanism in order to check the correspondence between the semantic privacy policy of a service provider and its actual privacy preservation practices is also necessary. This extension to the framework can involve mechanisms for privacy policy enforcement and a certification solution with the use of trusted third parties in order to deal with issues such as service providers that do not act according to their privacy policies and service providers that obscure the details of their privacy preservation practices in their privacy policies.

Another future work is the inclusion of a mechanism to support dynamic ontologies. This includes the necessity to update the ontologies as a consequence of their use and the need to improve them.

The inclusion of a component for inferring new knowledge from privacy ontologies is necessary in order to take advantage of the full benefits of using ontologies for enhancing service discovery in the proposed framework.

Another future work is to support the personalization of services for consumers with the sets of privacy preservation practices previously selected for the consumers. Related to this work is the need to keep track of the different requested policies in order to support auditing in the proposed framework.

The evaluation of the proposed framework in another domain is another future work. This evaluation would enable a more comprehensive validation of the privacy ontology approach used in the framework.

Another future work is to improve the policy model of the proposed framework with the inclusion of support to the definition of priorities among policy alternatives. The modification of the framework so that it could guarantee more controlled policies as default is also necessary.

A possible future work is the inclusion of an additional functionality in the mediator service, which would make it able to validate policies. This new functionality would enable the verification of the existence of conflicts in policies or among policies. This check could be performed as a step before the send of the policy to the privacy service for compatibility verification.

Finally, a quantitative evaluation of the proposed framework is another possible future work. By comparing the framework with basic SOA, it would be possible to evaluate the overhead of the extension.

Other privacy preservation solutions for SOA proposed in the literature have faced difficulties to reach applicability. These difficulties show that several issues should be addressed in order to guarantee the practical use of the framework proposed in this thesis, including the issues discussed in this section that have not been currently addressed in this work. Thus, the framework proposed in this thesis is an important step towards privacy preservation in the areas of service description and discovery, but other technical and non-technical solutions must be in place together with the proposed framework in order to support its applicability entirely.

Appendix A

Interfaces

This appendix includes the WSDL description of the mediator in Section A.1 and the one of the privacy in Section A.2.

A.1 Mediator

```
<definitions name="Mediator">

  <message name="PublicationAndDiscoveryServiceRegistrationMessage">
    <part name="PublicationAndDiscoveryServiceDescription"/>
  </message>
  <message name="PublicationAndDiscoveryServiceRegistrationResultMessage">
    <part name="PublicationAndDiscoveryServiceRegistrationResult"/>
  </message>
  <message name="PublicationAndDiscoveryServiceDeregistrationMessage">
    <part name="PublicationAndDiscoveryServiceIdentifier"/>
  </message>
  <message name="PublicationAndDiscoveryServiceDeregistrationResultMessage">
    <part name="PublicationAndDiscoveryServiceDeregistrationResult"/>
  </message>
  <message name="ServicePublicationMessage">
    <part name="ServiceDescription"/>
    <part name="PrivacyPolicy"/>
  </message>
  <message name="ServicePublicationFinalResultMessage">
    <part name="ServicePublicationFinalResult"/>
  </message>
  <message name="ServiceUnpublicationMessage">
    <part name="ServiceIdentifier"/>
    <part name="ServiceUnpublicationType"/>
  </message>
  <message name="ServiceUnpublicationFinalResultMessage">
    <part name="ServiceUnpublicationFinalResult"/>
  </message>
  <message name="ServiceDiscoveryMessage">
    <part name="ServiceRequest"/>
    <part name="PrivacyPolicy"/>
  </message>
  <message name="ServiceDiscoveryFinalResultMessage">
    <part name="ServiceIdentifier"/>
    <part name="IntersectionPolicy"/>
  </message>

  <portType name="MediatorPortType">
    <operation name="registerPublicationAndDiscoveryService">
      <input message="PublicationAndDiscoveryServiceRegistrationMessage"/>
      <output
message="PublicationAndDiscoveryServiceRegistrationResultMessage"/>
    </operation>
    <operation name="deregisterPublicationAndDiscoveryService">
      <input message="PublicationAndDiscoveryServiceDeregistrationMessage"/>
      <output
message="PublicationAndDiscoveryServiceDeregistrationResultMessage"/>
    </operation>
    <operation name="publishService">
      <input message="ServicePublicationMessage"/>
      <output message="ServicePublicationFinalResultMessage"/>
    </operation>
    <operation name="unpublishService">
      <input message="ServiceUnpublicationMessage"/>
      <output message="ServiceUnpublicationFinalResultMessage"/>
    </operation>
    <operation name="discoverService">
      <input message="ServiceDiscoveryMessage"/>
      <output message="ServiceDiscoveryFinalResultMessage"/>
    </operation>
  </portType>

  <service name="MediatorWS">
    <port name="MediatorPort" binding="MediatorBinding">
      <soap:address location="http://lh:8080/axis/Mediator.jws"/>
    </port>
  </service>

</definitions>
```

A.2 Privacy

```
<definitions name="Privacy">

  <message name="PrivacyOntologyRegistrationMessage">
    <part name="PrivacyOntology"/>
  </message>
  <message name="PrivacyOntologyRegistrationResultMessage">
    <part name="PrivacyOntologyRegistrationResult"/>
  </message>
  <message name="PrivacyOntologyDeregistrationMessage">
    <part name="PrivacyOntologyIdentifier"/>
  </message>
  <message name="PrivacyOntologyDeregistrationResultMessage">
    <part name="PrivacyOntologyDeregistrationResult"/>
  </message>
```

```
<message name="PrivacyPolicyPublicationMessage">
  <part name="PrivacyPolicy"/>
  <part name="ServiceIdentifier"/>
</message>
<message name="PrivacyPolicyPublicationResultMessage">
  <part name="PrivacyPolicyPublicationResult"/>
</message>
<message name="PrivacyPolicyUnpublicationMessage">
  <part name="PrivacyPolicyIdentifier"/>
</message>
<message name="PrivacyPolicyUnpublicationFinalResultMessage">
  <part name="PrivacyPolicyUnpublicationFinalResult"/>
</message>
<message name="PrivacyPolicyDiscoveryMessage">
  <part name="PrivacyPolicy"/>
</message>
<message name="PrivacyPolicyDiscoveryFinalResultMessage">
  <part name="IntersectionPolicies"/>
  <part name="ServiceIdentifiers"/>
</message>

<portType name="PrivacyPortType">
  <operation name="registerPrivacyOntology">
    <input message="PrivacyOntologyRegistrationMessage"/>
    <output message="PrivacyOntologyRegistrationResultMessage"/>
  </operation>
  <operation name="deregisterPrivacyOntology">
    <input message="PrivacyOntologyDeregistrationMessage"/>
    <output message="PrivacyOntologyDeregistrationResultMessage"/>
  </operation>
  <operation name="publishPrivacyPolicy">
    <input message="PrivacyPolicyPublicationMessage"/>
    <output message="PrivacyPolicyPublicationResultMessage"/>
  </operation>
  <operation name="unpublishPrivacyPolicy">
    <input message="PrivacyPolicyUnpublicationMessage"/>
    <output message="PrivacyPolicyUnpublicationFinalResultMessage"/>
  </operation>
```

```
    <operation name="discoverPrivacyPolicy">
      <input message="PrivacyPolicyDiscoveryMessage"/>
      <output message="PrivacyPolicyDiscoveryFinalResultMessage"/>
    </operation>
  </portType>

  <service name="PrivacyWS">
    <port name="PrivacyPort" binding="PrivacyBinding">
      <soap:address location="http://lh:8080/axis/Privacy.jws"/>
    </port>
  </service>

</definitions>
```

Bibliography

[1] A. Acquisti, A. Friedman, and R. Telang, “Is There a Cost to Privacy Breaches? An Event Study”, in Proceedings of the Workshop on the Economics of Information Security / International Conference on Information Systems (WEIS’06), Milwaukee, USA, December 2006, pp: 26 - 33.

[2] M. Ali, L. Bussard, and U. Pinsdorf, “Obligation Language and Framework to Enable Privacy-Aware SOA”, in Lecture Notes in Computer Science, Vol. 5939, J. Garcia-Alfaro, G. Navarro-Arribas, N. Cuppens-Boulahia, and Y. Roudier (Editors), Springer, 2010, pp: 18 - 32.

[3] D. S. Allison, “Privacy Protection Framework for Service-Oriented Architecture”, Master of Engineering Science Thesis, Department of Electrical and Computer Engineering, University of Western Ontario, London, Canada, 2009.

[4] M. Al-Nedhami and P. K. Sinha, “A Privacy Framework for Composite Web Services”, in Proceedings of the International Workshop on Service-Oriented Engineering and Optimization (SENOPT’08), Bangalore, India, December 2008, paper number: 2.

[5] G. Alonso, F. Casati, H. Kuno, and V. Machiraju, Web Services: Concepts, Architectures and Applications, Springer Verlag, Heidelberg, 2004.

[6] W. F. Boh and D. M. Yellin, “Enablers and Benefits of Implementing Service-Oriented Architecture: An Empirical Investigation”, in International Journal of Information Technology and Management, Vol. 9, No. 1, 2010, pp: 3 - 29.

[7] D. Booth, H. Haas, F. McCabe, E. Newcomer, M. Champion, C. Ferris, and D. Orchard. Web Services Architecture. W3C Working Group Note 11 February 2004. Available at: http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/. Last seen: 26 March 2011.

[8] D. Booth and C. K. Liu (Editors). Web Services Description Language (WSDL) Version 2.0 Part 0: Primer. W3C Recommendation 26 June 2007. Available at: http://www.w3.org/TR/2007/REC-wsdl20-primer-20070626/. Last seen: 26 March 2011.

[9] Canada. Personal Information Protection and Electronic Documents Act. 2000. Available at: http://laws.justice.gc.ca/en/P-8.6/FullText.html/. Last seen: 26 March 2011.

[10] Canadian Mental Health Association. Community Mental Health and Addictions Privacy Toolkit. 2005. Available at: http://www.ontario.cmha.ca/privacytoolkit/docs/privacy_toolkit.pdf/. Last seen: 26 March 2011.

[11] L. Clement, A. Hately, C. von Riegen, and T. Rogers (Editors). UDDI Version 3.0.2. UDDI Spec Technical Committee Draft, Dated 20041019. Available at: http://www.oasis-open.org/committees/uddi-spec/doc/spec/v3/uddi-v3.0.2-20041019.htm/. Last seen: 26 March 2011.

[12] EKOS Research Associates Inc. Canadians and Privacy. March 2009. Available at: http://www.priv.gc.ca/information/survey/2009/ekos_2009_01_e.cfm/. Last seen: 26 March 2011.

[13] EKOS Research Associates Inc. Electronic Health Information and Privacy Survey: What Canadians Think. August 2007. Available at: http://www.hc-sc.gc.ca/ahc-asc/pubs/_atip-aiprp/survey-sondage/index-eng.php/. Last seen: 26 March 2011.

[14] European Union. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data. 1995. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML/. Last seen: 26 March 2011.

[15] D. Fensel, Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce, Springer Verlag, Heidelberg, 2001.

[16] T. R. Gruber, "A Translation Approach to Portable Ontology Specifications", in Knowledge Acquisition, Vol. 5, No. 2, 1993, pp: 199 - 220.

[17] Harris Interactive. Privacy On and Off the Internet: What Consumers Want. 7 February 2002. Available at: http://www.aicpa.org/download/webtrust/priv_rpt_21mar02.pdf/. Last seen: 26 March 2011.

[18] M. N. Huhns and M. P. Singh, "Service-Oriented Computing: Key Concepts and Principles", in IEEE Internet Computing, Vol. 9, No. 1, 2005, pp: 75 - 81.

[19] J. M. Humber, R. F. Almeder (Editors), Privacy and Health Care, Humana Press, Clifton, 2001.

[20] P. C. K. Hung (Editor). Security and Privacy Technologies in SOA. IEEE Computer Society TechSet 2009. Available at: http://www.computer.org/portal/web/buildyourcareer/ts020/. Last seen: 26 March 2011.

[21] P. C. K. Hung, E. Ferrari, and B. Carminati, “Towards Standardized Web Services Privacy Technologies”, in Proceedings of the IEEE International Conference on Web Services (ICWS’04), San Diego, USA, June 2004, pp: 174 - 181.

[22] M. U. Iqbal and S. Lim, “A Survey on Users’ Willingness-to-Pay for Privacy in Mobility Pricing Systems”, in International Journal of Liability and Scientific Enquiry, Vol. 1, No. 3, 2008, pp: 306 - 317.

[23] D. Krafzig, K. Banke, and D. Slama, Enterprise SOA: Service-Oriented Architecture Best Practices, Prentice Hall, Upper Saddle River, 2004.

[24] N. K. Malhotra, S. S. Kim, and J. Agarwal, “Internet Users’ Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model”, in Information Systems Research, Vol. 15, No. 4, 2004, pp: 336 - 355.

[25] H. Meziane and S. Benbernou, “A Dynamic Privacy Model for Web Services”, in Computer Standards & Interfaces, Vol. 32, No. 5-6, 2010, pp: 288 - 304.

[26] N. Mitra and Y. Lafon (Editors). SOAP Version 1.2 Part 0: Primer (Second Edition). W3C Recommendation 27 April 2007. Available at: http://www.w3.org/TR/2007/REC-soap12-part0-20070427/. Last seen: 26 March 2011.

[27] Nadalin, C. Kaler, R. Monzillo, and P. Hallam-Baker (Editors). Web Services Security: SOAP Message Security 1.1 (WS-Security 2004). OASIS Standard Specification 1 February 2006. Available at: http://www.oasis-open.org/committees/download.php/16790/wss-v1.1-spec-os-SOAPMessageSecurity.pdf/. Last seen: 26 March 2011.

[28] Ontario. Personal Health Information Protection Act. 2004. Available at: http://www.e-laws.gov.on.ca/html/statutes/english/elaws_statutes_04p03_e.htm/. Last seen: 26 March 2011.

[29] Organisation for Economic Co-operation and Development. OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. 1980. Available at: http://www.oecd.org/document/18/0,3343,en_2649_34255_1815186_1_1_1_1,00.html/. Last seen: 26 March 2011.

[30] Y. Osawa, S. Imamura, A. Takeda, G. Kitagata, N. Shiratori, and K. Hashimoto, “A Proposal of Privacy Management Architecture”, in Proceedings of the IEEE/IPSJ

International Symposium on Applications and the Internet (SAINT'10), Seoul, Korea, July 2010, pp: 161 - 164.

[31] M. P. Papazoglou, P. Traverso, S. Dustdar, and F. Leymann, "Service-Oriented Computing: A Research Roadmap", in International Journal of Cooperative Information Systems, Vol. 17, No. 2, 2008, pp: 223 - 255.

[32] Pew Research Center's Internet & American Life Project. Use of Cloud Computing Applications and Services. September 2008. Available at: http://www.pewinternet.org/Reports/2008/Use-of-Cloud-Computing-Applications-and-Services.aspx?r=1/. Last seen: 26 March 2011.

[33] M.-T. Schmidt, B. Hutchison, P. Lambros, and R. Phippen, "The Enterprise Service Bus: Making Service-Oriented Architecture Real", in IBM Systems Journal, Vol. 44, No. 4, 2005, pp: 781 - 797.

[34] H. J. Smith, S. J. Milberg, and S. J. Burke, "Information Privacy: Measuring Individuals' Concerns about Organizational Practices", in MIS Quarterly, Vol. 20, No. 2, 1996, pp: 167 - 196.

[35] S. Spiekermann and L. F. Cranor, "Engineering Privacy", in IEEE Transactions on Software Engineering, Vol. 35, No. 1, 2009, pp: 67 - 82.

[36] S. Staab and R. Studer (Editors), Handbook on Ontologies, Springer Verlag, Heidelberg, 2009.

[37] J. Tsai, S. Egelman, L. Cranor, and A. Acquisti, "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study" , in Proceedings of the Workshop on the Economics of Information Security (WEIS'07), Pittsburgh, USA, June 2007, paper number: 57.

[38] Tumer, A. Dogac, and I. H. Toroslu, "A Semantic-Based User Privacy Protection Framework for Web Services", in Lecture Notes in Computer Science, Vol. 3169, B. Mobasher and S. S. Anand (Editors), Springer, 2005, pp: 289 - 305.

[39] J. Turow, J. King, C. J. Hoofnagle, A. Bleakley, and M. Hennessy. Americans Reject Tailored Advertising and Three Activities that Enable It. Social Science Research Network 29 September 2009. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1478214/. Last seen: 26 March 2011.

[40] United States of America. The Privacy Act of 1974. 1974. Available at: http://www.archives.gov/about/laws/privacy-act-1974.html/. Last seen: 26 March 2011.

[41] G. van Heijst, A. T. Schreiber, and B. J. Wielinga, “Using Explicit Ontologies in KBS Development”, in International Journal of Human-Computer Studies, Vol. 46, No. 2-3, 1997, pp: 183 - 292.

[42] S. Vedamuthu, D. Orchard, F. Hirsch, M. Hondo, P. Yendluri, T. Boubez, and Ü. Yalçinalp. Web Services Policy 1.5 - Framework. W3C Recommendation 04 September 2007. Available at: http://www.w3.org/TR/2007/REC-ws-policy-20070904/. Last seen: 26 March 2011.

[43] W3C OWL Working Group. OWL 2 Web Ontology Language Document Overview. W3C Recommendation 27 October 2009. Available at: http://www.w3.org/TR/2009/REC-owl2-overview-20091027/. Last seen: 26 March 2011.

[44] F. F. Wang and N. Griffiths, “Protecting Privacy in Automated Transaction Systems: A Legal and Technological Perspective in the European Union”, in International Review of Law, Computers & Technology, Vol. 24, No. 2, 2010, pp: 153 - 162.

[45] S. D. Warren and L. D. Brandeis, “The Right to Privacy”, in Harvard Law Review, Vol. 4, No. 5, 1890, pp: 193 - 220.

[46] F. Westin, Privacy and Freedom, Atheneum, New York, 1967.

[47] G. O. M. Yee, “A Privacy Controller Approach for Privacy Protection in Web Services”, in Proceedings of the ACM Workshop on Secure Web Services (SWS’07), Fairfax, USA, November 2007, pp: 44 - 51.

VITA

Name:

Diego Zuquim Guimarães Garcia

Post-secondary Education and Degrees:

- 1999-2000. Technician Diploma, Industrial Informatics, Federal Technical College of Sao Paulo, Brazil.

- 2001-2004. BSc, Information Systems, Pontifical Catholic University of Minas Gerais, Brazil.
  Supervisors: Prof. Marco Rodrigo Costa and Prof. Alisson Rabelo Arantes.

- 2005-2007. MSc, Computer Science, State University of Campinas, Brazil.
  Thesis: Inclusion of Quality of Service into the Web Service Model.
  Supervisor: Prof. Maria Beatriz Felgar de Toledo.

- 2007-pres. PhD, Computer Science, State University of Campinas, Brazil and Software Engineering, University of Western Ontario, Canada.
  Supervisors: Prof. Maria Beatriz Felgar de Toledo and Prof. Miriam A. M. Capretz.

Honours and Awards:

- 2001-2004. BSc Academic Distinction Scholarship, Pontifical Catholic University of Minas Gerais, Brazil.

- 2004-2004. Scientific Initiation Scholarship, Scientific Initiation Scholarship Program, Brazil.

- 2005-2007. MSc Scholarship, Ministry of Education, Brazil.

- 2007-2011. PhD Scholarship, Research Support Institution of Sao Paulo, Brazil.

- 2008-2009. PhD International Internship Scholarship, Ministry of Education, Brazil.

Related Work Experience:

- 2001-2003. Teaching Assistant, Institute of Informatics, Pontifical Catholic University of Minas Gerais, Brazil.

- 2005-2005. Teaching Assistant, Institute of Computing, State University of Campinas, Brazil.

- 2007-2008. Lecturer, Institute of Computing, State University of Campinas, Brazil.

- 2008-2009. Research Assistant, Computing Department, Lancaster University, England.

- 2010-2010. Teaching Assistant, Department of Electrical and Computer Engineering, University of Western Ontario, Canada.

Publications:

- Garcia, D. Z. G.; Costa, M. R. “Performance Evaluation and Comparison of Object-Oriented Middleware for Web Application Development”. In: 2nd Latin American Web Congress & 10th Brazilian Symposium on Multimedia and the Web Joint Conference (LA-Web & WebMedia). SBC, Brazil, 2004 (in Portuguese).
- Garcia, D. Z. G.; de Toledo, M. B. F. “A Policy-based Web Service Infrastructure for Autonomic Service Integration”. In: 1st Latin American Autonomic Computing Symposium (LAACS). SBC, Brazil, 2006.
- Garcia, D. Z. G.; de Toledo, M. B. F. “A Web Service Architecture Providing QoS Management”. In: 4th Latin American Web Congress (LA-Web). IEEE, Mexico, 2006.

- Garcia, D. Z. G.; de Toledo, M. B. F. “Semantics-enriched QoS Policies for Web Service Interactions”. In: 12th Brazilian Symposium on Multimedia and the Web (WebMedia). SBC/ACM, Brazil, 2006.
- Garcia, D. Z. G.; de Toledo, M. B. F. “A Fault Tolerant Web Service Architecture”. In: 5th Latin American Web Congress (LA-Web). IEEE, Chile, 2007.
- Garcia, D. Z. G.; de Toledo, M. B. F. “A Fault Tolerant Architecture for Web Service-based Business Processes”. In: 5th IADIS Iberian Conference WWW/Internet (CIAWI). IADIS, Portugal, 2007 (in Portuguese).
- Garcia, D. Z. G.; de Toledo, M. B. F. “A UDDI Extension for Business Process Management Systems”. In: 6th IADIS International Conference WWW/Internet (WWW/Internet). IADIS, Portugal, 2007.
- Garcia, D. Z. G.; de Toledo, M. B. F. “A Policy Approach Supporting Web Service-based Business Processes”. In: 1st Workshop on Business Process Management (WBPM). SBC, Brazil, 2007.
- Garcia, D. Z. G.; de Toledo, M. B. F. “Achieving Autonomic Web Service Integration: A Quality of Service Policy-Based Approach”. In: International Transactions on Systems Science and Applications (ITSSA), Volume 3, Number 1. Xiaglow Institute, UK, 2007.
- Garcia, D. Z. G.; de Toledo, M. B. F. “Semantic Policies for Web Services”. In: IC-Unicamp PhD Thesis Workshop (WTD). IC-Unicamp, Brazil, 2007.
- Garcia, D. Z. G.; de Toledo, M. B. F. “Web Service Security Management Using Semantic Web Techniques”. In: 23rd ACM Symposium on Applied Computing (SAC). ACM, Brazil, 2008.
- Garcia, D. Z. G.; de Toledo, M. B. F. “A Web Service Privacy Framework Based on a Policy Approach Enhanced with Ontologies”. In: 1st Workshop Web2Touch - Living Experience Through Web (Web2Touch). IEEE, Brazil, 2008.
- Garcia, D. Z. G.; de Toledo, M. B. F. “Quality of Service Management for Web Service Compositions”. In: 11th IEEE International Conference on Computational Science and Engineering (CSE). IEEE, Brazil, 2008.
- Garcia, D. Z. G.; de Toledo, M. B. F. “Ontology-based Security Policies for Supporting the Management of Web Service Business Processes”. In: 2nd IEEE International Conference on Semantic Computing (ICSC). IEEE, USA, 2008.
- Garcia, D.; Toledo, M. B. “Trust Management for Supporting Web Service Business Processes”. In: 5th IADIS Intl Conf e-Commerce (e-Commerce). IADIS, Holland, 2008.
- Garcia, D. Z. G.; de Toledo, M. B. F. “An Approach for Establishing Trust Relationships in the Web Service Technology”. In: 9th IFIP Working Conference on Virtual Enterprises (PRO-VE). Springer, Poland, 2008.
- Diego Garcia, M. Beatriz F. Toledo, Paul Grace, Gordon S. Blair, Miriam A. M. Capretz, David S. Allison. “Towards Protecting Consumer’s Privacy in Service-Oriented Architecture”. In: IEEE Toronto International Conference - Science and Technology for Humanity (TIC). IEEE, Canada, 2009.
- M. Beatriz F. Toledo, Marcos Souza, Diego Garcia. “A virtual web environment for researchers on Brazilian indigenous cultural heritage, Indians and general public”. In: IEEE Toronto International Conference - Science and Technology for Humanity (TIC). IEEE, Canada, 2009.

- Diego Garcia, M. Beatriz F. Toledo, Gordon S. Blair, Paul Grace, Carlos Flores, Miriam A. M. Capretz, David S. Allison. “Towards a Base Ontology for Privacy Protection in Service-Oriented Architecture”. In: IEEE International Conference on Service-Oriented Computing and Applications (SOCA). IEEE, Taiwan, 2009.
- Diego Garcia, Gordon S. Blair, Paul Grace, Carlos Flores, M. Beatriz F. Toledo, Miriam A. M. Capretz. “A Configurable Approach to Privacy Ontology and its Application to Mobile e-Health Services”. In: IFIP/PrimeLife International Summer School - Privacy and Identity Management for Life. IFIP/PrimeLife, France, 2009.
- Hany F. EL Yamany, Miriam A. M. Capretz, David S. Allison, Diego Garcia, M. Beatriz F. Toledo. “QoSS Policies within SOA”. In: IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies (WI). IEEE, Italy, 2009.
- M. Beatriz F. Toledo, Olga Nabuco, Marcos Rodrigues, Diego Garcia, Miriam A. M. Capretz, Marcelo Fantinato, Itana Gimenes, Rodrigo Bonacin, Ana Guerra, Tarcísio da Rocha, Laura Viana. “An SOA-based Collaborative Environment for Clinical Trials on Neglected Diseases”. In: IEEE International Workshop on Service Oriented Architectures in Converging Networked Environments (SOCNE). IEEE, UK, 2009.
- Miriam A. M. Capretz, M. Beatriz F. Toledo, Marcelo Fantinato, Diego Garcia, Shuying Wang, David S. Allison, Olga Nabuco, Marcos Rodrigues, Rodrigo Bonacin, Emma Chen Sasse, Itana Gimenes, Americo Brigido Cunha. “Web technologies in a collaborative platform for clinical trials”. In: Electronic journal of communication information and innovation in health (RECIIS). Vol. 3, 2009.
- Diego Garcia, David S. Allison, Miriam A. M. Capretz, M. Beatriz F. Toledo. “Privacy Protection Mechanisms for Web Service Technology”. In: ACIS International Conference on Software Engineering Research, Management and Applications (SERA). IEEE, Canada, 2010.
- Diego Garcia, Miriam A. M. Capretz, M. Beatriz F. Toledo. “Using Contract and Ontology for Privacy Protection in Service-Oriented Architecture”. In: International Conference on Digital Information Management (ICDIM). IEEE, Canada, 2010.